Entropy Coherent and Entropy Convex Measures of Risk

R. Laeven, M. Stadje

Entropy Coherent and Entropy Convex Measures of Risk*

Roger J. A. Laeven

Dept. of Econometrics and Operations Research

Tilburg University, CentER

and EURANDOM

R.J.A.Laeven@uvt.nl

Mitja Stadje

Dept. of Mathematics and Computer Sciences

Technical University of Eindhoven

and EURANDOM

Stadje@eurandom.tue.nl

This Version: April 2, 2010

Abstract

We introduce entropy coherent and entropy convex measures of risk and prove a collection of axiomatic characterization and duality results. We show in particular that entropy coherent and entropy convex measures of risk emerge as negative certainty equivalents in (the regular and a generalized version, respectively, of) the popular maxmin expected utility theory of Gilboa and Schmeidler [12] whenever the negative certainty equivalents are translation invariant. In addition, we derive the dual conjugate function for entropy coherent and entropy convex measures of risk, and prove their distribution invariant representation.

Keywords: Robust preferences; Convex risk measures; Exponential utility; Relative entropy; Translation invariance; Convexity.

AMS 2010 Classification: Primary: 91B06, 91B16, 91B30; Secondary: 60E15, 62P05.

JEL Classification: D81, G10, G20.

1 Introduction

Among the most popular theories for decision-making under uncertainty is the *robust Savage representation*, postulating that an economic agent evaluates the payoff of a choice alternative (financial position) X, defined on a measurable space $(\Omega, \mathcal{F})$, according to

$$U(X) = \inf_{Q \in \mathcal{Q}} \mathrm{E}_Q \left[u(X) \right], \tag{1.1}$$

where $u : \mathbb{R} \to \mathbb{R}$ is an increasing function, and $\mathcal{Q}$ is a set of probability measures on $(\Omega, \mathcal{F})$. The function u, referred to as a utility function, represents the agent's attitude towards wealth,

---

and the set $\mathcal{Q}$ represents the agent's uncertainty about the correct probabilistic model. Gilboa and Schmeidler [12] established a preference axiomatization of the robust Savage representation in an 'enlarged setting', generalizing Savage [21] in the framework of Anscombe and Aumann [1]: They assumed that the payoff in each scenario ω itself can be a lottery, i.e., every payoff corresponds to a stochastic kernel $\tilde{X}(\omega, dx)$. One-stage payoffs X can then be embedded into the space of stochastic kernels by setting $\tilde{X}(\omega, dx) = \delta_{X(\omega)}$, with $\delta_{X(\omega)}$ a Dirac point mass in $X(\omega)$. Now for a given preference order on the space of stochastic kernels satisfying certain axioms, Gilboa and Schmeidler [12] obtained a numerical representation which on the space of one-stage payoffs corresponds to (1.1). The representation of Gilboa and Schmeidler [12], also referred to as maxmin expected utility or multiple priors, was a decision-theoretic foundation of the classical decision rule of Wald [25] –see also Huber [19]– that had long seen little popularity outside (robust) statistics.

To measure the 'risk' related to a financial position X the theory sketched above would lead to the definition of a loss functional $L(X) = -U(X)$, satisfying

$$L(X) = \sup_{Q \in \mathcal{Q}} \mathrm{E}_Q \left[l(-X) \right],$$

where $l(x) := -u(-x)$. The disutility (or loss) function l should be interpreted as describing how much a loss hurts. One could, then, look at the capital amount $\bar{m}_X$ that is 'equivalent' to the potential loss of X, solving for $\bar{m}_X$ in $L(\bar{m}_X) = L(X)$. However, because we want to interpret $\bar{m}_X$ as a certain amount of capital one needs to hold in response to the position X, we will rather look at the *negative certainty equivalent* of X, m_X, given by $-\bar{m}_X$, satisfying $L(-m_X) = l(m_X) = L(X)$, or equivalently,

$$m_X = l^{-1} \left(\sup_{Q \in \mathcal{Q}} \mathrm{E}_Q \left[l(-X) \right] \right). \tag{1.2}$$

In a separate strand of the literature, in financial mathematics, convex risk measures have played an increasingly important role since their introduction by Föllmer and Schied [8], Fritelli and Rosazza Gianin [10] and Heath and Ku [18], generalizing the seminal Artzner et al. [2]; see also the early Deprez and Gerber [6]. For a given financial position X that an economic agent holds, a convex risk measure ρ returns the minimal amount of capital the agent is required to commit and add to the financial position in order to make it 'safe': The theory of convex risk measures postulates that from the viewpoint of the supervisory authority, the financial position $X + \rho(X)$ is acceptably insured against adverse shocks. Convex risk measures are characterized by the axioms of monotonicity, translation invariance and convexity. They can be represented in the form

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{ \mathrm{E}_Q \left[-X \right] - \alpha(Q) \}, \tag{1.3}$$

where α is a penalty function defined on probability measures on $(\Omega, \mathcal{F})$. With

$$\alpha(Q) = \begin{cases} 0, & \text{if } Q \in \mathcal{Q}; \\ \infty, & \text{otherwise}; \end{cases}$$

we obtain the particular subclass of coherent measures of risk, represented in the form

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \mathrm{E}_Q \left[-X \right].$$

2

One of the main goals of this paper is to find precise connections between risk measurement à la Gilboa-Schmeidler in the microeconomic theory of decision under uncertainty –(1.2)– and the notion of convex risk measurement in financial mathematics –(1.3). We will identify a subclass of convex risk measures that we call *entropy coherent measures of risk* and that includes all coherent risk measures. We then show that entropy coherent measures of risk constitute exactly those risk measures that satisfy (1.2) *and* (1.3). To study entropy coherent measures of risk we first study the more general class of *entropy convex measures of risk*, introduced in this paper. We show that these risk measures satisfy similar appealing properties as convex risk measures, the difference being that the expectation operator with respect to a probability measure Q is everywhere replaced by the entropic risk measure with respect to Q, see the definitions below. We axiomatize entropy convex measures of risk: We prove that negative certainty equivalents in the generalized Gilboa-Schmeidler setting, where every probability measure Q is discounted by an additional factor $\beta(Q)$, are translation invariant if and only if they are entropy convex measures of risk. The discount factor $\beta(Q)$ represents the esteemed plausibility of the probabilistic model under Q. Entropy coherent measures of risk occur whenever $\beta(Q) \equiv 1$, which is the case for the regular robust Savage representation. The mathematical details in the proofs of these representation results are delicate.

In the traditional setting of Von Neumann-Morgenstern, where the probability measure is known and given so that simply $U(X) = \mathrm{E}\,[u(X)]$, analogs of these representation results are relatively easy to obtain; see Hardy, Littlewood and Pólya [17] (p. 88, Theorem 106), Gerber [11] (Chapter 5) and Goovaerts, De Vylder and Haezendonck [14] (Chapter 3). It is intriguingly more complicated for the (regular and generalized) robust Savage representation considered here, and we will show that without richness assumptions on the probability space and subdifferentiability conditions on ρ, our representation theorems in fact break down, with interesting counterexamples. In recent work, Cheridito and Kupper [3] suggest without proof a connection between (1.2) and (1.3). They restrict, however, to a specific and simple probabilistic setting which, as we will see below, can be viewed as supplementary (and non-overlapping) to a special case of the general setting considered here. While there is a rich literature on both theories (1.2) and (1.3), to the best of our knowledge, we are not aware of other work establishing precise connections between these two dominant theories.

In addition, we prove various results on the dual conjugate function for entropy coherent and entropy convex measures of risk. We show in particular that, quite exceptionally, the dual conjugate function can explicitly be identified under some technical conditions. We also study entropy coherent and entropy convex measures of risk under the assumption of distribution invariance. Due to their convex nature, a feature that singles out entropy coherent risk measures in the class of negative Gilboa-Schmeidler certainty equivalents, we can obtain explicit representation results in this setting. As a bridge towards the distribution invariant representation, we axiomatize entropy coherent measures of risk within Schmeidler's [23] framework of Choquet expected utility.

The rest of this paper is organized as follows: In Section 2, we review some preliminaries for coherent and convex measures of risk. In Section 3, we introduce entropy coherent and entropy convex measures of risk and discuss some of their properties. In Section 4, we prove axiomatic characterization results for entropy coherent and entropy convex measures of risk. Section 5 studies the dual conjugate function for entropy coherent and entropy convex measures, and Section 6 proves their distribution invariant representation. Conclusions are in Section 7.

2   Preliminaries

We fix a probability space $(\Omega, \mathcal{F}, P)$. Throughout this paper, equalities and inequalities between random variables are understood in the P-almost sure sense. We let $L^\infty(\Omega, \mathcal{F}, P) \equiv L^\infty$ denote the space of all real-valued random variables X on $(\Omega, \mathcal{F}, P)$ for which $||X||_\infty := \inf\{c > 0 | P[|X| \leq c] = 1\} < \infty$, where two random variables are identified if they are P-almost surely equal. We denote $]0, \infty[$ by $\mathbb{R}^+$ and $] - \infty, 0]$ by $\mathbb{R}_0^-$.

Definition 2.1 *We call a mapping $\rho : L^\infty \to \mathbb{R}$ a convex risk measure if it has the following properties:*

- *Normalization: $\rho(0) = 0$*

- *Translation Invariance: $\rho(X + m) = \rho(X) - m$ for all $m \in \mathbb{R}$*

- *Monotonicity: If $X \leq Y$, then $\rho(X) \geq \rho(Y)$*

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

- *Continuity from above: If $X_n \in L^\infty$ is a decreasing sequence converging to $X \in L^\infty$, then*

$$\rho(X_n) \uparrow \rho(X).$$

Furthermore, ρ is called a coherent risk measure if additionally it is positively homogeneous, i.e.,

- *Positive Homogeneity: For $\lambda > 0 : \rho(\lambda X) = \lambda\rho(X)$.*

We denote by $\mathcal{Q}(P) \equiv \mathcal{Q}$ all probability measures that are absolutely continuous with respect to P. If $Q \in \mathcal{Q}$, we also write $Q \ll P$. It is well-known that if ρ is a convex risk measure then there exists a unique lower-semicontinuous and convex function $\alpha : \mathcal{Q} \to \mathbb{R} \cup \{\infty\}$, referred to as the *dual conjugate* of ρ, such that the following dual representation holds:

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \left\{ \mathrm{E}_Q[-X] - \alpha(Q) \right\}. \tag{2.1}$$

Furthermore,

$$\alpha(Q) = \sup_{X \in L^\infty} \left\{ \mathrm{E}_Q[-X] - \rho(X) \right\}; \tag{2.2}$$

α is minimal in the sense that for every other (possibly non-convex or non-lower-semicontinuous) function α' satisfying (2.1), $\alpha \leq \alpha'$; see, for instance, Föllmer and Schied [9]. We define the subdifferential of ρ by

$$\partial\rho(X) = \{Q \in \mathcal{Q} | \rho(X) = \mathrm{E}_Q[-X] - \alpha(Q)\}. \tag{2.3}$$

We say that ρ is subdifferentiable if for every $X \in L^\infty$ we have $\partial\rho(X) \neq \emptyset$. In this paper, we furthermore denote by $C^n(E)$ the space of all functions from $\mathbb{R}$ to $\mathbb{R}$ for which the first n-derivatives exist and which are continuous in an open set E. Finally, for a set $M \subset \mathcal{Q}$, we denote by $\bar{I}_M$ the penalty function that is zero if $Q \in M$ and ∞ otherwise.

4

3 Entropy Coherence and Entropy Convexity

Throughout this section we suppose that $\gamma \in [0, \infty]$ is fixed. A risk measure that is particularly popular in insurance and financial mathematics (Gerber [11], Föllmer and Schied [9] and Mania and Schweizer [20]), macroeconomics (Hansen and Sargent [15, 16]), and decision theory (Gollier [13] and Strzalecki [24]), is the (standard) entropic risk measure defined by

$$e_\gamma(X) = \gamma \log \left(\mathrm{E} \left[\exp \left\{ -\frac{X}{\gamma} \right\} \right] \right).$$

In a setting with distribution invariance, it is commonly referred to as the exponential premium; see Gerber [11]. As is well-known (Csiszár [4]),

$$e_\gamma(X) = \sup_{\bar{P} \ll P} \left\{ \mathrm{E}_{\bar{P}} \left[-X \right] - \gamma H(\bar{P}|P) \right\},$$

where $H(\bar{P}|P)$ is the relative entropy, i.e.,

$$H(\bar{P}|P) = \begin{cases} \mathrm{E}_{\bar{P}} \left[\log \left(\frac{d\bar{P}}{dP} \right) \right], & \text{if } \bar{P} \ll P; \\ \infty, & \text{otherwise.} \end{cases}$$

The relative entropy is also known as the Kullback-Leibler divergence; it measures the distance between the distributions $\bar{P}$ and P.

Risk measurement with the relative entropy is natural in the following setting: The economic agent has a reference measure P; the measure P is, however, an approximation to the probabilistic model of the payoff X rather than the true model. The agent therefore does not fully trust the measure P and considers many measures $\bar{P}$, with esteemed plausibility decreasing proportionally to their distance from the approximation P. Note that for every given X, the mapping $\gamma \to e_\gamma(X)$ is increasing. Consequently, the parameter γ may be viewed as measuring the degree of trust the agent puts in the reference measure P. If $\gamma = 0$, then $e_0(X) = -\operatorname{ess\,inf} X$, which corresponds to a maximal level of distrust; in this case only the zero sets of the measure P are considered reliable. If, on the other hand, $\gamma = \infty$, then $e_\infty(X) = -\mathrm{E}[X]$, which corresponds to a maximal level of trust in the measure P. In the case that $\gamma \in \mathbb{R}^+$, it is well-known that $\partial e_\gamma(X)$ is given by the *Esscher density* with respect to P: $\exp \left\{ -\frac{X}{\gamma} \right\} / \mathrm{E} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right]$.

