UvA-DARE (Digital Academic Repository)

Entropy coherent and entropy convex measures of risk

Laeven, R.J.A.; Stadje, M.

DOI

[10.1287/moor.1120.0559](https://doi.org/10.1287/moor.1120.0559)

Publication date

2013

Document Version

Final published version

Published in

Mathematics of operations research

[Link to publication](#)

Citation for published version (APA):

Laeven, R. J. A., & Stadje, M. (2013). Entropy coherent and entropy convex measures of risk. *Mathematics of operations research*, 38(2), 265-293. <https://doi.org/10.1287/moor.1120.0559>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (<https://dare.uva.nl>)

Entropy Coherent and Entropy Convex Measures of Risk

Roger J. A. Laeven

EURANDOM and CentER, Department of Quantitative Economics, University of Amsterdam,
1018 XE Amsterdam, The Netherlands, r.j.a.laeven@uva.nl

Mitja Stadje

Tilburg University and CentER, Department of Econometrics and Operations Research, Tilburg University,
5000 LE Tilburg, The Netherlands, m.a.stadje@uvt.nl

We introduce two subclasses of convex measures of risk, referred to as entropy coherent and entropy convex measures of risk. Entropy coherent and entropy convex measures of risk are special cases of φ -coherent and φ -convex measures of risk. Contrary to the classical use of coherent and convex measures of risk, which for a given probabilistic model entails evaluating a financial position by considering its expected loss, φ -coherent and φ -convex measures of risk evaluate a financial position under a given probabilistic model by considering its normalized expected φ -loss. We prove that (i) entropy coherent and entropy convex measures of risk are obtained by requiring φ -coherent and φ -convex measures of risk to be translation invariant; (ii) convex, entropy convex, and entropy coherent measures of risk emerge as certainty equivalents under variational, homothetic, and multiple priors preferences upon requiring the certainty equivalents to be translation invariant; and (iii) φ -convex measures of risk are certainty equivalents under variational and homothetic preferences if and only if they are convex and entropy convex measures of risk. In addition, we study the properties of entropy coherent and entropy convex measures of risk, derive their dual conjugate function, and characterize entropy coherent and entropy convex measures of risk in terms of properties of the corresponding acceptance sets.

Key words: multiple priors; variational and homothetic preferences; robustness; convex risk measures; exponential utility; relative entropy; translation invariance; convexity; certainty equivalent

MSC2000 subject classification: Primary: 91B06, 91B16, 91B30; secondary: 60E15, 62P05

OR/MS subject classification: Primary: risk, asset pricing, insurance

History: Published online in *Articles in Advance* October 24, 2012.

1. Introduction. Convex measures of risk have played an increasingly important role since their introduction by Föllmer and Schied [17], Frittelli and Gianin [20], and Heath and Ku [31], generalizing Artzner et al. [2]; see also the early work of Deprez and Gerber [14] and Ben-Tal and Teboulle [4, 5], and the more recent Ben-Tal and Teboulle [6] and Ruszczyński and Shapiro [38, 39]. For a given (discounted) financial position X , defined on a measurable space $(\Omega, \mathcal{F})$, a convex risk measure ρ returns the minimal amount of capital the economic agent holding X 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—under additional assumptions on the space of financial positions and on continuity properties of the risk measures; see §2—be represented in the form

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{E_Q[-X] - \alpha(Q)\}, \quad (1)$$

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

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

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

$$\rho(X) = \sup_{Q \in M} E_Q[-X].$$

Under the probabilistic model Q , the esteemed plausibility of which is measured by $\alpha(Q)$, convex measures of risk evaluate the financial position X by considering its expected loss $E_Q[-X]$. This is equivalent to assuming a risk-neutral valuation, given the probabilistic model Q . A more general, and potentially more cautious, approach to evaluating X under the probabilistic model Q consists in considering its normalized expected φ -loss $c_\varphi(X, Q) = \varphi^{-1}(E_Q[\varphi(-X)])$, with φ increasing. The case of a linear φ corresponds to a risk-neutral evaluation,

while a nonlinear and convex φ corresponds to a risk-averse evaluation. The generalized risk measure ρ then takes the form

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{c_\varphi(X, Q) - \theta(Q)\}, \quad (2)$$

where θ is a penalty function defined on probability measures on $(\Omega, \mathcal{F})$. Henceforth, we call risk measures of the form (2) φ -convex measures of risk. They reduce to φ -coherent measures of risk whenever θ is an indicator function that takes the value zero if $Q \in M$ and ∞ otherwise. The central objects of this paper, referred to as entropy convex measures of risk, are obtained by specifying $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in (0, \infty)$, in (2), so that $c_\varphi(X, Q) = \gamma \log(E_Q[\exp\{-X/\gamma\}])$. Entropy coherent measures of risk occur when $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in (0, \infty)$, and θ is an indicator function.

In a related strand of the literature, in decision theory, among the most popular theories for decision under uncertainty is the multiple priors model, postulating that an economic agent evaluates the consequences (payoffs) of a decision alternative (financial position) X , according to

$$U(X) = \inf_{Q \in M} E_Q[u(X)], \quad (3)$$

where $u: \mathbb{R} \rightarrow \mathbb{R}$ is an increasing function, and $M \subset \mathcal{Q}$ is a set of probability measures (priors). The function u , referred to as a utility function, captures the agent's risk aversion, and the set M represents the agent's uncertainty about the correct probabilistic model. Gilboa and Schmeidler [24] establish a preference axiomatization of this *robust* Savage representation, generalizing Savage [40] in the framework of Anscombe and Aumann [1]. The representation of Gilboa and Schmeidler [24], also referred to as *maxmin* expected utility, is a decision-theoretic foundation of the classical decision rule of Wald [43]—see also Huber [32]—that had long seen little popularity outside (robust) statistics.

The multiple priors model is a special case of interest in the class of variational preferences axiomatized by Maccheroni et al. [33]. Under variational preferences, the numerical representation takes the form

$$U(X) = \inf_{Q \in \mathcal{Q}} \{E_Q[u(X)] + \alpha(Q)\}, \quad (4)$$

where α is an ambiguity index (penalty function) on probability measures on $(\Omega, \mathcal{F})$. Multiple priors occurs when α is an indicator function that takes the value zero if $Q \in M$ and ∞ otherwise. Under multiple priors, the degree of ambiguity is reflected by the multiplicity of the priors. Under variational preferences, the degree of ambiguity is reflected by the multiplicity of the priors and the esteemed plausibility of the prior according to the ambiguity index. Recently, Chateauneuf and Faro [10] and, slightly more generally, Cerreia-Vioglio et al. [8] axiomatized a multiplicative analog of variational preferences, henceforth referred to as *homothetic* preferences, represented as

$$U(X) = \inf_{Q \in \mathcal{Q}} \{\beta(Q) E_Q[u(X)]\}, \quad (5)$$

with β a penalty function on probability measures on $(\Omega, \mathcal{F})$; it also includes multiple priors as a special case if only payoffs X are considered with $u(X) \geq 0$ and $\beta(Q) = 1$ if $Q \in M$, and ∞ otherwise.

To measure the “risk” related to a financial position X , the theories of variational and homothetic preferences sketched above would lead to the definition of a loss functional $L(X) = -U(X)$, satisfying

$$L(X) = \sup_{Q \in \mathcal{Q}} \{E_Q[\varphi(-X)] - \alpha(Q)\} \quad \text{and} \quad L(X) = \sup_{Q \in \mathcal{Q}} \{\beta(Q) E_Q[\varphi(-X)]\},$$

respectively, where $\varphi(x) = -u(-x)$. 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 $L(\bar{m}_X) = L(X)$. This number is commonly referred to as the *certainty equivalent* of X , which is a classical notion in decision theory to evaluate X ; see, e.g., Gollier [25]. However, because we are interested in the amount of capital needed to compensate or counterbalance the risk from the financial position X , we will instead look at the negative certainty equivalent of X , m_X , given by $-\bar{m}_X$, satisfying $L(-m_X) = \varphi(m_X) = L(X)$, or equivalently,

$$m_X = \varphi^{-1} \left(\sup_{Q \in \mathcal{Q}} \{E_Q[\varphi(-X)] - \alpha(Q)\} \right) \quad \text{and} \quad m_X = \varphi^{-1} \left(\sup_{Q \in \mathcal{Q}} \{\beta(Q) E_Q[\varphi(-X)]\} \right). \quad (6)$$

The contribution of this paper is twofold. First we derive precise connections between risk measurement using φ -convex measures of risk—(2)—and risk measurement under the theories of variational, homothetic, and multiple priors preferences—(6). Specifically, we prove the following three main results. Clearly, φ -coherent

measures of risk coincide with negative certainty equivalents under multiple priors preferences. In this case, θ in (2) is an indicator function,

$$\theta(Q) = \begin{cases} 0, & \text{if } Q \in M, \\ \infty, & \text{otherwise,} \end{cases}$$

meaning that all probabilistic models considered are esteemed equally plausible. But if θ is not an indicator function, we prove that φ -convex measures of risk are negative certainty equivalents under variational and homothetic preferences if and only if they are convex and entropy convex measures of risk, respectively (Theorem 4.18 and Remark 4.20). In the former case, φ is linear, inducing risk neutrality; in the latter case φ is exponential, inducing risk aversion.

Furthermore we prove that entropy coherent and entropy convex measures of risk are obtained by requiring φ -coherent and φ -convex measures of risk to be translation invariant (Theorem 4.1). It entails that entropy coherent and entropy convex measures of risk are the only convex (hence translation invariant) risk measures of the form (2) with a nonlinear φ , thus allowing for risk aversion. The property of translation invariance is motivated by the interpretation of a risk measure as a minimal amount of risk capital. It ensures that $\rho(X + \rho(X)) = 0$.

We also prove that negative certainty equivalents under variational, homothetic and multiple priors preferences are translation invariant if and only if they are convex, entropy convex, and entropy coherent measures of risk, respectively (Theorem 4.4, Corollary 4.9, and Theorem 4.11). It entails that convex, entropy convex, and entropy coherent measures of risk induce linear (risk-neutral) or exponential (risk-averse) utility functions in the theories of variational, homothetic, and multiple priors preferences. We show furthermore that, under a normalization condition, this characterization remains valid when the condition of translation invariance is replaced by requiring convexity (Theorem 4.15, Corollary 4.16, and Remark 4.17). The mathematical details in the proofs of these three main characterization results are delicate.

The three characterization results identify entropy coherent and entropy convex measures of risk as distinctive and important subclasses of convex measures of risk. Our second contribution, then, is to study the classes of entropy coherent and entropy convex measures of risk in detail. We show that they satisfy many appealing properties. 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 provide characterizations of entropy coherent and entropy convex risk measures in terms of their acceptance sets.