In certain situations the agent could possibly consider other reference measures $Q \ll P$. Then we define the entropy $e_{\gamma,Q}$ with respect to Q as

$$e_{\gamma,Q}(X) = \gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right).$$

Consider the following example:

Example 3.1 Suppose that the agent is only interested in downside tail risk. The standard risk measure focusing on tail risk is the Tail-Value-at-Risk ($TV@R$), also referred to as Conditional-Value-at-Risk or Average-Value-at-Risk. $TV@R$ is defined by

$$TV@R^\alpha(X) = \frac{1}{\alpha} \int_0^\alpha V@R^\lambda(X) d\lambda, \quad \alpha \in]0, 1],$$

with $V@R^\lambda(X) = -q_X^+(\lambda)$, where q_X^+ is the upper quantile function of X: $q_X^+(\lambda) = \inf\{x|P[X \leq x] > \lambda\}$. If the distribution of X is continuous, $TV@R^\alpha(X) = E\left[-X|X \leq q_X^+(\alpha)\right]$, so that $TV@R$ computes the average over the left tail of the distribution of X up to $q_X^+(\alpha)$. It is well-known that

$$TV@R^\alpha(X) = \sup_{Q \in M_\alpha} E_Q\left[-X\right],$$

where M_α is the set of all probability measures $Q \ll P$ such that $\frac{dQ}{dP} \leq \frac{1}{\alpha}$. Let $\frac{dQ}{dP} = \frac{1}{\alpha}I_{\{X<q_X^+(\alpha)\}} + cI_{\{X=q_X^+(\alpha)\}}$, where c should be chosen such that $E\left[\frac{dQ}{dP}\right] = 1$. Then one can show that

$$Q \in \arg\max\{E_{\bar{P}}\left[-X|\bar{P} \in M_\alpha\right]\},$$

i.e., $TV@R^\alpha(X) = E_Q\left[-X\right]$, and, for continuous distributions, $Q = P[.|X \leq q_X^+(\alpha)]$. Thus, the measure Q coincides with the original reference measure P, but concentrated on the left tail of X. The economic agent may, however, not fully trust the probabilistic model of X under P, hence under Q. Therefore, for every fixed Q, the agent considers the supremum over all measures absolutely continuous with respect to Q, where measures that are 'close' to Q are esteemed more plausible than measures that are 'distant' from Q. This leads to a risk measure ρ given by

$$\rho(X) = \sup_{\bar{P} \ll Q}\sup_{Q \in M_\alpha}\{E_{\bar{P}}\left[-X\right] - \gamma H(\bar{P}|Q)\} = \sup_{\bar{P} \ll P}\sup_{Q \in M_\alpha}\{E_{\bar{P}}\left[-X\right] - \gamma H(\bar{P}|Q)\}$$

$$= \sup_{Q \in M_\alpha}\sup_{\bar{P} \ll P}\{E_{\bar{P}}\left[-X\right] - \gamma H(\bar{P}|Q)\} = \sup_{Q \in M_\alpha} e_{\gamma,Q}(X),$$

where we have used in the second and last equalities that $H(\bar{P}|Q) = \infty$ if $\bar{P}$ is not absolutely continuous with respect to Q.

This example motivates the following definition:

Definition 3.2 *We call a mapping $\rho : L^\infty \to \mathbb{R}$ γ-entropy coherent, $\gamma \in [0,\infty]$, if there exists a set $M \subset \mathcal{Q}$ such that*

$$\rho(X) = \sup_{Q \in M} e_{\gamma,Q}(X).$$

It will be interesting to consider as well a more general class of risk measures:

Definition 3.3 *The mapping $\rho : L^\infty \to \mathbb{R}$ is γ-entropy convex, $\gamma \in [0,\infty]$, if there exists a penalty function $c : \mathcal{Q} \to [0,\infty]$ with $\inf_{Q \in \mathcal{Q}} c(Q) = 0$, such that*

$$\rho(X) = \sup_{Q \in \mathcal{Q}}\{e_{\gamma,Q}(X) - c(Q)\}. \tag{3.1}$$

Henceforth, we call a mapping entropy coherent (convex) if there exists a $\gamma \in [0,\infty]$ such that ρ is γ-entropy coherent (convex).

Considering

$$-\rho(X) = \inf_{Q \in \mathcal{Q}}\left\{-\gamma \log\left(E_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) + c(Q)\right\},$$

the definition of entropy convexity (whence the special case of entropy coherence as well) can also be motivated using microeconomic theory, as follows: An economic agent with a CARA (exponential) utility function $u(x) = 1 - e^{-\frac{x}{\gamma}}$ computes the certainty equivalent to the

6

payoff X with respect to the reference measure P. The agent is, however, uncertain about the probabilistic model under the reference measure, and therefore takes the infimum over all probability measures Q absolutely continuous with respect to P, where the penalty function $c(Q)$ represents the esteemed plausibility of the probabilistic model under Q. The certainty equivalent thus computed is precisely $-\rho(X)$.

Proposition 3.4 *Every γ-entropy convex functional is a convex risk measure.*

Proof. For every fixed Q with $Q \ll P$ we have that if $X = Y$ P-a.s. then also $X = Y$ Q-a.s., hence, $e_{\gamma,Q}(X) = e_{\gamma,Q}(Y)$ and therefore

$$\sup_{Q \in \mathcal{Q}} \{e_{\gamma,Q}(X) - c(Q)\} = \sup_{Q \in \mathcal{Q}} \{e_{\gamma,Q}(Y) - c(Q)\},$$

as well. Furthermore, $e_{\gamma,Q}(X) - c(Q)$ is translation invariant, monotone, convex and lower-semicontinuous (hence, continuous from above). Thus, also $\sup_{Q \in \mathcal{Q}} \{e_{\gamma,Q}(X) - c(Q)\}$ is translation invariant, monotone, convex and continuous from above. Normalization follows because $\inf_{Q \in \mathcal{Q}} c(Q) = 0$ by assumption. □

As $e_{\infty,Q}(X) = \mathrm{E}_Q[-X]$, (2.1) implies that ρ is a convex risk measure if and only if it is ∞-entropy convex. As we will see later (for example, Theorem 5.2 below), however, with $\gamma < \infty$, not every convex risk measure is γ-entropy convex. This is important: In Theorem 4.1 below we will see that, under some technical conditions, negative certainty equivalents in a generalized Gilboa-Schmeidler setting are translation invariant if and only if they are γ-entropy convex with $\gamma \in \mathbb{R}^+$ or ∞-entropy coherent, ruling out the general ∞-entropy convex case. But the following result is available:

Proposition 3.5 *Let ρ be a convex risk measure. Then for every $\gamma \in [0, \infty]$ there exists a γ-entropy convex risk measure $\rho_{\gamma,dom}$ dominating ρ.*

Proof. We have

$$e_{\gamma,Q}(X) = \sup_{\bar{P} \ll Q} \{\mathrm{E}_{\bar{P}}[-X] - \gamma H(\bar{P}|Q)\} \geq \mathrm{E}_Q[-X].$$

Thus, setting $\alpha = c$,

$$\rho(X) = \sup_{Q \ll P} \{\mathrm{E}_Q[-X] - \alpha(Q)\} \leq \sup_{Q \ll P} \{e_{\gamma,Q}(X) - c(Q)\} = \rho_{\gamma,dom}(X).$$

□

For a risk measure ρ we define

$$\rho^*(Q) = \sup_{X \in L^\infty} \{e_{\gamma,Q}(X) - \rho(X)\}$$

and

$$\rho^{**}(X) = \sup_{Q \ll P} \{e_{\gamma,Q}(X) - \rho^*(Q)\}.$$

Lemma 3.6 *If ρ is γ-entropy convex, then for every $X \in L^\infty$,*

$$\rho^{**}(X) \leq \rho(X). \tag{3.2}$$

Proof. As $\rho^*(Q) = \sup_{X \in L^\infty}\{e_{\gamma,Q}(X) - \rho(X)\}$ it follows that $e_{\gamma,Q}(X) - \rho^*(Q) \leq \rho(X)$ for all $X \in L^\infty$. Taking the supremum over all measures Q which are absolutely continuous with respect to P yields (3.2). $\qquad\square$

The next theorem establishes a duality result for γ-entropy convex risk measures:

Theorem 3.7 *A normalized mapping ρ is γ-entropy convex if and only if $\rho^{**} = \rho$. Furthermore, ρ^* is the minimal penalty function.*

Proof. The 'if' part holds because if $\rho(X) = \rho^{**}(X) = \sup_{Q \ll P}\{e_{\gamma,Q}(X) - \rho^*(Q)\}$ then by virtue of the equalities

$$0 = -\rho(0) = -\sup_{Q \in \mathcal{Q}} -\rho^*(Q) = \inf_{Q \in \mathcal{Q}} \rho^*(Q),$$

ρ is γ-entropy convex. Let us prove the 'only if' direction. We already know from Lemma 3.6 that $\rho^{**} \leq \rho$. We will prove that $\rho^{**} \geq \rho$. If ρ is γ-entropy convex there exists a penalty function c such that

$$\rho(X) = \sup_{Q \in \mathcal{Q}}\{e_{\gamma,Q}(X) - c(Q)\}.$$

Thus, for every $Q \ll P$ we have $c(Q) \geq e_{\gamma,Q}(X) - \rho(X)$. By the definition of ρ^* this yields $c(Q) \geq \rho^*(Q)$. This proves that every penalty function ρ is dominating ρ^*. Moreover,

$$\rho^{**}(X) = \sup_{Q \ll P}\{e_{\gamma,Q}(X) - \rho^*(Q)\} \geq \sup_{Q \ll P}\{e_{\gamma,Q}(X) - c(Q)\} = \rho(X).$$

$\qquad\square$

Theorem 3.7 suggests a way to find out whether a risk measure ρ is γ-entropy convex: compute ρ^* and ρ^{**}, and verify whether $\rho^{**} = \rho$.

Remark 3.8 ρ^* measures how much ρ deviates from below from the Q-entropy. If there exists a $Q \ll P$ such that $\rho(X) \leq e_{\gamma,Q}(X)$ then $\rho^*(Q) \geq e_{\gamma,Q}(X) - \rho(X) \geq 0$. This and the convexity of ρ^* jointly imply that ρ is entropy coherent if and only if $\rho^* = \bar{I}_M$ for a set $M \subset \mathcal{Q}$.

Remark 3.9 Let $\mathcal{A}$ be the acceptance set of ρ, i.e., $\mathcal{A} = \{X \in L^\infty | \rho(X) \leq 0\}$. ρ^* can be represented as

$$\rho^*(Q) = \sup_{X \in \mathcal{A}} e_{\gamma,Q}(X).$$

To see this, note that clearly,

$$\rho^*(Q) = \sup_{X \in L^\infty}\{e_{\gamma,Q}(X) - \rho(X)\} \geq \sup_{X \in \mathcal{A}}\{e_{\gamma,Q}(X) - \rho(X)\} \geq \sup_{X \in \mathcal{A}} e_{\gamma,Q}(X).$$

On the other hand, if $X \in \mathcal{A}$ then $X + \rho(X) \in \mathcal{A}$, which implies that

$$\rho^*(Q) = \sup_{X \in L^\infty}\{e_{\gamma,Q}(X) - \rho(X)\} = \sup_{X \in L^\infty}\{e_{\gamma,Q}(X + \rho(X))\}$$

$$= \sup_{X + \rho(X) = Y \in L^\infty} e_{\gamma,Q}(Y) \leq \sup_{Y \in \mathcal{A}} e_{\gamma,Q}(Y).$$

Definition 3.10 *For a γ-entropy convex function ρ we denote by*

$$\partial_{entropy}\rho(X) = \{Q^* \in \mathcal{Q} | \rho(X) = e_{\gamma,Q^*}(X) - c(Q^*)\}$$

the entropy subdifferential. Furthermore, if for every $X \in L^\infty$, $\partial_{entropy}\rho(X) \neq \emptyset$, then we say that ρ is entropy subdifferentiable.

Remark 3.11 If $\gamma \in \mathbb{R}^+$ and $Q^* \in \partial_{entropy}\rho(X)$, then $\dfrac{\exp\left\{-\frac{X}{\gamma}\right\}}{\mathrm{E}_{Q^*}\left[\exp\left\{-\frac{X}{\gamma}\right\}\right]} \in \partial\rho(X)$, where $\partial\rho(X)$ is the usual subdifferential defined by (2.3). In the case that there exists a c such that (3.1) holds and such that the domain of c is a separated compact space it follows directly from Theorem 2.4.18, Zalinscu [26] that every $\bar{P}$ in $\partial\rho(X)$ can be written as the L^1 limit of convex combinations of measures $\bar{P}^n$ given by $\dfrac{d\bar{P}^n}{dP} = \dfrac{\exp\left\{-\frac{X}{\gamma}\right\}}{\mathrm{E}_{Q_n^*}\left[\exp\left\{-\frac{X}{\gamma}\right\}\right]}$ with $Q_n^* \in \partial_{entropy}\rho(X)$. In particular, in this case $\partial_{entropy}\rho(X) \neq \emptyset$ if and only if $\partial\rho(X) \neq \emptyset$.

Proposition 3.12 *Suppose that ρ is a γ-entropy coherent risk measure with $\gamma \in\,]0, \infty]$. Then the following statements are equivalent:*

(a) *For every $X \in L^\infty$,*

$$\rho(X) = \max_{Q \in M} e_{\gamma,Q}(X).$$

(b) *M is weakly compact.*

(c) *ρ is continuous from below, i.e., $X_n \uparrow X \implies \rho(X_n) \downarrow \rho(X)$.*

Proof. Let

$$\bar{\rho}(X) = \sup_{Q \in M} \mathrm{E}_Q\left[-X\right]. \tag{3.3}$$

First of all, notice that by Corollary 4.35 in Föllmer and Schied [9] and the translation invariance of $\bar{\rho}$, M being weakly compact is equivalent to the maximum in (3.3) being attained for every $X < 0$.

(a)$\Rightarrow$(b): Suppose that $X < 0$. Then

$$\bar{\rho}(X) = \exp\left\{\frac{1}{\gamma}\rho(-\gamma\log(-X))\right\} = \exp\left\{\frac{1}{\gamma}\max_{Q \in M}\gamma\log(\mathrm{E}_Q\left[-X\right])\right\} = \max_{Q \in M}\mathrm{E}_Q\left[-X\right].$$

(b)$\Rightarrow$(a): We write

$$\rho(X) = \gamma\log\left(\sup_{Q \in M}\mathrm{E}_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) = \gamma\log\left(\max_{Q \in M}\mathrm{E}_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) = \max_{Q \in M}e_{\gamma,Q}(X).$$

(b)$\Leftrightarrow$(c): Corollary 4.35 in Föllmer and Schied [9] implies also that M being weakly compact is equivalent to $\bar{\rho}$ being continuous from below. Now clearly $\bar{\rho}$ being continuous from below implies that ρ is continuous from below. On the other hand, suppose that $X_n \uparrow X$. Since $\bar{\rho}$ is translation invariant we may assume without loss of generality that $X_n < 0$. Define $Y_n := -\gamma\log(-X_n) \uparrow Y =: -\gamma\log(-X)$. Then the continuity from below of ρ implies that

$$\bar{\rho}(X_n) = \exp\left\{\frac{\rho(Y_n)}{\gamma}\right\} \downarrow \exp\left\{\frac{\rho(Y)}{\gamma}\right\} = \bar{\rho}(X).$$

$\square$

4 Axiomatic Characterizations

In this section, we axiomatize entropy convex and entropy coherent measures of risk. A key role in this section is played by the axiom of translation invariance.

4.1 Entropy Convexity

We state the following theorem:

Theorem 4.1 *Suppose that the probability space is sufficiently rich to support a random variable with a uniform distribution, and that $\bar{\rho} : L^\infty \to \mathbb{R}$ is monotone, convex, positively homogeneous and continuous from above and for all $m \in \mathbb{R}_0^-$, $\bar{\rho}(m) = -m$. Let ϕ be a strictly increasing and continuous function satisfying $0 \in closure(Image(\phi))$, $\phi(\infty) = \infty$ and $\phi \in C^3(]\phi^{-1}(0), \infty[)$. Then the following statements are equivalent:*

(i) *$\rho(X) = \phi^{-1}(\bar{\rho}(-\phi(-X)))$ is translation invariant and the subdifferential of $\bar{\rho}$ is always nonempty.*

(ii) *ρ is γ-entropy convex with $\gamma \in \mathbb{R}^+$ or ρ is ∞-entropy coherent, and the entropy subdifferential is always nonempty.*

Remark 4.2 The direction (i)$\Rightarrow$(ii) in Theorem 4.1 does not hold (even not in the case that we additionally assume that $\bar{\rho}$ is translation invariant as in Corollary 4.10 below) if the probability space is not rich, or if the assumption on the subdifferential of $\bar{\rho}$ is omitted.

Suppose, for instance, that $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ and that, without loss of generality, $P[\{\omega_i\}] = p_i > 0$, $i = 1, \ldots, n$. Then for a payoff X we can define $\bar{\rho}(X) = \max_{Q \ll P} \mathrm{E}_Q[-X] = \max_{i=1,\ldots,n} -X(\omega_i)$, where the maximum is attained in the measure Q that sets $Q[\{\omega_{i_0}\}] = 1$, where $\omega_{i_0} = \arg\max_\omega -X(\omega)$. Let ϕ be a strictly increasing and continuous function. Then it *always* holds that

$$\phi^{-1}(\bar{\rho}(-\phi(-X))) = \phi^{-1}(\max_i \phi(-X(\omega_i)))$$
$$= \phi^{-1}(\phi(-X(\omega_{i_0}))) = -X(\omega_{i_0}) = \bar{\rho}(X).$$

In particular, $\rho(X) = \phi^{-1}(\bar{\rho}(-\phi(-X))) = \bar{\rho}(X)$ is translation invariant for every function ϕ that is strictly increasing and continuous. This shows that (i)$\Rightarrow$(ii) in Theorem 4.1 does not hold if the probability space is finite.

If, on the other hand, the probability space is rich but we omit the assumption that $\bar{\rho}$ is subdifferentiable, then the coherent risk measure $\bar{\rho}(X) = \mathrm{ess\,sup} -X$ satisfies for *every* strictly increasing and continuous function ϕ that $\rho(X) = \phi^{-1}(\bar{\rho}(-\phi(-X))) = \bar{\rho}(X)$ is a convex risk measure. The equality may be seen to hold as

$$\phi^{-1}(\bar{\rho}(-\phi(-X))) = \mathrm{ess\,sup}\, \phi^{-1}(\phi(-X)) = \mathrm{ess\,sup} -X = \bar{\rho}(X).$$

Remark 4.3 In the proof of Theorem 4.1 we will see that ρ is entropy coherent if and only if $\bar{\rho}$ is a coherent risk measure. In this case, $\bar{\rho}(X) = \sup_{Q \in M} \mathrm{E}_Q[-X]$ for a set $M \subset \mathcal{Q}$, and ρ is a negative certainty equivalent in the Gilboa-Schmeidler framework.

Furthermore, the case that ρ is entropy convex corresponds to ρ being the negative certainty equivalent of $\bar{\rho}(X) = \sup_{Q \in M} \beta(Q)\mathrm{E}_Q[-X]$, where $\beta : M \to [0, 1]$ can be viewed as a discount

factor. In this case, every model Q is discounted by a factor $\beta(Q)$ corresponding to its esteemed plausibility. If $\beta(Q) = 1$ for all $Q \in M$, we are back in the framework of Gilboa-Schmeidler. However, if there exists a $Q \in M$ such that $\beta(Q) < 1$, we will see that ρ is entropy convex with $\gamma \in \mathbb{R}^+$ but not entropy coherent.

Remark 4.4 In financial mathematics, translation invariance is typically motivated by the interpretation of a risk measure on L^∞ as a minimal amount of risk capital. It ensures that $\rho(X + \rho(X)) = 0$.

Remark 4.5 Notice that since ϕ is positive somewhere and $0 \in closure(Image(\phi))$ we have that $\phi^{-1}(\delta)$ is well-defined for all $\delta > 0$ small enough and we can define $\phi^{-1}(0) = \lim_{\delta \downarrow 0} \phi^{-1}(\delta)$. The common condition that $\phi(\infty) = \infty$ implies that ρ remains loss sensitive.

Before proving Theorem 4.1, we first present the following two lemmas:

Lemma 4.6 *Suppose that $\bar{\rho} : L^\infty \to \mathbb{R}$ is monotone, convex, positively homogeneous and continuous from above and for all $m \in \mathbb{R}_0^-$, $\bar{\rho}(m) = -m$. Then there exists a function $\beta : \mathcal{Q} \supset M \to [0, 1]$ with $\sup_{Q \in M} \beta(Q) = 1$, such that for all $X \in L^\infty$ with $X \leq 0$,*

$$\bar{\rho}(X) = \sup_{Q \in M} \beta(Q) \mathrm{E}_Q [-X]. \tag{4.1}$$

Furthermore, if additionally we have $\bar{\rho}(1) = -1$ then M can be chosen such that $\beta(Q) = 1$ for all $Q \in M$.

Proof. By standard arguments (see, for example, Lemma A64 in the appendix of Föllmer and Schied [9]), we may conclude that $\bar{\rho}$ is weak* lower-semicontinuous. Proposition 3.1.2 in Dana [5] implies that

$$\bar{\rho}(X) = \sup_{X' \in L_+^1} \{ \mathrm{E} \left[-X'X \right] - \hat{\bar{\rho}}(X') \},$$

and it follows from standard results in convex analysis that the positive homogeneity of $\bar{\rho}$ entails that $\hat{\bar{\rho}}$ is an indicator function of a convex nonempty set, say $H \subset L_+^1$. Hence,

$$\bar{\rho}(X) = \sup_{X' \in H} \mathrm{E} \left[-X'X \right]$$

$$= \sup_{X' \in H} \mathrm{E} \left[X' \right] \mathrm{E} \left[-\frac{X'}{\mathrm{E} \left[X' \right]} X \right] = \sup_{X' \in H} \mathrm{E} \left[X' \right] \mathrm{E}_{Q^{X'}} [-X], \tag{4.2}$$

where in the case that $X' \equiv 0$, we set $0/0 = 1$ and $Q^{X'} = P$. Now set $M = \Big\{ Q \in \mathcal{Q} \Big|$ there exists a $\lambda \geq 0$ such that $\lambda \frac{dQ}{dP} \in H \Big\}$. Then (4.2) entails that for all $X \in L^\infty$ with $X \leq 0$,

$$\bar{\rho}(X) = \sup_{Q \in M} \beta(Q) \mathrm{E}_Q [-X],$$

where for $Q \in M$, $\beta(Q) = \sup \{ \lambda \geq 0 | \lambda \frac{dQ}{dP} \in H \}$. This shows (4.1). Furthermore,

$$\sup_{Q \in M} \beta(Q) = \bar{\rho}(-1) = 1.$$

To see the last part of the lemma note that if $\bar{\rho}(1) = -1$ then we must have $-1 = \bar{\rho}(1) = \sup_{X' \in H} E[-X']$. This implies that

$$\inf_{X' \in H} E[X'] = 1.$$

On the other hand, since $\bar{\rho}(-1) = 1$, we also have that $\sup_{X' \in H} E[X'] = 1$. Hence, for every $X' \in H$ we get that $E[X'] = 1$ and by the definition of β we obtain that $\beta(Q) = 1$ for every $Q \in M$. $\qquad \square$

Subsequently, we will identify the measure $\beta(Q)Q$ (given by $(\beta(Q)Q)(A) = \beta(Q)Q(A)$ for every $A \in \mathcal{F}$) with its density $\beta(Q)\frac{dQ}{dP}$. We recall that an element $X' \in H \subset L_+^1$ is in $\partial\bar{\rho}(X)$ if it attains the supremum in (4.2), i.e., $\bar{\rho}(X) = E[-X'X]$.

Lemma 4.7 *Suppose that $\bar{\rho} : L^\infty \to \mathbb{R}$ is monotone, convex, positively homogeneous and continuous from above and for all $m \in \mathbb{R}_0^-$, $\bar{\rho}(m) = -m$. Let $X \in L^\infty$ with $X > 0$. Then for every Q with $\beta(Q)Q \in \partial\bar{\rho}(-X)$ we have that*

$$\beta(Q) \geq \frac{\operatorname{ess\,inf} X}{\operatorname{ess\,sup} X}.$$

Proof. Choose $Q \in M$ such that $\beta(Q)Q \in \partial\bar{\rho}(-X)$. Then by (4.1) and the monotonicity of $\bar{\rho}$

$$\operatorname{ess\,inf} X = \bar{\rho}(-\operatorname{ess\,inf} X) \leq \bar{\rho}(-X) = \beta(Q)E_Q[X] \leq \beta(Q)\operatorname{ess\,sup} X,$$

where the last inequality holds as $\beta(Q) \geq 0$. Dividing both sides by $\operatorname{ess\,sup} X$ completes the proof. $\qquad \square$

Proof of Theorem 4.1. (i)$\Rightarrow$(ii):
Since ϕ is positive somewhere and $0 \in closure(Image(\phi))$, there are two cases:

(H1) There exists an x_0 such that $\phi(x_0) = 0$.

(H2) $\lim_{x \to -\infty} \phi(x) = 0$ and for every $x \in \mathbb{R}$ we have $\phi(x) > 0$.

Let $\phi_z(\cdot) := \phi(\cdot + z)$ for $z \in \mathbb{R}$. By translation invariance,

$$\phi_z^{-1}(\bar{\rho}(-\phi_z(-X))) = \phi^{-1}(\bar{\rho}(-\phi_z(-X))) - z = \phi^{-1}(\bar{\rho}(-\phi(-X))).$$

Thus, by considering ϕ_z instead of ϕ, we may assume without loss of generality that:

- If (H1) holds then $\phi(0) = 0$ and $\phi \in C^3(]\phi^{-1}(0), \infty[) = C^3(\mathbb{R}^+)$.

- If (H2) holds then $\phi(0) > 0$ and $\phi \in C^3(]\phi^{-1}(0), \infty[) = C^3(]-\infty, \infty[)$.

In particular, we may always assume that $\phi^{-1}(0) \in \{-\infty, 0\}$ and

$$\phi(0) \geq 0. \tag{4.3}$$

Next, let us look at $X \in L^\infty$ such that $X < 0$. By assumption, $\partial\bar{\rho}(-\phi(-X)) \neq \emptyset$. As $-\phi(-X) < 0$ (since $\phi(0) \geq 0$ and ϕ is strictly increasing), by (4.1) and the assumption that the subdifferential of $\bar{\rho}$ is always nonempty we have that

$$\bar{\rho}(-\phi(-X)) = \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)E_Q[\phi(-X)]. \tag{4.4}$$

Now we need the following proposition:

Proposition 4.8 *Let $X \in L^\infty$ with $X < 0$. Under the assumptions of Theorem 4.1 (i) we have that*

$$\phi' \circ \phi^{-1}\left(\max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi(-X)\right] \right) = \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right]. \quad (4.5)$$

Proof. Note that as ϕ is in $C^3(]\phi^{-1}(0), \infty[)$ we have for $|m| < \operatorname{ess\,inf} -X$,

$$\bar{\rho}(-\phi(-X + m)) = \bar{\rho}(-\phi(-X) - \phi'(-X)m + O(m^2)).$$

As a result, we will find that

$$\lim_{m \to 0} \frac{\bar{\rho}(-\phi(-X + m)) - \bar{\rho}(-\phi(-X))}{m} = \lim_{m \to 0} \frac{\bar{\rho}(-\phi(-X) - \phi'(-X)m + O(m^2)) - \bar{\rho}(-\phi(-X))}{m}$$

$$= \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right]. \quad (4.6)$$

That the last equality holds is seen as follows: For arbitrary $\epsilon > 0$ we have for small m that $\left|\frac{O(m^2)}{m}\right| \le \epsilon$. Therefore,

$$\limsup_{m \downarrow 0} \frac{\bar{\rho}(-\phi(-X) - \phi'(-X)m + O(m^2)) - \bar{\rho}(-\phi(-X))}{m}$$

$$\le \limsup_{m \downarrow 0} \frac{\bar{\rho}(-\phi(-X) - (\phi'(-X) + \epsilon)m) - \bar{\rho}(-\phi(-X))}{m} = \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X) + \epsilon\right],$$

where the inequality holds by the monotonicity of $\bar{\rho}$ while the equality holds by Theorem 2.4.9 Zalinescu [26]. As ϵ can be chosen to be arbitrary small we find that

$$\limsup_{m \downarrow 0} \frac{\bar{\rho}(-\phi(-X) - \phi'(-X)m + O(m^2)) - \bar{\rho}(-\phi(-X))}{m} \le \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right].$$

Similarly, one can prove (with ϵ replaced by $-\epsilon$) that the same inequality holds when $\limsup_{m \downarrow 0}$ on the left-hand side is replaced by $\limsup_{m \uparrow 0}$. It means that

$$\limsup_{m \to 0} \frac{\bar{\rho}(-\phi(-X) - \phi'(-X)m + O(m^2)) - \bar{\rho}(-\phi(-X))}{m} \le \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right].$$

The reverse inequality

$$\liminf_{m \to 0} \frac{\bar{\rho}(-\phi(-X) - \phi'(-X)m + O(m^2)) - \bar{\rho}(-\phi(-X))}{m} \ge \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right]$$

is proven analogously. Hence, indeed (4.6) holds. In particular, the mapping $g(m) = \bar{\rho}(-\phi(-X + m))$ is differentiable in $m = 0$ and

$$g'(0) = \max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right]. \quad (4.7)$$

Now by assumption, $\phi^{-1}(\bar{\rho}(-\phi(-X)))$ is translation invariant and for all $m \in \mathbb{R}$,

$$\frac{\phi^{-1}(\bar{\rho}(-\phi(-X + m))) - \phi^{-1}(\bar{\rho}(-\phi(-X)))}{m} = 1. \quad (4.8)$$

13

Letting m converge to zero in (4.8) we get that

$$(\phi^{-1} \circ g)'(0) = 1. \tag{4.9}$$

On the other hand, applying the chain rule to $\phi^{-1} \circ g$, we obtain

$$(\phi^{-1} \circ g)'(0) = \frac{\partial\left[\phi^{-1}(\bar{\rho}(-\phi(-X+m)))\right]}{\partial m}\bigg|_{m=0} = \frac{g'(m)}{\phi' \circ \phi^{-1}\left(\bar{\rho}(-\phi(-X+m))\right)}\bigg|_{m=0}$$

$$= \frac{\max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right]}{\phi' \circ \phi^{-1}\left(\bar{\rho}(-\phi(-X))\right)}$$

$$= \frac{\max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right]}{\phi' \circ \phi^{-1}\left(\max_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi(-X)\right]\right)}, \tag{4.10}$$

where we applied (4.7) in the third and (4.4) in the last equality. Finally, (4.9) together with (4.10) entail that (4.5) holds true. $\square$