In the traditional setting of Von Neumann and Morgenstern [42], where the probabilistic model is known and given so that simply $U(X) = E[u(X)]$, analogs of (some of) the main characterization results established in this paper are relatively easy to obtain; see Hardy et al. [30, p. 88, Theorem 106], Gerber [22, Chapter 5], and Goovaerts et al. [26, Chapter 3]. It is intriguingly more complicated for the variational, homothetic, and multiple priors preferences 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. In recent work, Cheridito and Kupper [11, Example 3.6.3] suggest (without formal proof) a connection between certainty equivalents in the pure multiple priors model and convex measures of risk. They restrict, however, to a specific and simple probabilistic setting which, as we will see below, can be viewed as supplementary (and nonoverlapping) to a special case of the general setting considered here. While there is a rich literature on both theories (1) and (6), to the best of our knowledge, there is no other work establishing precise connections between these prominent paradigms, let alone between the more general (2) and (6).

The rest of this paper is organized as follows: in §2, we review some preliminaries for coherent and convex measures of risk. In §3, we introduce φ -coherent and φ -convex measures of risk, and their special cases referred to as entropy coherent and entropy convex measures of risk, provide some motivating examples, and discuss some of their basic properties. In §4 we prove axiomatic characterization results for convex, entropy convex, and entropy coherent measures of risk. Section 5 studies the dual conjugate function for entropy coherent and entropy convex measures. Section 6 provides characterizations of entropy coherent and entropy convex measures of risk in terms of acceptance sets. Conclusions are in §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 $\|X\|_\infty := \inf\{c > 0 \mid P[|X| \leq c] = 1\} < \infty$, where two random variables are identified if they are P -almost surely equal. We denote $(0, \infty)$ by $\mathbb{R}^+$ and $(-\infty, 0]$ by $\mathbb{R}_0^-$. We assume that a riskless asset exists and let X represent a discounted financial position.

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

- Normalization: $\rho(0) = 0$;
- Monotonicity: If $X \leq Y$, then $\rho(X) \geq \rho(Y)$;
- Translation invariance: $\rho(X + m) = \rho(X) - m$ for all $m \in \mathbb{R}$;
- 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 $\alpha: \mathcal{Q} \rightarrow \mathbb{R} \cup \{\infty\}$, referred to as the *dual conjugate* of ρ , such that the following dual representation holds:

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{E_Q[-X] - \alpha(Q)\}. \quad (7)$$

Furthermore,

$$\alpha(Q) = \sup_{X \in L^\infty} \{E_Q[-X] - \rho(X)\}; \quad (8)$$

α is minimal in the sense that for every other (possibly nonconvex or nonlower-semicontinuous) function α' satisfying (7), $\alpha \leq \alpha'$; see, for instance, Föllmer and Schied [18] and Ruszczyński and Shapiro [38, 39]. We define the subdifferential of ρ by

$$\partial \rho(X) = \{Q \in \mathcal{Q} \mid \rho(X) = E_Q[-X] - \alpha(Q)\}. \quad (9)$$

We say that ρ is subdifferentiable if for every $X \in L^\infty$, $\partial \rho(X) \neq \emptyset$. 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 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.

3. Entropy coherence and entropy convexity: Definitions and basic properties. From the representation of convex risk measures (1) it is apparent that under every probabilistic model Q , the financial position X is evaluated risk neutrally, that is, by considering its expected loss $E_Q[-X]$. We now consider, more generally, φ -convex measures of risk, which allow for a risk-averse evaluation of X , given the probabilistic model Q . Specifically, given Q , φ -convex measures of risk evaluate the financial position X by considering its normalized expected φ -loss $c_\varphi(X, Q) = \varphi^{-1}(E_Q[\varphi(-X)])$, with $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ increasing. If φ is linear, the evaluation under Q is risk neutral, but if φ is nonlinear and convex, the evaluation is risk averse. We state the following definitions:

DEFINITION 3.1. The mapping $\rho: L^\infty \rightarrow \mathbb{R}$ is a φ -convex measure of risk if there exists a strictly increasing and continuous function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ and a penalty function $\theta: \mathcal{Q} \rightarrow [0, \infty]$, with $\inf_{Q \in \mathcal{Q}} \theta(Q) = 0$, such that

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{c_\varphi(X, Q) - \theta(Q)\}, \quad (10)$$

with $c_\varphi(X, Q) = \varphi^{-1}(E_Q[\varphi(-X)])$.

DEFINITION 3.2. We call a mapping $\rho: L^\infty \rightarrow \mathbb{R}$ a φ -coherent measure of risk if there exists a strictly increasing and continuous function $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ and a set $M \subset \mathcal{Q}$ such that

$$\rho(X) = \sup_{Q \in \mathcal{Q}} c_\varphi(X, Q), \quad (11)$$

with $c_\varphi(X, Q) = \varphi^{-1}(E_Q[\varphi(-X)])$.

For a mapping $\rho: L^\infty \rightarrow \mathbb{R}$ we define

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

and

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

LEMMA 3.3. If ρ is a φ -convex measure of risk, then for every $X \in L^\infty$,

$$\rho^{**}(X) \leq \rho(X). \quad (12)$$

The next proposition establishes a basic duality result for φ -convex measures of risk:

PROPOSITION 3.4. A normalized mapping ρ is a φ -convex measure of risk if and only if $\rho^{**} = \rho$. Furthermore, ρ^* is the minimal penalty function.

PROOF. This duality result follows in principle from the general duality results in Moreau [35]. We provide a short proof to be self-contained. The “if” part holds because if $\rho(X) = \rho^{**}(X) = \sup_{Q \ll P} \{c_\varphi(X, Q) - \rho^*(Q)\}$, 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 a φ -convex measure of risk. Let us prove the “only if” direction. We already know from Lemma 3.3 that $\rho^{**} \leq \rho$. We will prove that $\rho^{**} \geq \rho$. If ρ is a φ -convex measure of risk, there exists a penalty function θ such that

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{c_\varphi(X, Q) - \theta(Q)\}.$$

Thus, for every $Q \ll P$ we have $\theta(Q) \geq c_\varphi(X, Q) - \rho(X)$. By the definition of ρ^* this yields $\theta(Q) \geq \rho^*(Q)$. This proves that every penalty function θ is dominating ρ^* . Moreover,

$$\rho^{**}(X) = \sup_{Q \ll P} \{c_\varphi(X, Q) - \rho^*(Q)\} \geq \sup_{Q \ll P} \{c_\varphi(X, Q) - \theta(Q)\} = \rho(X). \quad \square$$

Proposition 3.4 suggests a way to find out whether a mapping ρ is a φ -convex measure of risk: compute ρ^* and ρ^{**} , and verify whether $\rho^{**} = \rho$.

For later reference, we state the following definition:

DEFINITION 3.5. For a φ -convex measure of risk ρ we denote by

$$\begin{aligned} \partial_{\#}\rho(X) &= \{Q \in \mathcal{Q} \mid \rho(X) = c_\varphi(X, Q) - \theta(Q)\} \quad \text{and} \\ \partial_{\#}\rho^*(Q) &= \{X \in L^\infty \mid \rho^*(Q) = c_\varphi(X, Q) - \rho(X)\} \end{aligned}$$

the $\#$ -subdifferentials. Furthermore, if for every $X \in L^\infty$, $\partial_{\#}\rho(X) \neq \emptyset$, then we say that ρ is $\#$ -subdifferentiable. We define $\#$ -subdifferentiability of ρ^* similarly.

A risk measure that is particularly popular in insurance and financial mathematics (Gerber [22], Föllmer and Schied [18] and Mania and Schweizer [34]), macroeconomics (Hansen and Sargent [28, 29]), and decision theory (Gollier [25] and Strzalecki [41]), is the (standard) entropic risk measure defined by

$$e_\gamma(X) = \gamma \log \left(\mathbb{E} \left[\exp \left\{ -\frac{X}{\gamma} \right\} \right] \right), \quad \gamma \in \mathbb{R}^+,$$

with $e_0(X) = \lim_{\gamma \downarrow 0} e_\gamma(X) = -\text{ess inf } X$ and $e_\infty(X) = \lim_{\gamma \uparrow \infty} e_\gamma(X) = -\mathbb{E}[X]$. In a setting with distribution invariance, it is commonly referred to as the exponential premium; see Gerber [22] and Goovaerts et al. [27]. It occurs as a special case of $c_\varphi(X, P)$ by taking $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in \mathbb{R}^+$, and is the negative certainty equivalent of an economic agent with exponential (CARA) expected utility preferences. As is well known (Csiszár [12]),

$$e_\gamma(X) = \sup_{\bar{P} \ll P} \{\mathbb{E}_{\bar{P}}[-X] - \gamma H(\bar{P} \mid P)\},$$

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

$$H(\bar{P} \mid P) = \begin{cases} \mathbb{E}_{\bar{P}} \left[\log \left(\frac{d\bar{P}}{dP} \right) \right], & \text{if } \bar{P} \ll P, \\ 0, & \text{otherwise.} \end{cases}$$

The relative entropy is also known as the Kullback-Leibler divergence; it measures the distance between the measures $\bar{P}$ 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 $\gamma \rightarrow e_\gamma(X)$ 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 $e_0(X) = -\text{ess inf } X$, 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 $e_\infty(X) = -E[X]$, which corresponds to a maximal level of trust in the measure P .

Hence, on the one hand, $e_\gamma(X)$ has the interpretation of being the negative certainty equivalent of an economic agent with exponential expected utility preferences and coefficient of absolute risk aversion equal to γ . On the other hand, $e_\gamma(X)$ may be seen as a *robust* expectation with respect to a reference measure P , with the relative entropy as distance measure. It is well known that $\partial e_\gamma(X)$ is given by the *Esscher density* with respect to P : $\exp\{-X/\gamma\}/E[\exp\{-X/\gamma\}]$, $\gamma \in \mathbb{R}^+$.

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

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

Consider now the following two examples:

EXAMPLE 3.6. An economic agent with an exponential (CARA) utility function $u(x) = 1 - e^{-x/\gamma}$, $\gamma \in \mathbb{R}^+$, computes the certainty equivalent to the financial position X under the reference measure P ; that is, $u^{-1}(E[u(X)]) = -\gamma \log(E[\exp\{-X/\gamma\}])$. This evaluation of X coincides, upon a sign change, with the normalized expected φ -loss $c_\varphi(X, P)$ when $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in \mathbb{R}^+$. 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 , including an additive penalty function $\theta(Q)$ measuring the esteemed plausibility of the probabilistic model under Q . The robust certainty equivalent thus computed is

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

Upon a sign change, it is apparent that in this case $\rho(X)$ is a φ -convex measure of risk with risk-averse φ -loss function $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in \mathbb{R}^+$.