Continuation of the Proof of Theorem 4.1. (i)$\Rightarrow$(ii):
Next, we will show that Proposition 4.8 implies that there exists p, γ, q such that, for all $x \in]\phi^{-1}(0), \infty[$, $\phi(x) = p \exp\{\frac{x}{\gamma}\} + q$ or $\phi(x) = px + q$. We state the following lemma:

Lemma 4.9 *In the setting of Theorem 4.1, suppose that there does not exist p, γ, q such that, for all $x \in]\phi^{-1}(0), \infty[$, $\phi(x) = p \exp\{\frac{x}{\gamma}\} + q$ or $\phi(x) = px + q$. Then the function $\phi' \circ \phi^{-1}$ is not linear on $\phi(]\phi^{-1}(0), \infty[) = \mathbb{R}^+$.*

Proof. Suppose that there exists c, d such that $\phi' \circ \phi^{-1}(x) = cx + d$ for all $x \in \mathbb{R}^+$. As $\phi' \circ \phi^{-1} = \frac{1}{(\phi^{-1})'}$ we get that

$$(\phi^{-1})'(x) = \frac{1}{cx + d}.$$

If $c = 0$ then ϕ is linear on $]\phi^{-1}(0), \infty[$ contrary to our assumptions. As ϕ^{-1} is strictly increasing on $\mathbb{R}^+$, we must have that $c > 0$. This entails $\phi^{-1}(x) = \frac{1}{c} \log(cx + d)$, which yields that $\phi(x) = \frac{1}{c} \exp\{cx\} - \frac{d}{c}$ on $]\phi^{-1}(0), \infty[$. This contradicts again our assumptions. Hence, under the stated assumptions, $\phi' \circ \phi^{-1}$ is not linear on $\mathbb{R}^+$. $\square$

Continuation of the Proof of Theorem 4.1. (i)$\Rightarrow$(ii):
We will now assume that there does not exist p, γ, q such that, for all $x \in]\phi^{-1}(0), \infty[$, $\phi(x) = p \exp\{\frac{x}{\gamma}\} + q$ or $\phi(x) = px + q$, and prove that we obtain a contradiction to Proposition 4.8. By Lemma 4.9, this assumption implies that $\phi' \circ \phi^{-1}$ is not linear on $\phi(]\phi^{-1}(0), \infty[) = \mathbb{R}^+$. As ϕ is in $C^3(]\phi^{-1}(0), \infty[)$, $\phi' \circ \phi^{-1}$ is in $C^2(\mathbb{R}^+)$. Now the second derivative of $\phi' \circ \phi^{-1}$ cannot be constantly zero on $\mathbb{R}^+$ as $\phi' \circ \phi^{-1}$ is not linear. Let $u = \inf\left\{t > 0 \big| \left(\phi' \circ \phi^{-1}\right)''(t) \neq 0\right\} \geq 0$. Now, there are two cases:

(i) There exists a nonempty interval $J = (u, t) \subset \mathbb{R}^+$ such that $\left(\phi' \circ \phi^{-1}\right)'' < 0$, i.e., $\phi' \circ \phi^{-1}$ is strictly concave on J.

(ii) There exists a nonempty interval $J = (u, t) \subset \mathbb{R}^+$ such that $\left(\phi' \circ \phi^{-1}\right)'' > 0$, i.e., $\phi' \circ \phi^{-1}$ is strictly convex on J.

As $\phi' \circ \phi^{-1}$ is continuously differentiable on $]0, t[$ and linear on $]0, u]$ (by the definition of u), $\phi' \circ \phi^{-1}$ in case (i) is concave on $]0, t[$ and in case (ii) is convex on $]0, t[$. Let $\epsilon > 0$ such that $(1 - \epsilon)^2 t > u$. Since the probability space is rich we may choose $X \in L^\infty$ satisfying both of the following two properties:

(a) $-X \in \phi^{-1}(](1 - \epsilon)t, t[) \subset \phi^{-1}(J)$.

(b) $-X$ is diffuse.

From (a) it follows in particular that $\phi(-X) \in](1 - \epsilon)t, t[\subset J$. Denote

$$
\begin{aligned}
Q_1 &= \arg\max\nolimits_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi(-X)\right]. \\
Q_2 &= \arg\max\nolimits_{\beta(Q)Q \in \partial\bar{\rho}(-\phi(-X))} \beta(Q)\mathrm{E}_Q\left[\phi'(-X)\right].
\end{aligned}
$$

Since $Q_i \ll P$ and $-X$ is diffuse under P we have that $Q_i[-X = x] = 0$ for $i = 1, 2$ and every $x \in \phi^{-1}(J)$. Thus, $-X$ is also diffuse under Q_i. As by (a) and (4.3) $\phi(-X) \in J \subset \mathbb{R}^+$ and $\phi(0) \geq 0$, we have that $\phi(-X) > 0$. Since $\beta(Q_i)Q_i \in \partial\rho(-(\phi(-X)))$, Lemma 4.7 gives

$$
\beta(Q_i) \geq \frac{\operatorname{ess\,inf} \phi(-X)}{\operatorname{ess\,sup} \phi(-X)} \geq \frac{(1 - \epsilon)t}{t} = 1 - \epsilon > 0.
$$

Therefore, $\beta(Q_i)\phi(-X)$ is a diffuse random variable under Q_i and

$$
t > \phi(-X) \geq \beta(Q_i)\phi(-X) \geq (1 - \epsilon)\phi(-X) \geq (1 - \epsilon)^2 t > u,
$$

where the second inequality holds as $\beta(Q_i) \in]0, 1]$. In particular, $\beta(Q_i)\phi(-X) \in J$. Finally let us derive the contradiction. Assume case (i) above: Then

$$
\begin{aligned}
\phi' \circ \phi^{-1}&\left(\max_{i=1,2} \beta(Q_i)\mathrm{E}_{Q^i}\left[\phi(-X)\right]\right) \\
&= \max_{i=1,2} \phi' \circ \phi^{-1}\left(\mathrm{E}_{Q^i}\left[\beta(Q_i)\phi(-X)\right]\right) \\
&> \max_{i=1,2} \mathrm{E}_{Q^i}\left[\phi' \circ \phi^{-1}\left(\beta(Q_i)\phi(-X)\right)\right] \\
&= \max_{i=1,2} \lim_{\delta \downarrow 0} \mathrm{E}_{Q^i}\left[\phi' \circ \phi^{-1}\left(\beta(Q_i)\phi(-X) + (1 - \beta(Q_i))\delta\right)\right] \\
&\geq \max_{i=1,2} \liminf_{\delta \downarrow 0} \left\{\mathrm{E}_{Q^i}\left[\beta(Q_i)\phi' \circ \phi^{-1}\left(\phi(-X)\right)\right] + (1 - \beta(Q_i))\phi' \circ \phi^{-1}(\delta)\right\} \\
&= \max_{i=1,2} \left\{\beta(Q_i)\mathrm{E}_{Q^i}\left[\phi'(-X)\right] + (1 - \beta(Q_i))\liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta)\right\} \\
&\geq \max_{i=1,2} \beta(Q_i)\mathrm{E}_{Q^i}\left[\phi'(-X)\right],
\end{aligned}
$$

where the first inequality holds because of Jensen's inequality for strictly concave functions for the diffuse random variable $\beta(Q_i)\phi(-X) \in J$, with $i = 1, 2$, respectively (where we used that $\beta(Q_i)\phi(-X) \in J$ and the *strict* concavity of $\phi' \circ \phi^{-1}$ on J). The second inequality holds by the concavity of the function $\phi' \circ \phi^{-1}$ on $]0, t[$. The third inequality holds because $\phi' \circ \phi^{-1}(\delta) > 0$

15

for every $\delta > 0$ such that $\phi^{-1}(\delta)$ is well-defined, as ϕ' is positive. The (strict) inequality above is a contradiction to Proposition 4.8, applying to case (i).

Now consider the more challenging case (ii): Then the function $\phi' \circ \phi^{-1}$ is convex on $]0, t]$ and strictly convex on J. Choosing a sequence $\delta_n \downarrow 0$ such that

$$\liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta) = \lim_n \phi' \circ \phi^{-1}(\delta_n),$$

the same argumentation as before yields

$$\phi' \circ \phi^{-1}\left(\max_{i=1,2} \beta(Q_i) \mathrm{E}_{Q^i}\left[\phi(-X)\right] \right)$$
$$< \max_{i=1,2} \mathrm{E}_{Q^i}\left[\phi' \circ \phi^{-1}\left(\beta(Q_i)\phi(-X)\right)\right]$$
$$\leq \max_{i=1,2} \lim_n \left\{ \mathrm{E}_{Q^i}\left[\beta(Q_i)\phi' \circ \phi^{-1}\left(\phi(-X)\right)\right] + (1 - \beta(Q_i))\phi' \circ \phi^{-1}(\delta_n) \right\}$$
$$= \max_{i=1,2} \left\{ \beta(Q_i)\mathrm{E}_{Q^i}\left[\phi'(-X)\right] + (1 - \beta(Q_i))\liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta) \right\}. \tag{4.11}$$

Notice that if

$$(1 - \beta(Q_i))\liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta) = 0, \tag{4.12}$$

then (4.11) would imply that

$$\phi' \circ \phi^{-1}\left(\max_{i=1,2} \beta(Q_i)\mathrm{E}_{Q^i}\left[\phi(-X)\right] \right) < \max_{i=1,2} \beta(Q_i)\mathrm{E}_{Q^i}\left[\phi'(-X)\right],$$

which is a contradiction to Proposition 4.8. To see that $(1 - \beta(Q_i))\liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta) = 0$ note that there are two cases:

(1.) $\bar{\rho}(1) = -1$,

(2.) $\bar{\rho}(1) \neq -1$.

In the first case the second part of Lemma 4.6 implies that $\beta(Q_i) = 1$ for $i = 1, 2$ and in particular, (4.12) is satisfied. Let us look at the second case: By positive homogeneity (2.) entails that $\bar{\rho}(m) \neq -m$ for all $m > 0$. Now suppose that there exists $x_0 \in \mathbb{R}$ such that $\phi(-x_0) < 0$. Since by assumption there also exists x_1 such that $\phi(-x_1) > 0$ the continuity of ϕ yields that the assumption (H1) above holds. In particular, $\phi(0) = 0$. By (2.) and the positive homogeneity of $\bar{\rho}$, $\bar{\rho}(-\phi(-x_0)) \neq \phi(-x_0)$. This gives

$$\phi^{-1}(\bar{\rho}(-\phi(-x_0))) \neq -x_0. \tag{4.13}$$

However, by translation invariance and since $\bar{\rho}(0) = 0$,

$$\phi^{-1}(\bar{\rho}(-\phi(-x_0))) = -x_0 + \phi^{-1}(\bar{\rho}(-\phi(0)) = -x_0 + \phi^{-1}(\bar{\rho}(0)) = -x_0 + \phi^{-1}(0) = -x_0,$$

which is a contradiction to (4.13). Hence, $\phi(x) \geq 0$ for all $x \in \mathbb{R}$ and the assumption (H2) holds, i.e.,

$$\lim_{x \to -\infty} \phi(x) = 0. \tag{4.14}$$

16

By construction in H2 we have $\phi \in C^3(]-\infty, \infty[)$. Now (4.14) implies that the positive function $\phi'(x)$ cannot be bounded constantly away from zero on $(-\infty, z)$ for any $z \in \mathbb{R}$. This means that there is a sequence x_n converging to $-\infty$ such that

$$\liminf_n \phi'(x_n) = 0.$$

Choose $\bar{\delta}_n = \phi(x_n)$. By (4.14) we have that $\lim_n \bar{\delta}_n = 0$ and

$$0 \leq \liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta) \leq \lim_n \phi'(\phi^{-1}(\bar{\delta}_n)) = \lim_n \phi'(x_n) = 0.$$

Consequently,

$$\liminf_{\delta \downarrow 0} \phi' \circ \phi^{-1}(\delta) = 0.$$

This proves (4.12). Hence, we have derived a contradiction to Proposition 4.8, applying to case (ii). Furthermore, we have seen that the cases (H1) and (1.), and (H2) and (2.) coincide, respectively.

Hence, (4.5) of Proposition 4.8 implies that the function $\phi' \circ \phi^{-1}$ has to be linear, and by Lemma 4.9 this implies that there exist constants $p, \gamma, q \in \mathbb{R}$ such that $\phi(x) = pe^{x/\gamma} + q$ or $\phi(x) = px + q$ for all $x \in]\phi^{-1}(0), \infty[$ (where in case H1 $\phi^{-1}(0) = 0$ and in case H2 $\phi^{-1}(0) = -\infty$).

As ϕ is strictly increasing we have $p > 0$. Now in the case (H2) we must have that $\phi(x) = \exp\{x/\gamma\}$ (with $q = 0$) as only then $\lim_{x \to -\infty} \phi(x) = 0$. On the other hand, in the case (H1), condition (1.) holds and the second part of Lemma 4.6 implies that $\beta(Q) = 1$ for all $Q \in M$. Therefore, $\phi^{-1}(\bar{\rho}(-\phi(-X)))$ is invariant under positive affine transformations of ϕ. Thus, we may *always* assume that $q = 0$. Let us first consider the case that ϕ is not linear, i.e., $\phi(x) = e^{x/\gamma}$. Then

$$\phi^{-1}(\bar{\rho}(-\phi(-X))) = \phi^{-1}\left(\sup_{Q \in M} \beta(Q)\mathrm{E}_Q\left[\phi(-X)\right]\right)$$

$$= \gamma \log\left(\sup_{Q \in M} \beta(Q)\mathrm{E}_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right)$$

$$= \sup_{Q \in M}\left\{\gamma \log\left(\mathrm{E}_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) + \gamma \log(\beta(Q))\right\}$$

$$= \sup_{Q \in M}\{e_{\gamma, Q}(X) - c(Q)\},$$

with $c(Q) = -\gamma \log(\beta(Q)) \geq 0$ if $Q \in M$ and $c(Q) = \infty$ else. Thus, indeed $\phi^{-1}(\bar{\rho}(-\phi(-X)))$ is γ-entropy convex. As the supremum on the right-hand side of the first equality is always attained because $\partial\bar{\rho}(-\phi(-X)) \neq \emptyset$, (ii) follows.

Now in the case that ϕ is linear, we may assume that $\phi(x) = px$. But then by our assumptions, $\rho(X) = \phi^{-1}(\bar{\rho}(-\phi(-X))) = \bar{\rho}(X)$ is translation invariant. In particular, ρ is a coherent risk measure attaining its maximum. Thus, ρ is γ-entropy convex (even γ-entropy coherent) with $\gamma = \infty$ and its entropy subdifferential is always nonempty. This completes the proof of the implication (i)$\Rightarrow$(ii) of Theorem 4.1.