EXAMPLE 3.7. Suppose that the economic 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 (Rockafellar and Uryasev [36] and Rockafellar et al. [37]). $TV@R$ 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 : $q_X^+(\lambda) = \inf\{x \mid P[X \leq x] \geq \lambda\}$. If the distribution of X is continuous, $TV@R^\alpha(X) = E[-X \mid X \leq q_X^+(\alpha)]$, 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

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

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

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

i.e., $TV@R^\alpha(X) = E_Q[-X]$, and, for continuous distributions, $Q = P[\cdot \mid X \leq q_X^+(\alpha)]$. 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

$$\begin{aligned}\rho(X) &= \sup_{\bar{P} \ll Q} \sup_{Q \in M_\alpha} \{E_{\bar{P}}[-X] - \gamma H(\bar{P} | Q)\} = \sup_{Q \in M_\alpha} \sup_{\bar{P} \ll P} \{E_{\bar{P}}[-X] - \gamma H(\bar{P} | Q)\} = \sup_{Q \in M_\alpha} e_{\gamma, Q}(X) \\ &= \gamma \log \left(\sup_{Q \in M_\alpha} E_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) = \gamma \log \left(TV @ R^\alpha \left(-\exp \left\{ \frac{-X}{\gamma} \right\} \right) \right),\end{aligned}$$

where we have used in the second and third equalities that $H(\bar{P} | Q) = \infty$ if $\bar{P}$ is not absolutely continuous with respect to Q . The risk measure given by $\rho(X) = \gamma \log(TV @ R^\alpha(-\exp\{-X/\gamma\}))$ accounts for tail risk and model uncertainty. Furthermore, it is computationally attractive because all one needs is a reference model P for the payoff X , which in a probabilistic approach seems a mild presumption. One verifies that in this case $\rho(X)$ is a φ -coherent measure of risk with $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in \mathbb{R}^+$.

This motivates the following definitions:

DEFINITION 3.8. The mapping $\rho: L^\infty \rightarrow \mathbb{R}$ is γ -entropy convex if there exists a penalty function $\theta: \mathcal{Q} \rightarrow [0, \infty]$ with $\inf_{Q \in \mathcal{Q}} \theta(Q) = 0$, such that

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{e_{\gamma, Q}(X) - \theta(Q)\}, \quad \gamma \in \mathbb{R}^+. \quad (13)$$

DEFINITION 3.9. We call a mapping $\rho: L^\infty \rightarrow \mathbb{R}$ γ -entropy coherent, $\gamma \in \mathbb{R}^+$, if there exists a set $M \subset \mathcal{Q}$ such that

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

Henceforth, we call a mapping entropy coherent (convex) if there exists a $\gamma \in \mathbb{R}^+$ such that ρ is γ -entropy coherent (convex). Entropy convexity and entropy coherence occur when the φ -loss function in (10) and (11) is given by $\varphi(x) = \exp\{x/\gamma\}$, $\gamma \in \mathbb{R}^+$.

It is straightforward to see (by verifying the five properties listed in §2) that entropy convex risk measures, hence entropy coherent measures of risk as well, are convex risk measures. As a matter of fact, Theorem 4.1 below shows that among the φ -convex measures of risk with a nonlinear φ , thus allowing for risk aversion, and a nontrivial penalty function, entropy convex measures of risk are the only convex (hence translation invariant) risk measures. As we will see later (for example, see Theorem 5.2 below), however, not every convex risk measure is γ -entropy convex.

PROPOSITION 3.10. Suppose that ρ is a γ -entropy coherent risk measure. Then the following statements are equivalent:

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

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

(b) $M \subset \mathcal{Q}$ is weakly compact.

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

PROOF. Let

$$\bar{\rho}(X) = \sup_{Q \in M} E_Q[-X]. \quad (14)$$

First, notice that by Corollary 4.35 in Föllmer and Schied [18] and the translation invariance of $\bar{\rho}$, M being weakly compact is equivalent to the maximum in (14) being attained for every $X < 0$.

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

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

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

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

(b) $\Leftrightarrow$ (c): Corollary 4.35 in Föllmer and Schied [18] 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

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

Being the natural generalizations of the entropic measure of risk, entropy coherent and entropy convex measures of risk share in their analytic tractability in applications. This is illustrated in the following example on risk sharing.

EXAMPLE 3.11. Suppose that there are two economic agents A and B measuring risk using a general entropy convex measure of risk ρ^A and ρ^B with $\gamma^A, \gamma^B \in \mathbb{R}^+$. Let $V = -\rho$, $\bar{e}_{\gamma, Q} = -e_{\gamma, Q}$, and $\bar{\theta} = -\theta$. Suppose that A owns a financial position X^A , and B owns a financial position X^B . We solve explicitly the problem of optimal risk sharing given by

$$\begin{aligned} R^{A,B}(X^A, X^B) &= \sup_{F \in L^\infty} \{V^A(X^A - F + \Pi^F) + V^B(X^B + F - \Pi^F)\} \\ &= \sup_{\bar{F} \in L^\infty} \{V^A(X^A + X^B - \bar{F}) + V^B(\bar{F})\} =: V^A \square V^B(X^A + X^B), \end{aligned}$$

where Π^F is the agreed upon price of the financial derivative (risk transfer) F and where we have set $\bar{F} := F + X^B$. Assume that

$$\partial_{\#} V^A \left(\frac{\gamma^A}{\gamma^A + \gamma^B} (X^A + X^B) \right) \cap \partial_{\#} V^B \left(\frac{\gamma^B}{\gamma^A + \gamma^B} (X^A + X^B) \right) \neq \emptyset. \quad (15)$$

Then we have that

$$R^{A,B}(X^A, X^B) = \inf_{Q \in \mathcal{Q}} \{ \bar{e}_{\gamma^A + \gamma^B, Q}(X^A + X^B) - (\bar{\theta}^A(Q) + \bar{\theta}^B(Q)) \}. \quad (16)$$

Moreover, the optimal risk sharing is attained in the financial derivative $F^* = (\gamma^B/(\gamma^A + \gamma^B))X^A - (\gamma^A/(\gamma^A + \gamma^B))X^B$. To see this, let $X = X^A + X^B$ and $\bar{F}^* = F^* + X^B$. We write

$$\begin{aligned} R^{A,B}(X^A, X^B) &= \sup_{\bar{F} \in L^\infty} \left\{ \inf_{Q \in \mathcal{Q}} (\bar{e}_{\gamma^A, Q}(X - \bar{F}) - \bar{\theta}^A(Q)) + \inf_{Q \in \mathcal{Q}} (\bar{e}_{\gamma^B, Q}(\bar{F}) - \bar{\theta}^B(Q)) \right\} \\ &\leq \inf_{Q \in \mathcal{Q}} \sup_{\bar{F} \in L^\infty} \{ \bar{e}_{\gamma^A, Q}(X - \bar{F}) + \bar{e}_{\gamma^B, Q}(\bar{F}) - (\bar{\theta}^A(Q) + \bar{\theta}^B(Q)) \} \\ &= \inf_{Q \in \mathcal{Q}} \{ \bar{e}_{\gamma^A, Q}(X - \bar{F}^*) + \bar{e}_{\gamma^B, Q}(\bar{F}^*) - (\bar{\theta}^A(Q) + \bar{\theta}^B(Q)) \} \\ &= \inf_{Q \in \mathcal{Q}} \{ \bar{e}_{\gamma^A + \gamma^B, Q}(X) - (\bar{\theta}^A(Q) + \bar{\theta}^B(Q)) \} \\ &= \gamma_A \inf_{Q \in \mathcal{Q}} \left\{ \log \left(\mathbb{E}_Q \left[\exp \left\{ \frac{-X}{\gamma_A + \gamma_B} \right\} \right] \right) - \bar{\theta}^A(Q) \right\} + \gamma_B \inf_{Q \in \mathcal{Q}} \left\{ \log \left(\mathbb{E}_Q \left[\exp \left\{ \frac{-X}{\gamma_A + \gamma_B} \right\} \right] \right) - \bar{\theta}^B(Q) \right\} \\ &= \inf_{Q \in \mathcal{Q}} (\bar{e}_{\gamma^A, Q}(X - \bar{F}^*) - \bar{\theta}^A(Q)) + \inf_{Q \in \mathcal{Q}} (\bar{e}_{\gamma^B, Q}(\bar{F}^*) - \bar{\theta}^B(Q)) \\ &\leq \sup_{\bar{F} \in L^\infty} \left\{ \inf_{Q \in \mathcal{Q}} (\bar{e}_{\gamma^A, Q}(X - \bar{F}) - \bar{\theta}^A(Q)) + \inf_{Q \in \mathcal{Q}} (\bar{e}_{\gamma^B, Q}(\bar{F}) - \bar{\theta}^B(Q)) \right\}, \end{aligned}$$

where the first inequality holds by weak duality. The second equality holds by Borch [7]; see also Barrieu and El Karoui [3]. To verify the fourth equality, note that “ $\geq$ ” clearly holds, while “ $\leq$ ” follows from (15). Since the last line is equal to the term we started with, it follows that all inequalities must be equalities. This shows (16). The assumption on the $\#$ -subdifferentials is satisfied in particular if $\theta^A = \theta^B$, and for every $d \geq 0$, $\{Q \in \mathcal{Q} \mid \theta^A(Q) \leq d\}$ is weakly compact.

4. Axiomatic characterizations. In this section, we present our main characterization results. The characterization results make explicit that entropy coherent and entropy convex measures of risk are distinctive and important subclasses of convex risk measures. Recall that a rich probability space supports a random variable with a uniform distribution.

4.1. φ -convex and entropy convex measures of risk. The following theorem shows that φ -convex measures of risk are translation invariant if and only if φ is linear (classical convex risk measures) or exponential (entropy convex risk measures).

THEOREM 4.1. *Suppose that the probability space is rich, that ρ is a $\#$ -subdifferentiable φ -convex measure of risk, and that $\varphi \in C^3(\mathbb{R})$. Then the following statements are equivalent:*

- (i) ρ is translation invariant.
- (ii) φ is linear or exponential.

PROOF. The direction (ii) $\Rightarrow$ (i) is straightforward. Let us prove (i) $\Rightarrow$ (ii). As ρ is translation invariant, for every $X \in L^\infty$ the derivative of the mapping $g(m) := \rho(X + m) = \sup_{Q \in \mathcal{Q}} \{c_\varphi(X + m, Q) - \theta(Q)\}$ is constant and equal to -1 . From Lemma 7.1 in the Appendix and the chain rule, it then follows that, for $Q \in \partial_\# \rho(X)$,

$$-1 = g'(0) = \frac{-E_Q[\varphi'(-X)]}{\varphi'(\varphi^{-1}(E_Q[\varphi(-X)]))};$$

hence

$$\varphi'(\varphi^{-1}(E_Q[\varphi(-X)])) = E_Q[\varphi'(-X)]. \quad (17)$$

Now assume that φ is not linear and not exponential. By Lemma 7.2, this assumption implies that $\varphi' \circ \varphi^{-1}$ is not linear on the domain of φ^{-1} . As φ is in C^3 , $\varphi' \circ \varphi^{-1}$ is in C^2 on its domain. Now the second derivative of $\varphi' \circ \varphi^{-1}$ cannot be constantly zero on $\text{dom}(\varphi^{-1})$ as $\varphi' \circ \varphi^{-1}$ is not linear. Therefore, there are two cases:

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

Let $\epsilon > 0$ such that $(1 - \epsilon)t > u$. Since the probability space is rich we may choose $X \in L^\infty$ satisfying both of the following two properties:

- (a) $-X \in \varphi^{-1}(((1 - \epsilon)t, t)) \subset \varphi^{-1}(J)$.
- (b) $-X$ is diffuse.

From (a) it follows in particular that $\varphi(-X) \in ((1 - \epsilon)t, t) \subset J$. Since, with $Q \in \partial_\# \rho(X)$, $Q \ll P$, and $-X$ is diffuse under P , we have that $Q[-X = x] = 0$ for every $x \in \varphi^{-1}(J)$. Thus, $-X$ is also diffuse under Q . Finally, let us derive the contradiction. Assume case (i) above. Then

$$\varphi' \circ \varphi^{-1}(E_Q[\varphi(-X)]) > E_Q[\varphi' \circ \varphi^{-1}(\varphi(-X))] = E_Q[\varphi'(-X)],$$

where the inequality holds because of Jensen's inequality for strictly concave functions for the diffuse random variable $\varphi(-X)$ (where we used that $\varphi(-X) \in J$ and the *strict* concavity of $\varphi' \circ \varphi^{-1}$ on J). The (strict) inequality above is a contradiction to (17). Deriving a contradiction in case (ii) can be done similarly. This proves Theorem 4.1. $\square$

REMARK 4.2. Convex measures of risk are translation invariant by definition (see Definition 2.1). We recall that, in financial mathematics, translation invariance is typically motivated by the interpretation of a risk measure on L^∞ as a minimal amount of risk capital, as it ensures that $\rho(X + \rho(X)) = 0$. This induces that $\rho(X)$ can be viewed as the minimal amount of capital that the economic agent holding X is required to commit and add to the financial position to make the position acceptable. We note that translation invariance in this context is, in fact, an assumption of cash-additivity (El Karoui and Ravanelli [16] and Cerreia-Vioglio et al. [9]). It amounts to assuming that there are no frictions on the risk-free asset market and, in particular, that an institution can borrow and lend cash at the same risk-free rate. As $\rho(X)$ needs to be committed and added currently, while the financial position matures in the future, for the above construction of an acceptable position to be valid, it is crucial that cash itself does not carry any risk. This means that there should be a liquidly traded risk-less asset available in which the risk capital can be invested. Typically such an asset would be a zero-coupon bond without credit or liquidity risk.

While assuming that a liquidly traded risk-less asset exists is common in the literature on coherent and convex measures of risk, some interesting recent papers depart from this assumption, by replacing translation invariance by cash-subadditivity; see El Karoui and Ravanelli [16] and Cerreia-Vioglio et al. [9]. For a setting in which the regulator accepts an array of possibly risky securities instead of restricting to cash, see, e.g., Frittelli and Scandolo [21]. For a further discussion on the property of translation invariance, see Artzner et al. [2] and Föllmer and Schied [18]. Note that similar properties for premium principles can be found in, e.g., Deprez and Gerber [14].

Theorem 4.1 shows in particular that entropy convex measures of risk are the only convex risk measures among φ -convex measures of risk with nonlinear φ , thus allowing for risk aversion.

4.2. Homothetic preferences and entropy convex measures of risk. In this and the following subsection, we axiomatize convex, entropy convex, and entropy coherent measures of risk, showing that they emerge as certainty equivalents under variational, homothetic, and multiple priors preferences, respectively, upon requiring the certainty equivalents to be translation invariant. In the characterization theorems (Theorem 4.4, Corollary 4.9, and Theorem 4.11), we consider, more specifically, negative certainty equivalents of the form $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$, with

$$\bar{\rho}(X) = \begin{cases} \sup_{Q \in \mathcal{Q}} \beta(Q) E_Q[-X], \\ \sup_{Q \in M} E_Q[-X], \\ \sup_{Q \in \mathcal{Q}} \{E_Q[-X] - \alpha(Q)\}, \end{cases}$$

respectively. These constitute the negative certainty equivalents in the theories of homothetic, multiple priors, and variational preferences, respectively; cf. (6), and also Section I.3 in Föllmer et al. [19].

We first state the following proposition, which shows that $\bar{\rho}(X) = \sup_{Q \in \mathcal{Q}} \beta(Q) E_Q[-X]$; i.e., the mapping that induces negative certainty equivalents under homothetic preferences through $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$, is characterized (axiomatized) by the properties of monotonicity, convexity, positive homogeneity, and the property that $\bar{\rho}(m) = -m$ for all $m \leq 0$.

PROPOSITION 4.3. *Suppose that $\bar{\rho}: L^\infty \rightarrow \mathbb{R}$ 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 $\beta: \mathcal{Q} \supset M \rightarrow [0, 1]$ with $\sup_{Q \in M} \beta(Q) = 1$, such that for all $X \in L^\infty$ with $X \leq 0$,*

$$\bar{\rho}(X) = \sup_{Q \in M} \beta(Q) E_Q[-X]. \quad (18)$$

Furthermore, if additionally we have $\bar{\rho}(1) = -1$, then $\bar{\rho}$ is a coherent risk measure. In particular, 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 [18]), we may conclude that $\bar{\rho}$ is weak* lower-semicontinuous. Proposition 3.1.2 in Dana [13] implies that

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

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

$$\begin{aligned} \bar{\rho}(X) &= \sup_{X' \in H} E[-X'X] \\ &= \sup_{X' \in H} E[X'] E\left[-\frac{X'}{E[X']} X\right] = \sup_{X' \in H} E[X'] E_{Q^{X'}}[-X], \end{aligned} \quad (19)$$

where in the case that $X' \equiv 0$, we set $0/0 = 1$ and $Q^{X'} = P$. Now set $M = \{Q \in \mathcal{Q} \mid \text{there exists a } \lambda \geq 0 \text{ such that } \lambda(dQ/dP) \in H\}$. Then (19) entails that for all $X \in L^\infty$ with $X \leq 0$,

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

where for $Q \in M$, $\beta(Q) = \sup\{\lambda \geq 0 \mid \lambda(dQ/dP) \in H\}$. This shows (18). 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 $E[X'] = 1$. This entails that $\bar{\rho}$ is a coherent risk measure and by the definition of β we also obtain that $\beta(Q) = 1$ for every $Q \in M$. $\square$

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 $\beta(Q)(dQ/dP)$. We say that an element $X' \in H \subset L_+^1$ is in the subdifferential of $\bar{\rho}$, $\partial\bar{\rho}(X)$, if it attains the supremum in (19), i.e., $\bar{\rho}(X) = E[-X'X]$. We state the following theorem:

THEOREM 4.4. *Suppose that the probability space is rich and that $\bar{\rho}: L^\infty \rightarrow \mathbb{R}$ 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, nonlinear and continuous function satisfying $0 \in \text{closure}(\text{Image}(\varphi))$, $\varphi(\infty) = \infty$, and $\varphi \in C^3((\varphi^{-1}(0), \infty))$. Then the following statements are equivalent:*

- (i) $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is translation invariant and the subdifferential of $\bar{\rho}$ is always nonempty.
- (ii) ρ is γ -entropy convex with $\gamma \in \mathbb{R}^+$, and the #-subdifferential is always nonempty.

REMARK 4.5. In the proof of Theorem 4.4 (see also Corollary 4.9 below), it will become apparent that $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is entropy coherent if and only if $\bar{\rho}$ is a coherent risk measure. In this case, $\bar{\rho}(X) = \sup_{Q \in M} E_Q[-X]$ for a set $M \subset \mathcal{Q}$, and $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is a negative certainty equivalent in the multiple priors model. Furthermore, we will see that the case that $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is entropy convex corresponds to ρ being the negative certainty equivalent under homothetic preferences, with $\bar{\rho}(X) = \sup_{Q \in M} \beta(Q) E_Q[-X]$, where $\beta: M \rightarrow [0, 1]$ can be viewed as a confidence measure. In this case, every probabilistic 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 multiple priors framework. However, if there exists a $Q \in M$ such that $\beta(Q) < 1$, ρ is entropy convex but not entropy coherent. In both cases, φ turns out to be exponential, hence risk averse.

REMARK 4.6. The direction (i) $\Rightarrow$ (ii) in Theorem 4.4 does not hold (not even in the case where we additionally assume that $\bar{\rho}$ is translation invariant as in Corollary 4.9 below) if the probability space is not rich, or if the assumption on the subdifferential of $\bar{\rho}$ is omitted.

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

$$\begin{aligned} \varphi^{-1}(\bar{\rho}(-\varphi(-X))) &= \varphi^{-1}\left(\max_i \varphi(-X(\omega_i))\right) \\ &= \varphi^{-1}(\varphi(-X(\omega_{i_0}))) = -X(\omega_{i_0}) = \bar{\rho}(X). \end{aligned}$$

In particular, $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X))) = \bar{\rho}(X)$ is translation invariant for every function φ that is strictly increasing and continuous. This shows that (i) $\Rightarrow$ (ii) in Theorem 4.4 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 $\bar{\rho}(X) = \text{ess sup}\{-X\}$ satisfies for *every* strictly increasing and continuous function φ that $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X))) = \bar{\rho}(X)$ is a convex risk measure. The equality may be seen to hold as

$$\varphi^{-1}(\bar{\rho}(-\varphi(-X))) = \text{ess sup } \varphi^{-1}(\varphi(-X)) = \text{ess sup}\{-X\} = \bar{\rho}(X).$$

REMARK 4.7. Notice that since φ is positive somewhere and $0 \in \text{closure}(\text{Image}(\varphi))$ we have that $\varphi^{-1}(\delta)$ is well defined for all $\delta > 0$ small enough and we can define $\varphi^{-1}(0) = \lim_{\delta \downarrow 0} \varphi^{-1}(\delta)$. The common condition that $\varphi(\infty) = \infty$ implies that ρ remains loss sensitive.