Proof of Theorem 4.1. (ii)$\Rightarrow$(i):
To see the direction (ii)$\Rightarrow$(i), we distinguish between two cases: In the case that $\gamma < \infty$,

we let $\phi(x) = e^{x/\gamma}$, and $\bar{\rho}(X) = \sup_{Q \in \mathcal{Q}} \beta(Q)\mathrm{E}_Q[-X]$, with $\beta(Q) = e^{-\rho^*(Q)/\gamma} \geq 0$. Then $\rho(X) = \gamma \log\left(\bar{\rho}\left(-e^{-X/\gamma}\right)\right) = \phi^{-1}(\bar{\rho}(-\phi(-X)))$. Clearly, $\bar{\rho}$ is monotone, convex, positively homogeneous and continuous from above. As $\inf_{Q \in \mathcal{Q}} \rho^*(Q) = 0$, we get that $\sup_{Q \in \mathcal{Q}} \beta(Q) = 1$. This implies that for $m \in \mathbb{R}_0^-$, $\bar{\rho}(m) = -m$. Furthermore, because ρ is entropy convex it is translation invariant.

In the case that $\gamma = \infty$, we let $\phi(x) = x$ and $\bar{\rho}(X) = \rho(X)$. Notice that in both cases we have $\partial\rho(X) \neq \emptyset$ and hence $\partial\bar{\rho}(X) \neq \emptyset$. $\qquad\square$

Corollary 4.10 *In the setting of Theorem 4.1, if $\bar{\rho}$ is additionally assumed to be translation invariant, then statement (i) implies that ρ is γ-entropy coherent with $\gamma \in]0, \infty]$.*

Proof. As $\bar{\rho}$ is assumed to be translation invariant, we have that $\bar{\rho}(m) = -m$ for all $m \in \mathbb{R}$. By Lemma 4.6 this implies that in the proof of Theorem 4.1 we can choose $M \subset \mathcal{Q}$ such that $\beta(Q) = 1$ for all $Q \in M$. Hence, we get $c(Q) = \gamma \log(\beta(Q)) = 0$ if $\beta(Q) = 1$ and ∞ else. Thus, indeed $\phi^{-1}(\bar{\rho}(-\phi(-X)))$ is entropy coherent. $\qquad\square$

Remark 4.11 *In recent work, Cheridito and Kupper [3] (Example 3.6.3) suggest without proof a result quite similar to, but essentially different from, Corollary 4.10. Their suggested result can in a way be viewed as supplementary to the statement in Corollary 4.10: They restrict attention to a specific and simple probabilistic setting with a finite outcome space Ω and consider only strictly positive probability measures on Ω. By contrast, in Corollary 4.10, we consider a rich outcome space and allow for weakly positive probability measures.*

4.2 Entropy Coherence

While Corollary 4.10 in fact already presents an axiomatic characterization of entropy coherent measures of risk, this section axiomatizes entropy coherent measures of risk completely in terms of axioms with respect to a *preference order*. Just as for the robust Savage representation, it will be convenient to work in the framework of Anscombe and Aumann [1] where payoffs can be lotteries; see also Gilboa and Schmeidler [12].

We will therefore embed the space $L^\infty(\Omega, \mathcal{F}, P)$ into a space $\tilde{\mathcal{X}}$ of functions $\tilde{X}$ on $(\Omega, \mathcal{F})$ taking values in the convex set of probability measures $\mathcal{M}_b(\mathbb{R})$ defined by

$$\mathcal{M}_b(\mathbb{R}) = \left\{\mu \in \mathcal{M}_1(\mathbb{R}) \,|\, \mu([-u, u]) = 1 \text{ for some } u \geq 0\right\},$$

where $\mathcal{M}_1(\mathbb{R})$ is the set of all σ-additive measures on $\mathbb{R}$ with mass one. Recall that a stochastic kernel is a mapping $K : \Omega \to \mathcal{M}_1(\mathbb{R})$ such that $\omega \mapsto K(\omega, A)$ is measurable for each Borel set $A \subset \mathbb{R}$. Formally, then, $\tilde{\mathcal{X}}$ is defined as the convex set of all stochastic kernels $\tilde{X}$ on $(\Omega, \mathcal{F})$ for which there exists a constant $u \geq 0$ such that

$$\tilde{X}(\omega, [-u, u]) = 1 \text{ for } P\text{-a.s. all } \omega \in \Omega.$$

As mentioned already in the Introduction, $L^\infty(\Omega, \mathcal{F}, P)$ can then be embedded into $\tilde{\mathcal{X}}$ by the mapping

$$L^\infty(\Omega, \mathcal{F}, P) \ni X \to \delta_X \in \tilde{\mathcal{X}}.$$

We assume that a 'preference' (or rather: dispreference) order $\succeq$ is defined on $\tilde{X}$, where, loosely speaking, $\tilde{X} \succeq \tilde{Y}$ means that $\tilde{X}$ is at least as risky as $\tilde{Y}$. As usual, $\succ$ stands for strict

'preference' (or rather: dispreference) and $\sim$ for indifference. We suppose that $\succeq$ satisfies the following properties:

AXIOM A1–Non-degenerate Weak Order: $\succeq$ is complete, transitive and non-degenerate.

AXIOM A2–Continuity I: If $\tilde{X}, \tilde{Y}, \tilde{Z} \in \tilde{\mathcal{X}}$ are such that $\tilde{Z} \succ \tilde{Y} \succ \tilde{X}$, then there are $\lambda, \beta \in]0, 1[$ with

$$\lambda \tilde{Z} + (1 - \lambda)\tilde{X} \succ \tilde{Y} \succ \beta \tilde{Z} + (1 - \beta)\tilde{X}.$$

Moreover, for all $u > 0$ the restriction of $\succeq$ to $\mathcal{M}_1([-u, u])$ is continuous with respect to the weak topology.

AXIOM A3–Monotonicity: For all $\tilde{X}, \tilde{Y} \in \tilde{X}$: If $\tilde{X}(\omega) \succeq \tilde{Y}(\omega)$ for P-a.s. all $\omega \in \Omega$, then $\tilde{X} \succeq \tilde{Y}$. Moreover, $\succeq$ is compatible with the usual risk order on $\mathbb{R}$, i.e., $\delta_x \succ \delta_y$ if an only if $y > x$.

AXIOM A4–Uncertainty Aversion: If $\tilde{X}, \tilde{Y} \in \tilde{\mathcal{X}}$ and $\lambda \in [0, 1]$, then $\tilde{X} \sim \tilde{Y}$ implies $\tilde{Y} \succeq \lambda \tilde{X} + (1 - \lambda)\tilde{Y}$.

AXIOM A5–Certainty Independence: If $\tilde{X}, \tilde{Y} \in \tilde{\mathcal{X}}$ and $Z \equiv \mu \in \mathcal{M}_b(\mathbb{R})$, then

$$\tilde{X} \succeq \tilde{Y} \quad \Leftrightarrow \quad \lambda \tilde{X} + (1 - \lambda)Z \succeq \lambda \tilde{Y} + (1 - \lambda)Z \quad \text{for all } \lambda \in]0, 1[.$$

AXIOM A6–Continuity II: The induced preference order $\succeq$ on L^∞ satisfies the following continuity properties:

(i) $Y \succ X$ and $X_n \downarrow X$ a.s. $\Rightarrow$ $Y \succ X_n$ for all large n.

(ii) $X \succ Y$ and $X_n \uparrow X$ a.s. $\Rightarrow$ $X_n \succ Y$ for all large n.

Axioms A1-A6 are standard and similar to the ones imposed by Gilboa and Schmeidler [12]. We add the following two axioms to obtain entropy convex functionals on L^∞:

AXIOM A7–Translation Invariance: The induced preference order $\succeq$ on L^∞ satisfies for all $m \in \mathbb{R}$:

$$X \succeq Y \Rightarrow X + m \succeq Y + m.$$

AXIOM A8–Smoothness and Sensitivity: Let $x \in \mathbb{R}$ and $p \in]0, 1[$. Denote by $c_{p,x}$ the certainty equivalent of the lottery $\mu = p\delta_x$ that pays out x with probability p. Then, for all $x \in \mathbb{R}$, the function $p \to c_{x,p}$ is in $C^3(]0, 1[)$. Furthermore, for all $x \in \mathbb{R}$ and all $p, q \in]0, 1[$, there exists $m \in \mathbb{R}$ such that $c_{m,p} \leq c_{x,q}$.

With Axioms A1-A8 at hand, we state the following theorem:

Theorem 4.12 *Suppose that $(\Omega, \mathcal{F}, P)$ is rich. Then a preference order satisfies A1-A8 if and only if there exists a convex set $M \subset \mathcal{Q}$ and $\gamma \in]0, \infty]$ such that*

$$\tilde{X} \succeq \tilde{Y} \Leftrightarrow \max_{Q \in M} \gamma \log \left(\mathrm{E}_Q \left[\int \exp \left\{ \frac{-x}{\gamma} \right\} \tilde{X}(., dx) \right] \right) \geq \max_{Q \in M} \gamma \log \left(\mathrm{E}_Q \left[\int \exp \left\{ \frac{-x}{\gamma} \right\} \tilde{Y}(., dx) \right] \right),$$

where the case that $\gamma = \infty$ is identified with $\gamma \log(x) = x = e^{x/\gamma}$. In particular, for $X, Y \in L^\infty$ we have that

$$X \succeq Y \Leftrightarrow \rho(X) \geq \rho(Y)$$

for a γ-entropy coherent risk measure ρ, $\gamma \in]0, \infty]$, with a nonempty subdifferential.

Proof. The 'if' direction is straightforward (for Axiom A6, see Proposition 3.12). Let us prove the 'only if' direction. Define $\succeq^*$ by saying that $\tilde{Y} \succeq^* \tilde{X}$ if and only if $\tilde{X} \succeq \tilde{Y}$. Then our assumptions A1-A6 imply that the assumptions of Theorem 2.80, Föllmer and Schied [9], are satisfied by $\succeq^*$. Hence, there exists a convex set M of probability measures on $(\Omega, \mathcal{F})$ and a strictly increasing and continuous function u such that the preference order $\succeq^*$ has a numerical representation

$$\tilde{U}(\tilde{X}) = \min_{Q \in M} \mathrm{E}_Q \left[\int u(x) \tilde{X}(\cdot, dx) \right],$$

where u is unique up to positive affine transformations. This implies that one numerical representation L of $\succeq$ is given by

$$L(\tilde{X}) := -\tilde{U}(\tilde{X}) = \max_{Q \in M} \mathrm{E}_Q \left[\int -u(x) \tilde{X}(\cdot, dx) \right] = \max_{Q \in M} \mathrm{E}_Q \left[\int l(-x) \tilde{X}(\cdot, dx) \right],$$

where $l(x) = -u(-x)$. Because any positive affine transformation of l generates the same preferences, we may assume that $l(0) = 0$ and $l(1) = 1$. Suppose that the maximum is attained in a $\bar{Q} \notin \mathcal{Q}$. Then there exists $A \in \mathcal{F}$ such that $P[A] = 0$ and $\bar{Q}[A] > 0$. But then for $X \in L^\infty$ we get

$$L(X - I_A) = \max_{Q \in M_v} \mathrm{E}_Q \left[l(-(X - I_A)) \right] \geq \mathrm{E}_{\bar{Q}} \left[l(-(X - I_A)) \right] > \mathrm{E}_{\bar{Q}} \left[l(-X) \right] = L(X).$$

This is a contradiction to the fact that by A3 we must have $X \sim X - I_A$. Hence, we can choose M as a convex subset of $\mathcal{Q}$.

Next, notice that any strictly monotone transformation of L is also a numerical representation of $\succeq$. In particular, the functional

$$\rho(X) := l^{-1} \left(\max_{Q \in M} \mathrm{E}_Q \left[\int l(-x) \tilde{X}(\cdot, dx) \right] \right)$$

is a numerical representation of $\succeq$. Note that $-\rho$ is the certainty equivalent of $\succeq$. Hence, by A8 for all $x \in \mathbb{R}$ the function $f(p) := -\rho(p\delta_x) = -l^{-1}(l(-x)p)$ is in $C^3(]0, 1[)$ for $p \in]0, 1[$. As $l(0) = 0$ this implies that l^{-1} is in $C^3(]0, l(\infty)[)$ and consequently, the function l is in $C^3(]l^{-1}(0), \infty[) = C^3(\mathbb{R}^+)$ as well. Now as $X \sim -\rho(X)$ by A7 we have $X + m \sim -\rho(X) + m$. This yields

$$X + m \sim -\rho(X) + m.$$

As $\rho(m) = -m$ for all $m \in \mathbb{R}$ we obtain

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

Hence, ρ is translation invariant on L^∞. Set $\phi := l$ and $\bar{\rho}(X) := \sup_{Q \in M} \mathrm{E}_Q[-X]$. Since $\bar{\rho}$ is a coherent risk measure which always attains its supremum, the subdifferential of $\bar{\rho}$ is always nonempty. As we have seen that

$$\rho(X) = l^{-1} \left(\max_{Q \in M} \mathrm{E}_Q \left[\int l(-x) \tilde{X}(\cdot, dx) \right] \right) = \phi^{-1} \left(\bar{\rho}(-\phi(-X)) \right)$$

is translation invariant, Corollary 4.10, jointly with the sensitivity condition in A8 ensuring $l(\infty) = \infty$, implies that ρ is γ-entropy coherent with $\gamma \in]0, \infty]$. $\square$

4.3 Convexity Without the Translation Invariance Axiom

In the previous subsections the axiom of translation invariance played a key role. As is well-documented (see, for example, Cheridito and Kupper [3]), the axiom of translation invariance is equivalent to the axiom of convexity for general certainty equivalents under fairly weak conditions (e.g., continuity with respect to the L^∞-norm). In this subsection we adapt and apply this equivalence relation to the present setting, to replace the axiom of translation invariance by the axiom of convexity, which will now play the key role.

Throughout this subsection, we let the probability space $(\Omega, \mathcal{F}, P)$ be sufficiently rich to support a random variable with a uniform distribution. We state the following theorem:

Theorem 4.13 *Let $\bar{\rho} : L^\infty \to \mathbb{R}$ be monotone, convex, positively homogeneous and continuous from above, and let for all $m \in \mathbb{R}_0^-$, $\bar{\rho}(m) = -m$. Suppose that the subdifferential of $\bar{\rho}$ is always nonempty. Furthermore, suppose that $r : L^\infty \to \mathbb{R}$ is defined by $r(X) = l^{-1}(\bar{\rho}(-l(-X)))$, for a strictly increasing and continuous function $l \in C^3(]l^{-1}(0), \infty[)$. Finally, suppose that $0 \in closure(Image(l))$ and that $l(\infty) = \infty$. Then the following statements are equivalent:*

(i) r is convex and $r(m) = -m$ for all $m \in \mathbb{R}$.

(ii) r is γ-entropy convex with $\gamma \in \mathbb{R}^+$ or r is ∞-entropy coherent.