LEMMA 4.8. *Suppose that $\bar{\rho}: L^\infty \rightarrow \mathbb{R}$ 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*

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

PROOF. Choose $Q \in M$ such that $\beta(Q)Q \in \partial\bar{\rho}(-X)$. Then by (18) and the monotonicity of $\bar{\rho}$

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

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

PROOF OF THEOREM 4.4. (i) $\Rightarrow$ (ii): Since φ is positive somewhere and $0 \in \text{closure}(\text{Image}(\varphi))$, there are two cases:

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

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

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

$$\varphi_z^{-1}(\bar{\rho}(-\varphi_z(-X))) = \varphi^{-1}(\bar{\rho}(-\varphi(-X))) - z = \varphi^{-1}(\bar{\rho}(-\varphi(-X))).$$

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

- If (H1) holds, then $\varphi(0) = 0$ and $\varphi \in C^3((\varphi^{-1}(0), \infty)) = C^3(\mathbb{R}^+)$.
- If (H2) holds, then $\varphi(0) > 0$ and $\varphi \in C^3((\varphi^{-1}(0), \infty)) = C^3((-\infty, \infty))$.

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

$$\varphi(0) \geq 0. \quad (20)$$

Next, let us look at $X \in L^\infty$ such that $X < 0$. By assumption, $\partial \bar{\rho}(-\varphi(-X)) \neq \emptyset$. As $-\varphi(-X) < 0$ (since $\varphi(0) \geq 0$ and φ is strictly increasing), by (18) and the assumption that the subdifferential of $\bar{\rho}$ is always nonempty we have that, for $\beta(Q)Q \in \partial \bar{\rho}(-\varphi(-X))$,

$$\bar{\rho}(-\varphi(-X)) = \beta(Q) E_Q[\varphi(-X)]. \quad (21)$$

Let $X \in L^\infty$ with $X < 0$. Taking the derivative of the function $m \rightarrow \varphi^{-1}(\bar{\rho}(-\varphi(-X + m)))$ it follows from Lemma 7.1 in the Appendix and translation invariance that

$$1 = \frac{\beta(Q) E_Q[\varphi'(-X)]}{\varphi' \circ \varphi^{-1}(\beta(Q) E_Q[\varphi(-X)])};$$

hence

$$\varphi' \circ \varphi^{-1}(\beta(Q) E_Q[\varphi(-X)]) = \beta(Q) E_Q[\varphi'(-X)]. \quad (22)$$

Assume that (i) $\Rightarrow$ (ii) does not hold; i.e., there does not exist p, γ, q such that, for all $x \in (\varphi^{-1}(0), \infty)$, $\varphi(x) = p \exp\{x/\gamma\} + q$. Furthermore, by assumption it cannot hold that $\varphi(x) = px + q$. We will then derive a contradiction to (22). As in the proof of Theorem 4.1 it may be seen that there are two cases:

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

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

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) $-X \in \varphi^{-1}(((1 - \epsilon)t, t)) \subset \varphi^{-1}(J)$.
- (b) $-X$ is diffuse.

From (a) it follows in particular that $\varphi(-X) \in ((1 - \epsilon)t, t) \subset J$. Since, with $\beta(Q)Q \in \partial \bar{\rho}(-\varphi(-X))$, $Q \ll P$, and $-X$ is diffuse under P , we have that $Q[-X = x] = 0$ for every $x \in \varphi^{-1}(J)$. Thus, $-X$ is also diffuse under Q . As, by (a) and (20), $\varphi(-X) \in J \subset \mathbb{R}^+$ and $\varphi(0) \geq 0$, we have that $\varphi(-X) > 0$. Since $\beta(Q)Q \in \partial \bar{\rho}(-\varphi(-X))$, Lemma 4.8 gives

$$\beta(Q) \geq \frac{\text{ess inf } \varphi(-X)}{\text{ess sup } \varphi(-X)} \geq \frac{(1 - \epsilon)t}{t} = 1 - \epsilon > 0.$$

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

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

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

$$\begin{aligned} \varphi' \circ \varphi^{-1}(\beta(Q) E_Q[\varphi(-X)]) &> E_Q[\varphi' \circ \varphi^{-1}(\beta(Q)\varphi(-X))] \\ &= \lim_{\delta \downarrow 0} E_Q[\varphi' \circ \varphi^{-1}(\beta(Q)\varphi(-X) + (1 - \beta(Q))\delta)] \\ &\geq \liminf_{\delta \downarrow 0} \{E_Q[\beta(Q)\varphi' \circ \varphi^{-1}(\varphi(-X))] + (1 - \beta(Q))\varphi' \circ \varphi^{-1}(\delta)\} \\ &= \left\{ \beta(Q) E_Q[\varphi'(-X)] + (1 - \beta(Q)) \liminf_{\delta \downarrow 0} \varphi' \circ \varphi^{-1}(\delta) \right\} \\ &\geq \beta(Q) E_Q[\varphi'(-X)], \end{aligned}$$

where the first inequality holds because of Jensen's inequality for strictly concave functions for the diffuse random variable $\beta(Q)\varphi(-X)$ (where we used that $\beta(Q)\varphi(-X) \in J$ and the *strict* concavity of $\varphi' \circ \varphi^{-1}$ on J). The second inequality holds by the concavity of the function $\varphi' \circ \varphi^{-1}$ on $(0, t)$. The third inequality holds

because $\varphi' \circ \varphi^{-1}(\delta) > 0$ for every $\delta > 0$ such that $\varphi^{-1}(\delta)$ is well defined, as φ' is positive. The (strict) inequality above is a contradiction to (22), applying to case (i).

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

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

the same argumentation as before yields

$$\begin{aligned} \varphi' \circ \varphi^{-1}(\beta(Q) E_Q[\varphi(-X)]) &< E_Q[\varphi' \circ \varphi^{-1}(\beta(Q)\varphi(-X))] \\ &\leq \lim_n \{E_Q[\beta(Q)\varphi' \circ \varphi^{-1}(\varphi(-X))] + (1 - \beta(Q))\varphi' \circ \varphi^{-1}(\delta_n)\} \\ &= \beta(Q) E_Q[\varphi'(-X)] + (1 - \beta(Q)) \liminf_{\delta \downarrow 0} \varphi' \circ \varphi^{-1}(\delta). \end{aligned} \quad (23)$$

Notice that if

$$(1 - \beta(Q)) \liminf_{\delta \downarrow 0} \varphi' \circ \varphi^{-1}(\delta) = 0, \quad (24)$$

then (23) would imply that

$$\varphi' \circ \varphi^{-1}(\beta(Q) E_Q[\varphi(-X)]) < \beta(Q) E_Q[\varphi'(-X)],$$

which is a contradiction to (22). To see that $(1 - \beta(Q)) \liminf_{\delta \downarrow 0} \varphi' \circ \varphi^{-1}(\delta) = 0$ 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 Proposition 4.3 implies that $\beta(Q) = 1$ and, in particular, (24) 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 $\varphi(-x_0) < 0$. Since by assumption there also exists x_1 such that $\varphi(-x_1) > 0$ the continuity of φ yields that the assumption (H1) above holds. In particular, $\varphi(0) = 0$. By (2) and the positive homogeneity of $\bar{\rho}$, $\bar{\rho}(-\varphi(-x_0)) \neq \varphi(-x_0)$. This gives

$$\varphi^{-1}(\bar{\rho}(-\varphi(-x_0))) \neq -x_0. \quad (25)$$

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

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

which is a contradiction to (25). Hence, $\varphi(x) \geq 0$ for all $x \in \mathbb{R}$ and the assumption (H2) holds; i.e.,

$$\lim_{x \rightarrow -\infty} \varphi(x) = 0. \quad (26)$$

By construction in (H2) we have $\varphi \in C^3(\mathbb{R})$. Now (26) implies that the positive function $\varphi'(x)$ 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

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

Choose $\bar{\delta}_n = \varphi(x_n)$. By (26) we have that $\lim_n \bar{\delta}_n = 0$ and

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

Consequently,

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

This proves (24). Hence, we have derived a contradiction to (22), applying to case (ii). Furthermore, we have seen that the cases (H1) and (1), and (H2) and (2) coincide, respectively.

Hence, (22) implies that the function $\varphi' \circ \varphi^{-1}$ has to be linear, and by Lemma 7.2 this implies that there exist constants $p, \gamma, q \in \mathbb{R}$ such that $\varphi(x) = pe^{x/\gamma} + q$ or $\varphi(x) = px + q$ for all $x \in (\varphi^{-1}(0), \infty)$ (where in case (H1) $\varphi^{-1}(0) = 0$ and in case (H2) $\varphi^{-1}(0) = -\infty$). However, since φ is assumed to be nonlinear, φ has to be exponential.

As φ is strictly increasing and $\varphi(\infty) = \infty$, we must have that $\gamma, p > 0$ and by scale invariance of φ we may set $p = 1$. Now in the case (H2) we must have that $\varphi(x) = \exp\{x/\gamma\}$ (with $q = 0$) as only then does $\lim_{x \rightarrow -\infty} \varphi(x) = 0$.

On the other hand, in the case (H1), condition (1) holds and the second part of Proposition 4.3 implies that $\beta(Q) = 1$ for all $Q \in M$. Therefore, $\varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is invariant under positive affine transformations of φ . Thus, we may *always* assume that $q = 0$. We write

$$\begin{aligned} \varphi^{-1}(\bar{\rho}(-\varphi(-X))) &= \varphi^{-1}\left(\sup_{Q \in M} \beta(Q) E_Q[\varphi(-X)]\right) \\ &= \gamma \log\left(\sup_{Q \in M} \beta(Q) E_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) \\ &= \sup_{Q \in M} \left\{ \gamma \log\left(E_Q\left[\exp\left\{\frac{-X}{\gamma}\right\}\right]\right) + \gamma \log(\beta(Q)) \right\} \\ &= \sup_{Q \in M} \{e_{\gamma, Q}(X) - \theta(Q)\}, \end{aligned}$$

with $\theta(Q) = -\gamma \log(\beta(Q)) \geq 0$ if $Q \in M$ and $\theta(Q) = \infty$ otherwise. Thus, indeed $\varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is γ -entropy convex. As the supremum on the right-hand side of the first equality is always attained because $\partial \bar{\rho}(-\varphi(-X)) \neq \emptyset$, (ii) follows. This completes the proof of the implication (i) $\Rightarrow$ (ii) of Theorem 4.4.

PROOF OF THEOREM 4.4. (ii) $\Rightarrow$ (i): To see the direction (ii) $\Rightarrow$ (i), we let $\varphi(x) = e^{x/\gamma}$, and $\bar{\rho}(X) = \sup_{Q \in \mathcal{Q}} \beta(Q) E_Q[-X]$, with $\beta(Q) = e^{-\rho^*(Q)/\gamma} \geq 0$. Then $\rho(X) = \gamma \log(\bar{\rho}(-e^{-X/\gamma})) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$. 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. $\square$

COROLLARY 4.9. *In the setting of Theorem 4.4, if $\bar{\rho}$ is additionally assumed to be translation invariant, then statement (i) implies that ρ is γ -entropy coherent with $\gamma \in \mathbb{R}^+$.*

PROOF. As $\bar{\rho}$ is assumed to be translation invariant, we have that $\bar{\rho}(m) = -m$ for all $m \in \mathbb{R}$. By Proposition 4.3 this implies that in the proof of Theorem 4.4 we can choose $M \subset \mathcal{Q}$ such that $\beta(Q) = 1$ for all $Q \in M$. Hence, we get $\theta(Q) = \gamma \log(\beta(Q)) = 0$ if $\beta(Q) = 1$ and ∞ otherwise. Thus, indeed $\varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is entropy coherent. $\square$

REMARK 4.10. In recent work, Cheridito and Kupper [11, Example 3.6.3] suggest (without formal proof) a result quite similar to, but essentially different from, Corollary 4.9. Their suggested result can, in a way, be viewed as supplementary to the statement in Corollary 4.9: 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 4.9, we consider a rich outcome space and allow for weakly positive probability measures.

4.3. Variational preferences and convex measures of risk. We state the following theorem:

THEOREM 4.11. *Suppose that the probability space is rich and that $\bar{\rho}: L^\infty \rightarrow \mathbb{R}$ is a convex risk measure with dual conjugate α that has uniformly integrable sublevel sets. Let φ be a strictly increasing and convex function with $\varphi \in C^3(\mathbb{R})$ satisfying either $\varphi(-\infty) = -\infty$ or $\varphi(x)/x \xrightarrow{x \rightarrow \infty} \infty$. Then the following statements are equivalent:*

- (i) $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is translation invariant and the subdifferential of $\bar{\rho}$ is always nonempty.
- (ii) ρ is a convex risk measure and the subdifferential is always nonempty. Furthermore, in the case that $\varphi(x)/x \xrightarrow{x \rightarrow \infty} \infty$, ρ is γ -entropy coherent with $\gamma \in \mathbb{R}^+$ and $\#$ -subdifferentiable.

REMARK 4.12. Note that under the conditions of Theorem 4.11, $\bar{\rho}(X) = \sup_{Q \in \mathcal{Q}} \{E_Q[-X] - \alpha(Q)\}$, so that $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is a negative certainty equivalent under variational preferences. In the proof of Theorem 4.11, we will see that φ is either linear or exponential. In the latter case, α , the dual conjugate of $\bar{\rho}$, only takes the values zero and ∞ and ρ is γ -entropy coherent with $\gamma \in \mathbb{R}^+$. It means that entropy coherent measures of risk are the only convex risk measures among negative certainty equivalents under variational preferences with nonlinear utility.

PROOF OF THEOREM 4.11. The direction (ii) $\Rightarrow$ (i) is straightforward. Let us prove (i) $\Rightarrow$ (ii). Clearly, for $m' \in \mathbb{R}$ we can consider $\varphi(x) + m'$ instead of $\varphi(x)$. Hence, we may assume without loss of generality that $\varphi(0) = \varphi^{-1}(0) = 0$. Now, similarly as in (22), it follows by translation invariance and Lemma 7.1 in the Appendix that, for $Q \in \partial \bar{\rho}(-\varphi(-X))$,

$$\varphi' \circ \varphi^{-1}(\mathbb{E}_Q[\varphi(-X)] - \alpha(Q)) = \mathbb{E}_Q[\varphi'(-X)]. \quad (27)$$

We also need:

LEMMA 4.13. For any $X \in L^\infty$ such that $Q \in \partial \bar{\rho}(X)$,

$$0 \leq \alpha(Q) \leq \text{ess sup}\{-X\} - \text{ess inf}\{-X\}.$$

PROOF. Since $\bar{\rho}(0) = 0$ we must have that $\alpha(Q) \geq 0$. Furthermore, by monotonicity and translation invariance of $\bar{\rho}$,

$$\alpha(Q) = \mathbb{E}_Q[-X] - \bar{\rho}(X) \leq \text{ess sup}\{-X\} - \text{ess inf}\{-X\}. \quad \square$$

CONTINUATION OF THE PROOF OF THEOREM 4.11. (i) $\Rightarrow$ (ii): First, note that as φ is strictly increasing and convex we must have that $\varphi(\infty) = \infty$. Assume now that there does not exist p, γ, q such that, for all $x \in (\varphi^{-1}(-\infty), \infty)$, $\varphi(x) = p \exp\{x/\gamma\} + q$ or $\varphi(x) = px + q$. Let us derive a contradiction to (27). By Lemma 7.2 in the Appendix, this assumption implies that $\varphi' \circ \varphi^{-1}$ is not linear on $(\varphi(-\infty), \infty)$. As φ is in $C^3(\mathbb{R})$ we have that $\varphi' \circ \varphi^{-1}$ is in $C^2((\varphi(-\infty), \infty))$. Now the second derivative of $\varphi' \circ \varphi^{-1}$ cannot be constantly zero on $(\varphi(-\infty), \infty)$ as $\varphi' \circ \varphi^{-1}$ is not linear. Hence, one may see as in the proof of Theorem 4.4 that there are the following two cases:

- (a) There exists a nonempty interval $J = (u, t)$ such that $\varphi' \circ \varphi^{-1}$ is strictly convex on J .
- (b) $\varphi' \circ \varphi^{-1}$ is concave on $(\varphi(-\infty), \infty)$. Furthermore, there exists a nonempty interval $J = (u, t)$ such that $\varphi' \circ \varphi^{-1}$ is strictly concave on J .

Assume case (a). Choose an $\epsilon > 0$ such that $((1 - \epsilon)t, t) \subset J$. Since the probability space is rich we may choose $X \in L^\infty$ satisfying both of the following two properties:

- (a') $-X \in \varphi^{-1}(((1 - (2/3)\epsilon)t, (1 - (1/3)\epsilon)t)) \subset \varphi^{-1}(J)$.
- (b') $-X$ is diffuse.

Similar to the proof of Theorem 4.4, it may be seen that, with $Q \in \partial \bar{\rho}(-\varphi(-X))$, $-X$ is diffuse under Q . From (a') and Lemma 4.13 it follows in particular that $\varphi(-X) - \alpha(Q)$ is in $((1 - \epsilon)t, t) \subset J$. Now let us derive the contradiction. We write

$$\begin{aligned} \varphi' \circ \varphi^{-1}(\mathbb{E}_Q[\varphi(-X)] - \alpha(Q)) &< \mathbb{E}_Q[\varphi' \circ \varphi^{-1}(\varphi(-X) - \alpha(Q))] \\ &\leq \mathbb{E}_Q[\varphi' \circ \varphi^{-1}(\varphi(-X))] = \mathbb{E}_Q[\varphi'(-X)], \end{aligned}$$

where the strict inequality holds because of Jensen's inequality for strictly concave functions for the diffuse random variable $\varphi(-X) - \alpha(Q) \in J$. The second inequality holds since $\alpha(Q) \geq 0$. The (strict) inequality above is a contradiction to (27).

Now assume that case (a) does not hold. Then we are in case (b), and $\varphi' \circ \varphi^{-1}$ is concave on $(\varphi(-\infty), \infty)$. Note that, by assumption, $\varphi' \circ \varphi$ is also increasing and positive (as φ is convex and strictly increasing). Since no nonconstant concave function having domain equal to $\mathbb{R}$ is bounded from below, $\varphi(-\infty) \neq -\infty$. Hence, by our assumptions on φ , we must have that in this case $\lim_{x \rightarrow \infty} (\varphi(x)/x) = \infty$.

Next, note that since the derivative of $\varphi' \circ \varphi^{-1}$ is decreasing and positive it must converge to a constant, say $c \geq 0$. By the monotonicity of $\varphi' \circ \varphi^{-1}$ (as φ is assumed to be convex) there exists a constant $d \in \mathbb{R}$ such that for every $\epsilon > 0$ there exists $M_\epsilon > 0$ such that

$$cx + d \leq \varphi' \circ \varphi^{-1}(x) \leq cx + d + \epsilon, \quad \text{for all } x > M_\epsilon.$$

As $\varphi' \circ \varphi^{-1} = 1/(\varphi^{-1})'$ we get that for any $\epsilon > 0$ there exists a constant $M_\epsilon > 0$ such that for all $x > M_\epsilon$,

$$\frac{1}{cx + d} \leq (\varphi^{-1})'(x) \leq \frac{1}{cx + d - \epsilon}. \quad (28)$$

If $c = 0$, then (28) would imply that φ grows at most linearly, contradicting $\lim_{x \rightarrow \infty} (\varphi(x)/x) = \infty$. Hence, $c > 0$ and (28) implies

$$\varphi^{-1}(M_\epsilon) + \frac{1}{c} \log\left(\frac{cx + d}{cM_\epsilon + d}\right) \leq \varphi^{-1}(x) \leq \varphi^{-1}(M_\epsilon) + \frac{1}{c} \log\left(\frac{cx + d}{cM_\epsilon + d - \epsilon}\right), \quad (29)$$

for all $x \in (M_\epsilon, \infty)$, which yields that

$$\begin{aligned} & \frac{1}{c}((cM_\epsilon + d) \exp\{c(x - \varphi^{-1}(M_\epsilon))\} - d) \geq \varphi(x) \\ & \geq \frac{1}{c}((cM_\epsilon + d - \epsilon) \exp\{c(x - \varphi^{-1}(M_\epsilon))\} - d), \quad \text{for all } x \in (\varphi^{-1}(M_\epsilon), \infty). \end{aligned} \quad (30)$$

From Lemma 4.14 below, we may conclude that (29)–(30) entail that $\bar{\rho}$ must be coherent. Now it follows from Theorem 4.4 (since $\varphi(0) = 0$) that φ must be linear or exponential, which is a contradiction to our starting assumption that this is not the case. Hence, indeed, φ must be linear or exponential. Furthermore, if $\varphi(-\infty) = -\infty$, we must have that φ is linear, while if $\lim_{x \rightarrow \infty} (\varphi(x)/x) = \infty$, φ is exponential.