Proof. The direction from (ii) to (i) holds by virtue of Proposition 3.4. Let us show the reverse direction. First, notice that r is continuous with respect to the L^∞-norm. This can be seen as follows: From the proof of Lemma 4.6, we have that $\bar{\rho}(X) = \sup_{X' \in H}\{E[-X'X]\}$ with $H \subset L^1_+$ and $\sup_{X' \in H}\{E[|X'|]\} = \sup_{X' \in H}\{E[X']\} = 1$. Hence, for $X, Y \in L^\infty$,

$$\bar{\rho}(Y) - \bar{\rho}(X) = \sup_{X' \in H}\{E[-X'Y]\} - \sup_{X' \in H}\{E[-X'X]\}$$
$$\leq \sup_{X' \in H}\{E[-X'Y] - E[-X'X]\}$$
$$\leq ||Y - X||_\infty \sup_{X' \in H} E[|X'|] = ||Y - X||_\infty.$$

Switching the roles of X and Y it follows that $\bar{\rho}$ is indeed continuous with respect to the L^∞-norm. Now as l is continuous we can conclude that r is continuous with respect to the L^∞-norm as well. But then it follows from Proposition 2.5-(8) in Cheridito and Kupper [3] that r is translation invariant. The argument is simple, namely, for $\lambda \in (0, 1)$ we have

$$r(X + m) \leq \lambda r\left(\frac{X}{\lambda}\right) + (1 - \lambda)r\left(\frac{m}{1 - \lambda}\right) = \lambda r\left(\frac{X}{\lambda}\right) - m.$$

Letting λ converge to one and using the continuity of r with respect to the L^∞-norm we find that $r(X + m) \leq r(X) - m$. Replacing X by $X + m$ and m by $-m$ yields the stated result. Therefore, r is indeed translation invariant. Now upon application of Theorem 4.1, the direction from (i) to (ii) follows. $\square$

Using Corollary 4.10, we now obtain directly the following corollary:

Corollary 4.14 *In the setting of Theorem 4.13, suppose that $\bar{\rho}$ is additionally assumed to be translation invariant. Then the following statements are equivalent:*

(i) *r is convex.*

(ii) *r is γ-entropy coherent with $\gamma \in]0, \infty]$.*

5   The Dual Conjugate

In this section we study the dual conjugate function, defined in (2.2), for entropy coherent and entropy convex measures of risk. Quite unusually, some explicit results on the dual conjugate function can be obtained. Let $\gamma \in [0, \infty]$. We state the following proposition:

Proposition 5.1 *Suppose that ρ is γ-entropy convex. Then*

$$\rho^*(Q) = \sup_{\bar{P} \ll P} \{\alpha(\bar{P}) - \gamma H(\bar{P}|Q)\}. \tag{5.1}$$

Proof. We write

$$\rho^*(Q) = \sup_{X \in L^\infty} \{e_{\gamma,Q}(X) - \rho(X)\} = \sup_{X \in L^\infty} \sup_{\bar{P} \ll P} \{\mathrm{E}_{\bar{P}}[-X] - \gamma H(\bar{P}|Q) - \rho(X)\}$$

$$= \sup_{\bar{P} \ll P} \sup_{X \in L^\infty} \{\mathrm{E}_{\bar{P}}[-X] - \rho(X) - \gamma H(\bar{P}|Q)\} = \sup_{\bar{P} \ll P} \{\alpha(\bar{P}) - \gamma H(\bar{P}|Q)\}.$$

$\square$

Notice that (5.1) yields that $\alpha(\bar{P}) \le \rho^*(Q) + \gamma H(\bar{P}|Q)$. Hence,

$$\alpha(\bar{P}) \le \inf_{Q \in \mathcal{Q}} \{\rho^*(Q) + \gamma H(\bar{P}|Q)\}.$$

The next penalty function duality theorem will show that this inequality is sharp. It also establishes the explicit (!) relationship between the dual conjugate α and the penalty function c for γ-entropy convex measures of risk.

Theorem 5.2 *Suppose that ρ is γ-entropy convex with penalty function c. Then:*

(i) *The dual conjugate of ρ, defined in (2.2), is given by the largest convex and lower-semicontinuous function α being dominated by $\inf_{Q \in \mathcal{Q}} \{\gamma H(\bar{P}|Q) + c(Q)\}$.*

(ii) *If c is convex and lower-semicontinuous, then α is the largest lower-semicontinuous function being dominated by $\inf_{Q \in \mathcal{Q}} \{\gamma H(\bar{P}|Q) + c(Q)\}$.*

(iii) *If c is convex and lower-semicontinuous and for every $r \in \mathbb{R}^+$ the set $B_r = \{Q \in \mathcal{Q} | c(Q) \le r\}$ is uniformly integrable, then*

$$\alpha(\bar{P}) = \min_{Q \in \mathcal{Q}} \{\gamma H(\bar{P}|Q) + c(Q)\}. \tag{5.2}$$

22

Proof. If $\gamma = 0$ or $\gamma = \infty$ the theorem follows by standard arguments. Let us therefore assume that $\gamma \in \mathbb{R}^+$.

(i): We write

$$
\rho(X) = \sup_{Q \in \mathcal{Q}} \left\{ \gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) - c(Q) \right\} = \sup_{Q \in \mathcal{Q}} \sup_{\bar{P} \ll P} \left\{ \mathrm{E}_{\bar{P}} \left[-X \right] - \gamma H(\bar{P}|Q) - c(Q) \right\}
$$
$$
= \sup_{\bar{P} \ll P} \sup_{Q \in \mathcal{Q}} \left\{ \mathrm{E}_{\bar{P}} \left[-X \right] - \gamma H(\bar{P}|Q) - c(Q) \right\} = \sup_{\bar{P} \ll P} \left\{ \mathrm{E}_{\bar{P}} \left[-X \right] - \inf_{Q \in \mathcal{Q}} \left\{ \gamma H(\bar{P}|Q) + c(Q) \right\} \right\},
$$

$$(5.3)$$

where we have used in the second equality that $H(\bar{P}|Q) = \infty$ if $\bar{P}$ is not absolutely continuous with respect to Q. Since α is the minimal lower-semicontinuous and convex function satisfying (2.1), statement (i) follows.

(ii): Now assume that c is convex and lower-semicontinuous. We will first show that:

(a) $\gamma H(\bar{P}|Q)$ is jointly convex in $(\bar{P}, Q)$.

(b) If $\bar{P}_n$ and Q_n converge weakly to $\bar{P}$ and $\bar{Q}$, respectively, then $\gamma H(\bar{P}|Q) \leq \liminf_n \gamma H(\bar{P}_n|Q_n)$.

To see (a), note that for every $X \in L^\infty$, $-\gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right)$ is convex in Q, and $\mathrm{E}_{\bar{P}} \left[-X \right]$ is convex in $\bar{P}$. Hence, $\mathrm{E}_{\bar{P}} \left[-X \right] - \gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right)$ is jointly convex in $(\bar{P}, Q)$ and therefore

$$
\gamma H(\bar{P}|Q) = \sup_{X \in L^\infty} \left\{ \mathrm{E}_{\bar{P}} \left[-X \right] - \gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\}
$$

is jointly convex in $(\bar{P}, Q)$ as well.

(b) If $Q_n \in \mathcal{Q}$ converges weakly to Q, and $\bar{P}_n \in \mathcal{Q}$ converges weakly to $\bar{P}$, then for every $X \in L^\infty$ we have $\mathrm{E}_{Q_n} \left[-X \right] \overset{n \to \infty}{\to} \mathrm{E}_Q \left[-X \right]$ and $\mathrm{E}_{\bar{P}_n} \left[-X \right] \overset{n \to \infty}{\to} \mathrm{E}_{\bar{P}} \left[-X \right]$. Since

$$
\mathrm{E}_{\bar{P}} \left[-X \right] - \gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) = \lim_n \left\{ \mathrm{E}_{\bar{P}_n} \left[-X \right] - \gamma \log \left(\mathrm{E}_{Q_n} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\}
$$
$$
\leq \liminf_n \sup_{X \in L^\infty} \left\{ \mathrm{E}_{\bar{P}_n} \left[-X \right] - \gamma \log \left(\mathrm{E}_{Q_n} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\},
$$

it follows that

$$
\gamma H(\bar{P}|Q) = \sup_{X \in L^\infty} \left\{ \mathrm{E}_{\bar{P}} \left[-X \right] - \gamma \log \left(\mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\}
$$
$$
\leq \liminf_n \sup_{X \in L^\infty} \left\{ \mathrm{E}_{\bar{P}_n} \left[-X \right] - \gamma \log \left(\mathrm{E}_{Q_n} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\}
$$
$$
= \liminf_n \gamma H(\bar{P}_n|Q_n).
$$

This proves (b).

(a) and (b) imply that $\gamma H(\bar{P}|Q)$ is jointly convex and lower-semicontinuous in $(\bar{P}, Q)$. Furthermore, $c(Q)$ is convex and lower-semicontinuous. Therefore $\gamma H(\bar{P}|Q) + c(Q)$ is jointly convex and lower-semicontinuous as well. By Theorem 2.1.3 (v) of Zalinescu [26] it follows

that $\inf_{Q \in \mathcal{Q}}\{\gamma H(\bar{P}|Q) + c(Q)\}$ is convex in $\bar{P}$. Now (ii) follows since α is the minimal lower-semicontinuous and convex function satisfying (2.1).

(iii): If we could show that

$$\beta(\bar{P}) = \inf_{Q \in \mathcal{Q}}\{\gamma H(\bar{P}|Q) + c(Q)\} \tag{5.4}$$

is also lower-semicontinuous and that the infimum is attained, then (5.2) would follow from the uniqueness of α. First of all let us show that the infimum in (5.4) is attained. Let $Q_k \ll P$ be the minimizing sequence. Since $c \neq \infty$ we have for all $\bar{P}$ that $\beta(\bar{P}) < \infty$. Thus,

$$\limsup_k c(Q_k) \leq \limsup_k \gamma H(\bar{P}|Q_k) + c(Q_k) = \beta(\bar{P}) < \infty.$$

In particular, $(c(Q_k))_k$ is a bounded sequence. By our assumptions, Q_k must be a uniformly integrable sequence and by the Theorem of Dunford-Pettis, see for instance Theorem IV.8.9 in Dunford and Schwartz [7], the sequence Q_k is weakly relatively compact. Hence, for fixed $\bar{P}$ we may take the infimum in (5.4) over the weakly compact set $\{Q_1, Q_2, \ldots\}$. As by (b) above $Q \to \gamma H(\bar{P}|Q) + c(Q)$ is lower-semicontinuous we may infer that the infimum is attained.

So suppose that $\bar{P}_n$ converges weakly to $\bar{P}$. For the lower-semicontinuity we have to show that

$$\beta(\bar{P}) \leq \liminf_n \beta(\bar{P}_n). \tag{5.5}$$

If $\liminf_n \beta(\bar{P}_n) = \infty$ then clearly (5.5) holds. So assume that $r := \liminf_n \beta(\bar{P}_n) < \infty$. Denote by $(n_j)_j$ the subsequence such that $\liminf_n \beta(\bar{P}_n) = \lim_j \beta(\bar{P}_{n_j})$. Let

$$\bar{Q}_{n_j} \in \arg\min_{Q \in \mathcal{Q}}\{\gamma H(\bar{P}_{n_j}|Q) + c(Q)\}.$$

As $\limsup_j c(Q_{n_j}) \leq \lim_j \gamma H(\bar{P}_{n_j}|Q_{n_j}) + c(Q_{n_j}) = r$, the sequence Q_{n_j} is uniformly integrable. Again by the Theorem of Dunford-Pettis, Q_{n_j} has a subsequence, denoted by n_{j_k}, converging weakly to a measure $\bar{Q} \in \mathcal{Q}$. Hence, by the lower-semicontinuity of the mapping $(\bar{P}, Q) \to H(\bar{P}|Q)$ proved in (b),

$$\beta(\bar{P}) = \min_{Q \in \mathcal{Q}}\{\gamma H(\bar{P}|Q) + c(Q)\} \leq \gamma H(\bar{P}|\bar{Q}) + c(\bar{Q})$$
$$\leq \liminf_k \gamma H(\bar{P}_{n_k}|\bar{Q}_{n_{j_k}}) + c(Q_{n_{j_k}}) = \liminf_n \beta(\bar{P}_n),$$

where the second equality holds because n_{j_k} was a subsequence of the sequence n_j. Hence, indeed β is lower-semicontinuous and we can conclude that $\beta = \alpha$. □

Corollary 5.3 *Suppose that*

$$\rho(X) = \sup_{Q \in M} e_{\gamma, Q}(X)$$

for a convex set $M \subset \mathcal{Q}$. Then the dual conjugate of ρ is given by the largest lower-semicontinuous function α being dominated by $\inf_{Q \in M} \gamma H(\bar{P}|Q)$. Furthermore, if M is weakly relatively compact, then

$$\alpha(\bar{P}) = \min_{Q \in M} \gamma H(\bar{P}|Q). \tag{5.6}$$

Proof. The first part of the corollary is precisely (ii) of Theorem 5.2 with $c = \bar{I}_M$. The second part follows as for all $r \in \mathbb{R}^+$ we have $\{Q \in \mathcal{Q} | c(Q) \leq r\} = M$. (5.6) now follows as by the Theorem of Dunford-Pettis, M is weakly relatively compact if and only if M is uniformly integrable. $\square$

Corollary 5.4 *Suppose that ρ is a convex risk measure with dual conjugate α for which*

$$\alpha(P) = 0 \text{ and } \alpha(Q) > 0 \text{ if } Q \neq P.$$

Then ρ is γ-entropy coherent if and only if $\rho(X) = e_\gamma(X)$ for $\gamma \in]0, \infty]$.