Now all that is left to show is that if φ has an exponential form, then α , the dual conjugate of $\bar{\rho}$, has to be an indicator function that only takes the values zero and ∞ . In this case ρ would be γ -entropy coherent with $\gamma \in \mathbb{R}^+$. However, note that if φ has an exponential form, then (28) holds even for a certain d and $\epsilon = M_\epsilon = 0$. This implies that (29)–(30) also hold for every $\epsilon > 0$ and every M_ϵ chosen large enough such that $\epsilon < cM_\epsilon + d$. Hence, the fact that in this case α is an indicator function also follows from Lemma 4.14 below. That ρ is $\#$ -subdifferentiable follows directly from the fact that the supremum in $\bar{\rho}$ is attained. This completes the proof. $\square$

LEMMA 4.14. *Suppose Theorem 4.11(i) and that there exist $c > 0$ and $d \in \mathbb{R}$ such that for every $\epsilon > 0$ there exists $M_\epsilon > 0$ such that (29)–(30) hold. Then $\bar{\rho}$ is coherent.*

PROOF. The lemma would be proved if we could show that α , the dual conjugate of $\bar{\rho}$, is an indicator function. Let

$$b(\epsilon) := \frac{cM_\epsilon + d - \epsilon}{cM_\epsilon + d},$$

and denote $b^{-1}(\epsilon) = 1/b(\epsilon)$. Without loss of generality we may assume that M_ϵ converges to ∞ as ϵ tends to zero so that $b(\epsilon)$ tends to one. We will prove the lemma by contradiction. So assume that there exists Q_0 such that $0 < \alpha(Q_0) < \infty$. Let

$$M = \left\{ Q \in \mathcal{Q} \mid \alpha(Q) \leq \frac{\alpha(Q_0)}{2} \right\}. \quad (31)$$

As M is closed and convex, by the Hahn-Banach Theorem there exists an $X'_0 \in L^\infty$ such that

$$\mathbb{E}_{Q_0}[-X'_0] > \sup_{Q \in M} \mathbb{E}_Q[-X'_0].$$

By considering $X_0 := X'_0 + m$ we may, if we choose m suitably, assume that

$$\mathbb{E}_{Q_0}[-X_0] > 0 > \sup_{Q \in M} \mathbb{E}_Q[-X_0].$$

For $\epsilon > 0$ with $\epsilon < cM_\epsilon + d$, let $\lambda^\epsilon := ((cb(\epsilon)\alpha(Q_0) + \epsilon)/(cM_\epsilon + d - \epsilon)) \mathbb{E}_{Q_0}[X_0]^{-1}$. Then

$$\mathbb{E}_{Q_0}[-\lambda^\epsilon X_0] > \frac{cb(\epsilon)\alpha(Q_0)}{cM_\epsilon + d - \epsilon} > 0 > \sup_{Q \in M} \mathbb{E}_Q[-\lambda^\epsilon X_0] \geq \sup_{Q \in M} \left\{ \mathbb{E}_Q[-\lambda^\epsilon X_0] - \frac{cb(\epsilon)\alpha(Q)}{cM_\epsilon + d - \epsilon} \right\},$$

where we used that $\alpha \geq 0$ in the last inequality. Hence,

$$\mathbb{E}_{Q_0}[-\lambda^\epsilon X_0] - \frac{cb(\epsilon)\alpha(Q_0)}{cM_\epsilon + d - \epsilon} > 0 > \sup_{Q \in M} \left\{ \mathbb{E}_Q[-\lambda^\epsilon X_0] - \frac{cb(\epsilon)\alpha(Q)}{cM_\epsilon + d - \epsilon} \right\}. \quad (32)$$

Clearly this inequality also holds for $-\lambda^\epsilon X_0 + \bar{m}$ for any constant $\bar{m} \in \mathbb{R}$. Let us choose a suitable constant so that $-Z_\epsilon = -\lambda^\epsilon X_0 + \bar{m} > 1$, and consequently $\log(-Z_\epsilon)/c$ is well defined and positive. Define

$$Z_\epsilon := \lambda^\epsilon X_0 - \|\lambda^\epsilon X_0\|_\infty - \epsilon - 1.$$

Then $\log(-Z_\epsilon)/c > 0$, and

$$\|Z_\epsilon\|_\infty \leq \frac{cb(\epsilon)\alpha(Q_0) + \epsilon}{cM_\epsilon + d - \epsilon} 2\|X_0\|_\infty \mathbb{E}_{Q_0}[X_0]^{-1} + \epsilon + 1. \quad (33)$$

Let $\tilde{\rho}(X) := \sup_{Q \in \mathcal{Q}} \{E_Q[-X] - (cb(\epsilon)/(cM_\epsilon + d - \epsilon))\alpha(Q)\}$. By assumption, the sublevel sets of α are weakly compact. This entails that for every ϵ with $\epsilon < cM_\epsilon + d$ there exists $Q_\epsilon^* \in \mathcal{Q}$ such that

$$\tilde{\rho}(Z_\epsilon) = E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon}.$$

It follows from (32) that

$$\begin{aligned} E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon} &= \sup_{Q \in \mathcal{Q}} \left\{ E_Q[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q)}{cM_\epsilon + d - \epsilon} \right\} \\ &\geq E_{Q_0}[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q_0)}{cM_\epsilon + d - \epsilon} > \sup_{Q \in M} \left\{ E_Q[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q)}{cM_\epsilon + d - \epsilon} \right\}. \end{aligned}$$

Therefore, $Q_\epsilon^* \notin M$, which by (31) implies that for every $\epsilon > 0$ we have that $\alpha(Q_\epsilon^*) > \alpha(Q_0)/2 > 0$. Next, choose $\epsilon > 0$ small enough such that $c/(cM_\epsilon + d) < 1$, $\epsilon < cM_\epsilon + d$, and

$$\begin{aligned} (b^{-1}(\epsilon) - b(\epsilon)) E_{Q_\epsilon^*}[-Z_\epsilon] &= ((b^{-1}(\epsilon) - 1) + (1 - b(\epsilon))) E_{Q_\epsilon^*}[-Z_\epsilon] \\ &= \left(\frac{\epsilon}{cM_\epsilon + d - \epsilon} + \frac{\epsilon}{cM_\epsilon + d} \right) E_{Q_\epsilon^*}[-Z_\epsilon] \\ &\leq \left(\frac{\epsilon}{cM_\epsilon + d - \epsilon} + \frac{\epsilon}{cM_\epsilon + d} \right) \left(\frac{cb(\epsilon)\alpha(Q_0) + \epsilon}{cM_\epsilon + d - \epsilon} 2\|X_0\|_\infty E_{Q_0}[X_0]^{-1} + \epsilon + 1 \right) \\ &< \frac{c\alpha(Q_0)}{2(cM_\epsilon + d - \epsilon)} < \frac{c\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon}, \end{aligned}$$

where we used (33) in the first inequality. In particular,

$$(b^{-1}(\epsilon) - b(\epsilon)) E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{c\alpha(Q_\epsilon^*)}{cM_\epsilon + d} < 0.$$

Next, choose $m' > 0$ large enough such that

$$(b^{-1}(\epsilon) - b(\epsilon)) E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{c\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon} < -e^{-cm'} \alpha(Q_\epsilon^*).$$

This is equivalent to

$$b^{-1}(\epsilon) \left\{ E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon} \right\} < b(\epsilon) \{ E_{Q_\epsilon^*}[-Z_\epsilon] - e^{-cm'} b^{-1}(\epsilon) \alpha(Q_\epsilon^*) \}. \quad (34)$$

Finally, let us derive a contradiction. We write

$$\begin{aligned} \rho\left(\frac{-\log(-Z_\epsilon)}{c} - \varphi^{-1}(M_\epsilon) - m'\right) &= \varphi^{-1}\left(\tilde{\rho}\left(-\varphi\left(\frac{\log(-Z_\epsilon)}{c} + \varphi^{-1}(M_\epsilon) + m'\right)\right)\right) \\ &\geq \frac{1}{c} \log\left(\frac{1}{cM_\epsilon + d} \left[c\tilde{\rho}\left(-\frac{1}{c}[(cM_\epsilon + d - \epsilon) \right. \right. \right. \\ &\quad \times \exp\{\log(-Z_\epsilon) + c\varphi^{-1}(M_\epsilon) + cm' - c\varphi^{-1}(M_\epsilon)\} - d] + d) \Big] + \varphi^{-1}(M_\epsilon) \\ &= \frac{1}{c} \log\left(\frac{c}{cM_\epsilon + d} \tilde{\rho}\left(\frac{e^{cm'}}{c}(cM_\epsilon + d - \epsilon)Z_\epsilon\right)\right) + \varphi^{-1}(M_\epsilon) \\ &\geq \frac{1}{c} \log\left(\tilde{\rho}\left(\frac{cM_\epsilon + d - \epsilon}{cM_\epsilon + d} e^{cm'} Z_\epsilon\right)\right) + \varphi^{-1}(M_\epsilon) \\ &= \frac{1}{c} \log\left(\sup_{Q \in \mathcal{Q}} \{E_Q[-e^{cm'} b(\epsilon)Z_\epsilon] - \alpha(Q)\}\right) + \varphi^{-1}(M_\epsilon) \\ &= \frac{1}{c} \log\left(e^{cm'} b(\epsilon) \sup_{Q \in \mathcal{Q}} \{E_Q[-Z_\epsilon] - e^{-cm'} b^{-1}(\epsilon) \alpha(Q)\}\right) + \varphi^{-1}(M_\epsilon) \end{aligned}$$

$$\begin{aligned}
&= m' + \frac{1}{c} \log \left(b(\epsilon) \sup_{Q \in \mathcal{Q}} \{E_Q[-Z_\epsilon] - e^{-cm'} b^{-1}(\epsilon) \alpha(Q)\} \right) + \varphi^{-1}(M_\epsilon) \\
&\geq m' + \frac{1}{c} \log(b(\epsilon) \{E_{Q_\epsilon^*}[-Z_\epsilon] - e^{-cm'} b^{-1}(\epsilon) \alpha(Q_\epsilon^*)\}) + \varphi^{-1}(M_\epsilon), \quad (35)
\end{aligned}$$

where we have used (29)–(30) in the first inequality as $\log(-Z_\epsilon)/c + m' + \varphi^{-1}(M_\epsilon) > \log(-Z_\epsilon)/c + \varphi^{-1}(M_\epsilon) > \varphi^{-1}(M_\epsilon)$. In the second equality we have used translation invariance and performed obvious simplifications. Also, in the second inequality we have used that, as $\bar{\rho}$ is convex and $\bar{\rho}(0) = 0$, we must have, for $0 \leq \lambda = c/(cM_\epsilon + d) \leq 1$, that $\lambda \bar{\rho}(X) = \lambda \bar{\rho}(X) + (1 - \lambda) \bar{\rho}(0) \geq \bar{\rho}(\lambda X)$. On the other hand, we obtain

$$\begin{aligned}
&\rho \left(\frac{-\log(-Z_\epsilon)}{c} - \varphi^{-1}(M_\epsilon) \right) \\
&\leq \frac{1}{c} \log \left(\frac{1}{cM_\epsilon + d - \epsilon} \left[c \bar{\rho} \left(-\frac{1}{c} \left[(cM_\epsilon + d) \exp \left\{ c \left(\frac{\log(-Z_\epsilon)}{c} \right) \right\} - d \right] \right) + d \right] \right) + \varphi^{-1}(M_\epsilon) \\
&= \frac{1}{c} \log \left(\frac{c}{cM_\epsilon + d - \epsilon} \bar{\rho} \left(\frac{cM_\epsilon + d}{c} Z_\epsilon \right) \right) + \varphi^{-1}(M_\epsilon) \\
&= \frac{1}{c} \log \left(\frac{c}{cM_\epsilon + d - \epsilon} \sup_{Q \in \mathcal{Q}} \left\{ E_Q \left[-\frac{cM_\epsilon + d}{c} Z_\epsilon \right] - \alpha(Q) \right\} \right) + \varphi^{-1}(M_\epsilon) \\
&= \frac{1}{c} \log \left(b^{-1}(\epsilon) \sup_{Q \in \mathcal{Q}} \left\{ E_Q[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q)}{cM_\epsilon + d - \epsilon} \right\} \right) + \varphi^{-1}(M_\epsilon) \\
&= \frac{1}{c} \log(b^{-1}(\epsilon) \tilde{\rho}(Z_\epsilon)) + \varphi^{-1}(M_\epsilon) \\
&= \frac{1}{c} \log \left(b^{-1}(\epsilon) \left\{ E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon} \right\} \right) + \varphi^{-1}(M_\epsilon), \quad (36)
\end{aligned}$$

where we have used (29)–(30) in the first inequality.