Proof. The 'if' direction is trivial. Let us prove the 'only if' direction. If ρ is γ-entropy coherent, then by Corollary 5.3 we must have $\alpha(\bar{P}) \leq \inf_{Q \in M} \gamma H(\bar{P}|Q)$ for a convex set M. Note that if $\bar{P} \in M$ then $0 \leq \alpha(\bar{P}) \leq \inf_{Q \in M} \gamma H(\bar{P}|Q) = 0$. By the assumptions on α this implies that M can at most contain P. Hence, either $\alpha(\bar{P}) = \gamma H(\bar{P}|P)$ for all $\bar{P} \ll P$, or $M = \emptyset$ and $\alpha = \infty$. However, as $\inf_Q \alpha(Q) = \rho(0) = 0$ we must have that $\alpha(\bar{P}) = \gamma H(\bar{P}|P)$. Therefore, by (2.1) indeed

$$\rho(X) = \sup_{\bar{P} \in \mathcal{Q}} \{E_{\bar{P}}[-X] - \gamma H(\bar{P}|P)\} = e_\gamma(X).$$

$\square$

Corollary 5.5 *Let ρ be a γ-entropy convex risk measure with penalty function c. Let $p \in (1, \infty]$. Suppose that there exists a function $f : \mathbb{R}^+ \to \mathbb{R}^+$, with $f(x) \to \infty$ as $x \to \infty$, such that for all $Q \in \mathcal{Q}$ we have $c(Q) \geq f\left(E\left[\left|\frac{dQ}{dP}\right|^p\right]\right)$. Then*

$$\alpha(\bar{P}) = \min_{Q \in \mathcal{Q}} \{\gamma H(\bar{P}|Q) + c(Q)\}.$$

Proof. Let $r \in \mathbb{R}^+$. By assumption, there exists $\bar{r} \geq 0$ such that $f(x) > r$ for all $x > \bar{r}$. This entails that

$$\{Q \ll P | c(Q) \leq r\} \subset \left\{Q \ll P | f\left(E\left[\left|\frac{dQ}{dP}\right|^p\right]\right) \leq r\right\} \subset \left\{Q \ll P \left|\left|\frac{dQ}{dP}\right|^p \leq \bar{r}\right.\right\}.$$

By the Lemma of de la Vallée-Poussin, the last set is uniformly integrable. Hence, $\{Q \ll P | c(Q) \leq r\}$ is uniformly integrable as well. The corollary now follows from Theorem 5.2, (iii). $\square$

Corollary 5.6 *Let ρ be a convex risk measure. Then the following statements are equivalent:*

(i) *For a convex and lower-semicontinuous function c from $\mathcal{Q}$ to $[0, \infty]$ with $\inf_{Q \in \mathcal{Q}} c(Q) = 0$ and uniformly integrable sublevel sets we have*

$$\alpha(\bar{P}) = \min_{Q \in \mathcal{Q}} \{\gamma H(\bar{P}|Q) + c(Q)\}. \tag{5.7}$$

(ii) *ρ is γ-entropy convex with a convex and lower-semicontinuous penalty function c which has uniformly integrable sublevel sets.*

25

Proof. The direction from (ii) to (i) is precisely part (iii) of Theorem 5.2. The reverse direction holds since

$$\rho(X) = \sup_{\bar{P}\in\mathcal{Q}} \{E_{\bar{P}}[-X] - \alpha(\bar{P})\} = \sup_{\bar{P}\in\mathcal{Q}} \{E_{\bar{P}}[-X] - \min_{Q\in\mathcal{Q}}[\gamma H(\bar{P}|Q) + c(Q)]\}$$

$$= \sup_{Q\in\mathcal{Q}} \sup_{\bar{P}\in\mathcal{Q}} \{E_{\bar{P}}[-X] - \gamma H(\bar{P}|Q) + c(Q)\} = \sup_{Q\in\mathcal{Q}} \{e_{\gamma,Q}(X) - c(Q)\}.$$

□

In the case that the penalty functions admit uniformly integrable sublevel sets, the next theorem establishes a complete characterization of entropy convexity involving only the dual conjugate α. It shows that entropy convexity is equivalent to a min-max being a max-min.

Theorem 5.7 *Suppose that ρ is a convex risk measure. Furthermore, let c be defined by $c(Q) := \sup_{\hat{P}\ll P}\{\alpha(\hat{P}) - \gamma H(\hat{P}|Q)\}$. Then the following statements are equivalent:*

(i) ρ is γ-entropy convex with ρ^* having uniformly integrable sublevel sets.

(ii) c is convex and lower-semicontinuous with $\inf_{Q\in\mathcal{Q}} c(Q) = 0$ and uniformly integrable sublevel sets, and for every $\bar{P}\in\mathcal{Q}$,

$$\inf_{Q\in\mathcal{Q}} \sup_{\hat{P}\in\mathcal{Q}} \left\{\gamma H(\bar{P}|Q) + \alpha(\hat{P}) - \gamma H(\hat{P}|Q)\right\} \tag{5.8}$$

$$= \sup_{\hat{P}\in\mathcal{Q}} \inf_{Q\in\mathcal{Q}} \left\{\gamma H(\bar{P}|Q) + \alpha(\hat{P}) - \gamma H(\hat{P}|Q)\right\}.$$

Proof. If $\gamma = 0$, both (i) and (ii) cannot hold, so that the theorem holds trivially. Suppose, therefore, that $\gamma \in]0,\infty]$. We can write the right-hand side of (5.8) as

$$\sup_{\hat{P}\in\mathcal{Q}} \inf_{Q\in\mathcal{Q}} \left\{\gamma E_Q\left[\log\left(\frac{d\bar{P}}{dQ}\right) - \log\left(\frac{d\hat{P}}{dQ}\right)\right] + \alpha(\hat{P})\right\} = \sup_{\hat{P}\in\mathcal{Q}} \inf_{Q\in\mathcal{Q}} \left\{\gamma E_Q\left[\log\left(\frac{d\bar{P}}{d\hat{P}}\right)\right] + \alpha(\hat{P})\right\}.$$

If $\frac{d\bar{P}}{d\hat{P}} \neq 1$ on a non-zero set we have that $\log\left(\frac{d\bar{P}}{d\hat{P}}\right) < 0$ on a non-zero set. But then

$$\inf_{Q\in\mathcal{Q}} \gamma E_Q\left[\log\left(\frac{d\bar{P}}{d\hat{P}}\right)\right] = -\infty.$$

Consequently, we have to choose $\hat{P} = \bar{P}$ in the supremum above, which implies that the right-hand side in (5.8) is equal to $\alpha(\bar{P})$. Moreover, clearly for the left-hand side we have that

$$\inf_{Q\in\mathcal{Q}} \sup_{\hat{P}\in\mathcal{Q}} \left\{\gamma H(\bar{P}|Q) + \alpha(\hat{P}) - \gamma H(\hat{P}|Q)\right\} = \inf_{Q\in\mathcal{Q}} \left\{\gamma H(\bar{P}|Q) + \sup_{\hat{P}\in\mathcal{Q}}\{\alpha(\hat{P}) - \gamma H(\hat{P}|Q)\}\right\}$$

$$= \inf_{Q\in\mathcal{Q}} \left\{\gamma H(\bar{P}|Q) + c(Q)\right\}.$$

Now the theorem follows from Proposition 5.1 and Corollary 5.6. □

6 Distribution Invariant Entropy Convex Measures of Risk

In this section, we derive the distribution invariant representation for entropy coherent and entropy convex measures of risk. As a bridge towards the distribution invariant representation, we first present a preference axiomatization of entropy coherent measures of risk, taking Schmeidler's [23] preference axioms rather than the preference axioms of Gilboa and Schmeidler [12] (see Section 4.2) as a starting point.

Let $\gamma \in [0, \infty]$. For a normalized, monotone and possibly non-additive measure (or, set function) $v : \mathcal{F} \to [0,1]$ and a (truly) bounded random variable X we define

$$\mathrm{E}_v[X] := \int X dv := \int_0^\infty v[X > t] dt + \int_{-\infty}^0 (v[X > t] - 1) dt.$$

We say that v is submodular if

$$v(A \cap B) + v(A \cup B) \leq v(A) + v(B) \text{ for } A, B \in \mathcal{F}.$$

By Schmeidler [22, 23], v is submodular if and only if for every bounded X,

$$\mathrm{E}_v[X] = \max_{Q \in M_v} \mathrm{E}_Q[X], \tag{6.1}$$

with $M_v = \{Q \text{ is additive on } \mathcal{F} | Q(A) \leq v(A) \text{ for all } A \in \mathcal{F}\}$. M_v is also called the core of v.

We note that every bounded random variable is an element in L^∞, and that every P-almost surely bounded random variable can be identified with a (truly) bounded random variable (by (re)defining $X \in L^\infty$ to be equal to its original value for those $\omega \in \Omega$ for which $|X(\omega)| \leq ||X||_\infty$ and by setting $X(\omega) = 0$ otherwise). Then, for $X \in L^\infty$, we define

$$e_{\gamma,v}(X) := \gamma \log \left(\int \exp \left\{ \frac{-X}{\gamma} \right\} dv \right).$$

In the case that v is continuous from above, that is, if $v(A_n) \downarrow 0$ for any decreasing sequence of events (A_n) such that $\bigcap_n A_n = \emptyset$, we have that (6.1) holds with $M_v = \{Q \in \mathcal{Q} | Q(A) \leq v(A) \text{ for all } A \in \mathcal{F}\}$. We state the following proposition:

Proposition 6.1 *The following statements are equivalent:*

(i) $\rho(X) = e_{\gamma,v}(X)$ is γ-entropy coherent and continuous from below.

(ii) v is submodular and continuous from above, and $\rho(X) = \max_{Q \in M_v} e_{\gamma,Q}(X)$ with $M_v = \{Q \in \mathcal{Q} | Q(A) \leq v(A) \text{ for all } A \in \mathcal{F}\}$.

Proof. The direction (ii)⇒(i) follows from (6.1) and the fact that v being continuous from above implies that $M_v \subset \mathcal{Q}$. Furthermore, ρ is continuous from below by virtue of Proposition 3.12.

To see the reverse direction, let $M \subset \mathcal{Q}$ with $\bar{I}_M = \rho^*$. Since $\gamma \log \left(\int \exp \left\{ \frac{-X}{\gamma} \right\} dv \right) = \rho(X) = \sup_{Q \in M} e_{\gamma,Q}(X)$, we have that

$$\int \exp \left\{ \frac{-X}{\gamma} \right\} dv = \sup_{Q \in M} \mathrm{E}_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right].$$

27

Now on the one hand, for $A \in \mathcal{F}$, setting $X = -\gamma \log(e^{\frac{-1}{\gamma}} + 1) < 1$ on A and 1 else, we get for $Q \in M$

$$
\begin{aligned}
(e^{-1/\gamma} + 1 - e^{-1/\gamma})v(A) + e^{-1/\gamma} &= \int \exp\left\{\frac{-X}{\gamma}\right\} dv \\
&\geq \mathrm{E}_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right] = (e^{-1/\gamma} + 1 - e^{-1/\gamma})Q(A) + e^{-1/\gamma}.
\end{aligned}
$$

Thus, $Q(A) \leq v(A)$ and we may infer that $\sup_{Q \in M} Q(A) \leq v(A)$, which implies $M \subset M_v \subset \mathcal{Q}$. On the other hand, if $Q \in M_v$, then

$$
\rho(X) = \int \exp\left\{\frac{-X}{\gamma}\right\} dv \geq \int \exp\left\{\frac{-X}{\gamma}\right\} dQ = e_{\gamma,Q}(X).
$$

Since $Q \in \mathcal{Q}$, this entails that $\rho^*(Q) = 0$. In particular, $Q \in M$. $\qquad\square$

Next, we present an axiomatic characterization of $e_{\gamma,v}$. We need the following definition:

Definition 6.2 $\tilde{X}$ and $\tilde{Y}$ in $\tilde{\mathcal{X}}$ are said to be comonotone if for no $\omega, \omega' \in \Omega$ we have $\tilde{X}(\omega) \succ \tilde{X}(\omega')$ and $\tilde{Y}(\omega') \succ \tilde{Y}(\omega)$.

Consider the following axiom of Schmeidler [23]:

AXIOM A5'–Comonotone Independence: For all pairwise comonotone $\tilde{X}, \tilde{Y}, \tilde{Z} \in \tilde{\mathcal{X}}$,

$$
\tilde{X} \succeq \tilde{Y} \iff \lambda \tilde{X} + (1 - \lambda)\tilde{Z} \succeq \lambda \tilde{Y} + (1 - \lambda)\tilde{Z} \ \text{ for all } \lambda \in]0, 1[.
$$

Axiom A5' is a stronger version of Axiom A5, that is, Axiom A5' implies Axiom A5. Then we state the following theorem:

Theorem 6.3 *Suppose the probability space $(\Omega, \mathcal{F}, P)$ is rich. A preference order satisfies A1-A4, A5' and A6-A8 if and only if there exists a normalized, monotone and submodular set function $v : \mathcal{F} \to [0, 1]$ that is continuous from above with a core in $\mathcal{Q}$ and $\gamma \in]0, \infty]$ such that*

$$
\tilde{X} \succeq \tilde{Y} \Leftrightarrow \gamma \log\left(\int \int \exp\left\{\frac{-x}{\gamma}\right\} \tilde{X}(., dx)dv\right) \geq \gamma \log\left(\int \int \exp\left\{\frac{-x}{\gamma}\right\} \tilde{Y}(., dx)dv\right),
$$

where as usual the case that $\gamma = \infty$ is identified with $\gamma \log(x) = x = e^{x/\gamma}$. In particular, for $X, Y \in L^\infty$ and $\gamma \in]0, \infty]$ we have

$$
X \succeq Y \Leftrightarrow e_{\gamma,v}(X) \geq e_{\gamma,v}(Y).
$$

Proof. The 'if' direction is straightforward using (6.1) and Proposition 3.12. Let us prove the 'only if' direction. Define $\tilde{X} \succeq^* \tilde{Y}$ if and only if $\tilde{Y} \succeq \tilde{X}$. Note that A1-A3 and A5'-A6 jointly imply that the Axioms (i)-(vii) of Schmeidler [23] are satisfied. Thus, there exists a real-valued function u and a normalized and monotone set function $v : \mathcal{F} \to [0, 1]$ such that

$$
\tilde{X} \succeq^* \tilde{Y} \Leftrightarrow \int \int u(x)\tilde{X}(\omega, dx)v(d\omega) \geq \int \int u(x)\tilde{Y}(\omega, dx)v(d\omega).
$$

Hence,

$$
\tilde{X} \succeq \tilde{Y} \Leftrightarrow -\int \int u(x)\tilde{X}(\omega, dx)v(d\omega) \geq -\int \int u(x)\tilde{Y}(\omega, dx)v(d\omega) \tag{6.2}
$$

28

and

$$\tilde{X} \succeq \tilde{Y} \Leftrightarrow -u^{-1}\left(\int\int u(x)\tilde{X}(\omega, dx)v(d\omega)\right) \geq -u^{-1}\left(\int\int u(x)\tilde{Y}(\omega, dx)v(d\omega)\right).$$

By standard arguments, the monotonicity and continuity axioms entail that u is strictly increasing and continuous. Furthermore, by Schmeidler [23], A4 implies that v is submodular which is equivalent to that for all $X \in L^\infty$,

$$\int Xdv = \min_{Q \in M_v} E_Q[X] \text{ for } M_v = \{Q \text{ is additive on } \mathcal{F}|Q(A) \leq v(A)\}.$$

Suppose without loss of generality that $u(1) = 1$ and $u(0) = 0$. Then, Axiom A6 and (6.2) imply that for every decreasing sequence (A_n) of events such that $\bigcap_n A_n = \emptyset$ we have

$$v(A_n) = \int\int u(x)\tilde{X}_n(\omega, dx)v(d\omega) \downarrow u(0) = 0, \tag{6.3}$$

with $\tilde{X}_n = \delta_{I_{A_n}} \downarrow \delta_0$. Clearly (6.3) implies that all $Q \in M_v$ are σ-additive. For $X \in L^\infty$ define

$$\rho(X) := -u^{-1}\left(\int u(X(\omega))v(d\omega)\right) = \phi^{-1}\left(\max_{Q \in M_v} E_Q[\phi(-X)]\right),$$

with $\phi(x) = -u(-x)$. We have for $X, Y \in L^\infty$ that $X \succeq Y$ if and only if $\rho(X) \geq \rho(Y)$. Now suppose that the maximum is attained in a $\bar{Q} \notin \mathcal{Q}$. Then there exists $A \in \mathcal{F}$ such that $P[A] = 0$ and $\bar{Q}[A] > 0$. But then

$$\rho(X - I_A) = \phi^{-1}\left(\max_{Q \in M_v} E_Q[\phi(-(X - I_A))]\right)$$
$$\geq \phi^{-1}\left(E_{\bar{Q}}[\phi(-(X - I_A))]\right) > \phi^{-1}\left(E_{\bar{Q}}[\phi(-X)]\right) = \rho(X).$$

This is a contradiction to the fact that by A3 we have $X \sim X - I_A$. Hence, $M_v \subset \mathcal{Q}$.

Next, notice that $\rho(m) = -m$ for all $m \in \mathbb{R}$. Hence, $X \sim -\rho(X)$ and it follows by A7 that $X + m \sim -\rho(X) + m$. As $\rho(m) = -m$ for all $m \in \mathbb{R}$ we obtain

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

The result now follows upon applying Theorem 4.1, since by A8 $\phi \in C^3(]\phi^{-1}(0), \infty[) = C^3(\mathbb{R}^+)$ and $\phi(\infty) = \infty$. □

Subsequently, let

$$\Psi = \{\psi : [0, 1] \to [0, 1] | \psi \text{ is concave, right-continuous at zero with } \psi(0+) = 0 \text{ and } \psi(1) = 1\}.$$

For $\psi \in \Psi$ and $X \in L^\infty$ we define $E_\psi[X] := \int Xd\psi(P)$. Furthermore, we define

$$e_{\gamma, \psi}(X) := \gamma \log\left(E_\psi\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) =: e_{\gamma, \psi(P)}(X).$$

We state the following proposition:

Proposition 6.4 *For a given $\psi \in \Psi$, $e_{\gamma,\psi}$ is γ-entropy coherent and its entropy dual $e_{\gamma,\psi}^*$ is given by*

$$e_{\gamma,\psi}^*(Q) = \bar{I}_M(Q),$$

where $M = \{Q \in \mathcal{Q}|Q \leq \psi(P)\}$. Furthermore, the dual conjugate of $e_{\gamma,\psi}$, defined by (2.2), is given by

$$\alpha(\bar{P}) = \min_{Q \in M} \gamma H(\bar{P}|Q).$$

Proof. As ψ is concave and right-continuous in zero, the corresponding v, defined by setting $v(A) = \psi(P[A])$ for all $A \in \mathcal{F}$, is submodular and continuous from above. Hence, by (6.1),

$$e_{\gamma,\psi}(X) = \max_{Q \in M} \gamma \log \left(E_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right),$$

where $M = \{Q \in \mathcal{Q}|Q \leq \psi(P)\}$ and the first part of the proposition follows. To see that the second part holds observe that, as ψ is right-continuous in zero, the set M is weakly compact, see Corollary 4.74, Lemma 4.63, and Corollary 4.35 in Föllmer and Schied [9]. Now Corollary 5.3 yields the proof of the second statement. $\qquad\square$

In the remainder of this section, we assume that the probability space is sufficiently rich to support a random variable with a uniform distribution. If ρ is distribution invariant we can identify ρ with a functional ρ' on the space of distributions with bounded support by setting $\rho'(q_X^+) = \rho(X)$, with q_X^+ the upper (right-continuous) quantile function of X. Furthermore, we will identify a function $\psi_Q \in \Psi$ with dQ/dP by setting $\psi_Q(t) = \int_0^t q_{\frac{dQ}{dP}}^+(1-s)ds$. For $X \in L^\infty$ with $X \geq 0$, we have, using Fubini's theorem,

$$E_{\psi_Q}[X] = \int_0^\infty \psi_Q(P[X > t])dt$$

$$= \int_0^\infty \int_0^1 I_{\{F_X(t) \leq 1-s\}} \psi_Q'^+(s)ds\,dt = \int_0^1 q_X^+(1-s)\psi_Q'^+(s)ds \tag{6.4}$$

$$= \int_0^1 q_X^+(1-s)q_{\frac{dQ}{dP}}^+(1-s)ds = \int_0^1 q_X^+(s)q_{\frac{dQ}{dP}}^+(s)ds. \tag{6.5}$$

For a general $X \in L^\infty$, (6.4)-(6.5) hold by translation invariance of $E_{\psi_Q}[\cdot]$. On the other hand, given a function $\psi \in \Psi$, we can define a measure $Q^\psi \in \mathcal{Q}$ by setting $\frac{dQ^\psi}{dP} = \psi'^+(1-U)$. Note that

$$\{Q \ll P\} = \left\{ Q \ll P | \frac{dQ}{dP} \overset{D}{=} \psi'^+(1-U) \text{ for } \psi \in \Psi \right\}, \tag{6.6}$$

where $\frac{dQ}{dP} \overset{D}{=} \psi_Q'^+(1-U)$ indicates that $\frac{dQ}{dP}$ and $\psi_Q'^+(1-U)$ have the same distribution under P. To see (6.6), note first that for every $\psi \in \Psi$, $\psi'^+(1-U)$ defines a density and thus a measure $Q \in \mathcal{Q}$. On the other hand, for every measure $Q \in \mathcal{Q}$, we have $\frac{dQ}{dP} \overset{D}{=} \psi_Q'^+(1-U)$, with $\psi_Q(t) = \int_0^t q_{\frac{dQ}{dP}}^+(1-s)ds$. Therefore (6.6) holds. Now we can identify ρ^* with a function $(\rho^*)' : \Psi \to \mathbb{R}$ by setting

$$(\rho^*)'(\psi) = \rho^*(Q^\psi).$$

Next, we need Lemma 4.55 of Föllmer and Schied [9]:

Lemma 6.5 *For $X \in L^\infty$ and $Y \in L^1$,*

$$\int_0^1 q_X^+(t)q_Y^+(t)dt = \sup_{\bar{X} \overset{D}{=} X} \mathrm{E}\left[\bar{X}Y\right].$$

Then we state the final theorem, presenting the distribution invariant representation for entropy convex measures of risk (including the special case of entropy coherent measures of risk). It extends the well-known distribution invariant representation results for coherent and convex measures of risk (see, for instance, Dana [5]), which arise whenever $\gamma = \infty$.

Theorem 6.6 *Suppose that ρ is γ-entropy convex. Then the following statements are equivalent:*

(i) ρ is distribution invariant.

(ii) ρ^* is distribution invariant and $(\rho^*)'(\psi) = \sup_{X \in L^\infty} \{e_{\gamma,\psi}(X) - \rho(X)\}.$

(iii) $\rho(X) = \sup_{\psi \in \Psi} \{e_{\gamma,\psi}(X) - (\rho^*)'(\psi)\}.$

Proof. (i)$\Rightarrow$(ii): We write

$$\begin{aligned}
\rho^*(Q) &= \sup_{X \in L^\infty} \left\{\gamma \log \left(\mathrm{E}\left[\frac{dQ}{dP}\exp\{-X/\gamma\}\right]\right) - \rho(X)\right\} \\
&= \sup_{X \in L^\infty} \sup_{\bar{X} \overset{D}{=} X} \left\{\gamma \log \left(\mathrm{E}\left[\frac{dQ}{dP}\exp\{-\bar{X}/\gamma\}\right]\right) - \rho(\bar{X})\right\} \\
&= \sup_{X \in L^\infty} \sup_{\bar{X} \overset{D}{=} X} \left\{\gamma \log \left(\mathrm{E}\left[\frac{dQ}{dP}\exp\{-\bar{X}/\gamma\}\right]\right) - \rho(X)\right\} \\
&= \sup_{X \in L^\infty} \left\{\gamma \log \left(\sup_{\bar{X} \overset{D}{=} X}\mathrm{E}\left[\frac{dQ}{dP}\exp\{-\bar{X}/\gamma\}\right]\right) - \rho(X)\right\} \\
&= \sup_{X \in L^\infty} \left\{\gamma \log \left(\int_0^1 q_{\frac{dQ}{dP}}^+(s)q_{\exp\{-X/\gamma\}}^+(s)ds\right) - \rho(X)\right\} \\
&= \sup_{X \in L^\infty} \left\{\gamma \log \left(\mathrm{E}_{\psi_Q}\left[\exp\{-X/\gamma\}\right]ds\right) - \rho(X)\right\} = \sup_{X \in L^\infty} \left\{e_{\gamma,\psi_Q}(X) - \rho(X)\right\},
\end{aligned}$$

where we have used the distribution invariance of ρ in the third, Lemma 6.5 in the fifth, and (6.5) in the sixth equality. In particular, ρ^* is distribution invariant. It follows that

$$(\rho^*)'(\psi) = \rho^*(Q^\psi) = \sup_{X \in L^\infty} \{e_{\gamma,\psi}(X) - \rho(X)\}.$$

(ii)$\Rightarrow$(iii): Similar as (i)$\Rightarrow$(ii).

(iii)$\Rightarrow$(i): Clearly, ρ is distribution invariant. Set $c(Q) = (\rho^*)'(\psi_Q)$. Notice that if $\frac{dQ}{dP} \overset{D}{=}$

$\psi'^+(1-U)$ then by the definition of $(\rho^*)'$ we have that $c(Q) = (\rho^*)'(\psi)$. We write

$$
\begin{aligned}
\rho(X) &= \sup_{\psi \in \Psi} \left\{ e_{\gamma,\psi}(X) - (\rho^*)'(\psi) \right\} \\
&= \sup_{\psi \in \Psi} \left\{ \gamma \log \left(\mathrm{E}_\psi \left[\exp\left\{ -\frac{X}{\gamma} \right\} \right] \right) - (\rho^*)'(\psi) \right\} \\
&= \sup_{\psi \in \Psi} \left\{ \gamma \log \left(\mathrm{E} \left[\psi'^+(1-U) q^+_{\exp\left\{ -\frac{X}{\gamma} \right\}}(U) \right] \right) - (\rho^*)'(\psi) \right\} \\
&= \sup_{\psi \in \Psi} \sup_{Q, \frac{dQ}{dP} \overset{D}{=} \psi'^+(1-U)} \left\{ \gamma \log \left(\mathrm{E} \left[\frac{dQ}{dP} \exp\left\{ -\frac{X}{\gamma} \right\} \right] \right) - (\rho^*)'(\psi) \right\} \\
&= \sup_{\psi \in \Psi} \sup_{Q, \frac{dQ}{dP} \overset{D}{=} \psi'^+(1-U)} \left\{ \gamma \log \left(\mathrm{E} \left[\frac{dQ}{dP} \exp\left\{ -\frac{X}{\gamma} \right\} \right] \right) - c(Q) \right\} \\
&= \sup_{Q \in \mathcal{Q}} \left\{ \gamma \log \left(\mathrm{E} \left[\frac{dQ}{dP} \exp\left\{ -\frac{X}{\gamma} \right\} \right] \right) - c(Q) \right\} = \sup_{Q \in \mathcal{Q}} \left\{ e_{\gamma,Q}(X) - c(Q) \right\},
\end{aligned}
$$

where we applied (6.4) in the third equality. In the fourth equality we used Lemma 6.5. The fifth equality holds by the definition of c and $(\rho^*)'$, and in the sixth equality we applied (6.6). This proves (ii)$\Rightarrow$(i). $\qquad \square$

7 Conclusions

In this paper, we have introduced two new classes of risk measures: entropy coherent and entropy convex measures of risk. The latter class is a natural generalization of the former. We have proved that entropy coherent and entropy convex measures of risk constitute the connection between two dominant theories: (regular and generalized) robust preferences à la Gilboa and Schmeidler [12] in the microeconomic theory of decision under uncertainty, and the theory of convex risk measures in financial mathematics. A variety of representation and duality results has made explicit that entropy coherent and entropy convex measures of risk satisfy many appealing properties. The theory developed in this paper is of a static nature. In future research we intend to develop its dynamic counterpart.

References

[1] ANSCOMBE, F.J. AND R.J. AUMANN (1963). A definition of subjective probability. *Annals of Mathematical Statistics* 34, 199-205.

[2] ARTZNER, PH., F. DELBAEN, J.-M. EBER AND D. HEATH (1999). Coherent measures of risk. *Mathematical Finance* 9, 203-228.

[3] CHERIDITO, P. AND M. KUPPER (2009). Recursiveness of indifference prices and translation-invariant preferences. *Mathematical and Financial Economics* 2, 173-188.

[4] CSISZÁR, I. (1975). *I*-divergence geometry of probability distributions and minimization problems. *Annals of Probability* 3, 146-158.

[5] DANA, R.-A. (2005). A representation result for concave Schur concave functions. *Mathematical Finance* 15, 613-634.

[6] DEPREZ, O. AND H.U. GERBER (1985). On convex principles of premium calculation. *Insurance: Mathematics and Economics* 4, 179-189.

[7] DUNFORD, N. AND J.T. SCHWARTZ (1958). *Linear Operators. Part I: General Theory.* Interscience, New York.

[8] FÖLLMER, H. AND A. SCHIED (2002). Convex measures of risk and trading constraints. *Finance & Stochastics* 6, 429-447.

[9] FÖLLMER, H. AND A. SCHIED (2004). *Stochastic Finance.* 2nd ed., De Gruyter, Berlin.

[10] FRITTELLI, M. AND E. ROSAZZA GIANIN (2002). Putting order in risk measures. *Journal of Banking and Finance* 26, 1473-1486.

[11] GERBER, H.U. (1979). *An Introduction to Mathematical Risk Theory.* S.S. Huebner Foundation Monograph 8, Homewood, Illinois.

[12] GILBOA, I. AND D. SCHMEIDLER (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18, 141-153.

[13] GOLLIER, C. (2001). *The Economics of Risk and Time*, MIT Press, Cambridge.

[14] GOOVAERTS, M.J., F.E.C. DE VYLDER AND J. HAEZENDONCK (1984). *Insurance Premiums*, North-Holland Publishing, Amsterdam.

[15] HANSEN, L.P. AND T.J. SARGENT (2001). Robust control and model uncertainty. *American Economic Review* 91, 60-66.

[16] HANSEN, L.P. AND T.J. SARGENT (2007). *Robustness.* Princeton University Press, Princeton.

[17] HARDY, G.H., J.E. LITTLEWOOD AND G. PÓLYA (1934). *Inequalities.* Cambridge University Press, Cambridge (2nd ed. 1952, reprinted 1978).

[18] HEATH, D. AND H. KU (2004). Pareto equilibria with coherent measures of risk. *Mathematical Finance* 14, 163-172.

[19] HUBER, P.J. (1981). *Robust Statistics.* Wiley, New York.

[20] MANIA, M. AND M. SCHWEIZER (2005). Dynamic exponential utility indifference valuation. *Annals of Applied Probability* 15, 2113-2143.

[21] SAVAGE, L.J. (1954). *The Foundations of Statistics.* Wiley, New York (2nd ed. 1972, Dover, New York).

[22] SCHMEIDLER, D. (1986). Integral representation without additivity. *Proceedings of the American Mathematical Society* 97, 253-261.

[23] SCHMEIDLER, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* 57, 571-587.

[24] STRZALECKI, T. (2008). Axiomatic foundations of multiplier preferences. Mimeo, Harvard University.

[25] WALD, A. (1950). *Statistical Decision Functions.* Wiley, New York.

[26] ZALINESCU, C. (2002). *Convex Analysis in General Vector Spaces.* World Scientific, Singapore.