Finally, we may conclude that

$$\begin{aligned}
\rho \left(\frac{-\log(-Z_\epsilon)}{c} - \varphi^{-1}(M_\epsilon) - m' \right) &= m' + \rho \left(\frac{-\log(-Z_\epsilon)}{c} - \varphi^{-1}(M_\epsilon) \right) \\
&\leq m' + \frac{1}{c} \log \left(b^{-1}(\epsilon) \left\{ E_{Q_\epsilon^*}[-Z_\epsilon] - \frac{cb(\epsilon)\alpha(Q_\epsilon^*)}{cM_\epsilon + d - \epsilon} \right\} \right) + \varphi^{-1}(M_\epsilon) \\
&< m' + \frac{1}{c} \log(b(\epsilon) \{E_{Q_\epsilon^*}[-Z_\epsilon] - e^{-cm'} b^{-1}(\epsilon) \alpha(Q_\epsilon^*)\}) + \varphi^{-1}(M_\epsilon) \\
&\leq \rho \left(\frac{-\log(-Z_\epsilon)}{c} - \varphi^{-1}(M_\epsilon) - m' \right), \quad (37)
\end{aligned}$$

where we have used (36) in the first inequality, (34) in the strict inequality, and (35) in the last inequality. The equality holds by translation invariance. The strict inequality (37) is a contradiction. $\square$

4.4. Convexity without the translation invariance axiom. In the previous two subsections the axiom of translation invariance played a key role; see Theorems 4.4(i) and 4.11(i). As is well documented (see, for example, Cheridito and Kupper [11]), 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 a key role.

Throughout this subsection, we suppose the probability space $(\Omega, \mathcal{F}, P)$ is rich. We state the following theorem:

THEOREM 4.15. *Let $\bar{\rho}: L^\infty \rightarrow \mathbb{R}$ 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 $\rho: L^\infty \rightarrow \mathbb{R}$ is defined by $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$, for a strictly increasing, nonlinear, and continuous function $\varphi \in C^3((\varphi^{-1}(0), \infty))$. Finally, suppose that $0 \in \text{closure}(\text{Image}(\varphi))$ and that $\varphi(\infty) = \infty$. Then the following statements are equivalent:*

- (i) ρ is convex and $\rho(m) = -m$ for all $m \in \mathbb{R}$.
- (ii) ρ is γ -entropy convex with $\gamma \in \mathbb{R}^+$.

PROOF. The direction from (ii) to (i) is straightforward. Let us show the reverse direction. First, notice that ρ is continuous with respect to the L^∞ -norm. This can be seen as follows: from the proof of Proposition 4.3, 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 $X, Y \in L^\infty$,

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

Switching the roles of X and Y , it follows that $\bar{\rho}$ is indeed continuous with respect to the L^∞ -norm. Now as φ is continuous we can conclude that ρ is continuous with respect to the L^∞ -norm as well. But, then it follows from Proposition 2.5-(8) in Cheridito and Kupper [11] that ρ is translation invariant. The argument is simple; namely, for $\lambda \in (0, 1)$ we have

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

Letting λ converge to one and using the continuity of ρ with respect to the L^∞ -norm we find that $\rho(X + m) \leq \rho(X) - m$. Replacing X by $X + m$ and m by $-m$ yields the stated result. Therefore, ρ is indeed translation invariant. Now upon application of Theorem 4.4, the direction from (i) to (ii) follows. $\square$

Using Corollary 4.9, we now obtain the following corollary directly:

COROLLARY 4.16. *In the setting of Theorem 4.15, suppose that $\bar{\rho}$ is additionally assumed to be translation invariant. Then the following statements are equivalent:*

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

REMARK 4.17. It is straightforward to adapt the proof of Theorem 4.15 to show that, similarly, the condition of translation invariance in Theorem 4.11 can also be replaced by convexity and the condition that $\rho(m) = -m$ for all $m \in \mathbb{R}$.

4.5. φ -convex measures of risk and homothetic and variational preferences. It is straightforward to verify that φ -coherent measures of risk coincide with negative certainty equivalents under multiple priors preferences, for any φ -loss function. Except for the subclasses of coherent and entropy coherent measures of risk, φ -coherent measures of risk are not translation invariant or convex (cf. Theorem 4.1 and Corollaries 4.9 and 4.16), and therefore do not belong to the class of convex risk measures. Under φ -coherent measures of risk and multiple priors preferences, all probabilistic models $M \subset \mathcal{Q}$ are esteemed equally plausible. The following theorem (remarks) show(s) that, when the penalty function θ is not an indicator function, that is, when the probabilistic models are not considered equally plausible, φ -convex measures of risk are negative certainty equivalents under homothetic (variational) preferences if and only if they are entropy convex (classical convex) measures of risk.

THEOREM 4.18. *Suppose that $\bar{\rho}: L^\infty \rightarrow \mathbb{R}$ is monotone, convex, positively homogeneous, and continuous from above, and for all $m \in \mathbb{R}_0^-$, $\bar{\rho}(m) = -m$. Let φ satisfy $0 \in \text{closure}(\text{Image}(\varphi))$ and $\varphi \in C^2(\mathbb{R})$. Furthermore, suppose that ρ is a φ -convex measure of risk with minimal penalty function θ that is L^1 -continuous, $\#$ -subdifferentiable in its interior, and not an indicator function. Then the following statements are equivalent:*

- (i) $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$, and the subdifferential of $\bar{\rho}$ is always nonempty.
- (ii) ρ is γ -entropy convex with $\gamma \in \mathbb{R}^+$, and the $\#$ -subdifferential is always nonempty.

PROOF. The direction (ii) $\Rightarrow$ (i) is straightforward. Let us prove (i) $\Rightarrow$ (ii). First, note that, as both φ -convex measures of risk and $\varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ are invariant under scaling of φ , we may assume without loss of generality that $\varphi'(0) = 1$. Note further that if $\bar{\rho}(1) = -1$ holds then, by Proposition 4.3, $\bar{\rho}$ would be a coherent risk measure. In particular, in that case there exists a convex closed set $M \subset \mathcal{Q}$ such that for all $X \in L^\infty$,

$$\bar{\rho}(X) = \max_{Q \in M} E_Q[X];$$

hence

$$\rho(X) = \max_{Q \in \mathcal{M}} \varphi^{-1}(\mathbb{E}_Q[\varphi(-X)]).$$

But this entails that ρ^* is given by I_M , which is a contradiction to the assumption that the minimal penalty function of ρ is not an indicator function.

Therefore, we may conclude that $\bar{\rho}(1) \neq -1$. Then, by positive homogeneity, $\bar{\rho}(m) \neq -m$ for all $m > 0$. Since at the same time $\rho(m) = -m$ for all $m \in \mathbb{R}$ as ρ is a φ -convex measure of risk, $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ can only hold if $\varphi \geq 0$. As φ is strictly increasing and 0 is in the closure of the image of φ , we must have that $\varphi > 0$ and $\text{Image}(\varphi) = \mathbb{R}^+$.

Applying the transformation $Y = \varphi(-X)$ and using that ρ is a φ -convex measure of risk, we obtain from $\rho(X) = \varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ that

$$\max_{Q \in \mathcal{Q}} \{\varphi(\varphi^{-1}(\mathbb{E}_Q[Y]) - \theta(Q))\} = \bar{\rho}(-Y). \quad (38)$$

Note that (38) holds only for random variables in the image of φ . Define

$$\bar{\rho}^Q(-Y) := \varphi(\varphi^{-1}(\mathbb{E}_Q[Y]) - \theta(Q)).$$

By (38), for all Y taking values in $\mathbb{R}^+ = \text{Image}(\varphi)$ a.s., we have

$$\bar{\rho}(-Y) = \max_{Q \in \mathcal{Q}} \bar{\rho}^Q(-Y). \quad (39)$$

For every such Y , we denote by Q^Y the measure which attains the maximum in (39).

By positive homogeneity of $\bar{\rho}$, we have for every $\lambda > 0$, $\bar{\rho}(-\lambda Y) = \lambda \bar{\rho}(-Y)$. Note that $Y \in \text{Image}(\varphi) = \mathbb{R}^+$ a.s. if and only if $\lambda Y \in \text{Image}(\varphi)$ a.s. with $\lambda > 0$. Now (38) entails that for all $Y > 0$ and $\lambda > 0$,

$$\bar{\rho}^{Q^Y}(-\lambda Y) \leq \lambda \bar{\rho}^{Q^Y}(-Y).$$

This may be seen as follows: Suppose that $\bar{\rho}^{Q^Y}(-\lambda Y) > \lambda \bar{\rho}^{Q^Y}(-Y)$. Then

$$\bar{\rho}(-\lambda Y) \geq \bar{\rho}^{Q^Y}(-\lambda Y) > \lambda \bar{\rho}^{Q^Y}(-Y) = \lambda \bar{\rho}(-Y),$$

which is a contradiction to the positive homogeneity property of $\bar{\rho}$. Hence, indeed $\bar{\rho}^{Q^Y}(-\lambda Y) \leq \lambda \bar{\rho}^{Q^Y}(-Y)$ for all $Y > 0$ and $\lambda > 0$. Dividing both sides by λ we get $\bar{\rho}^{Q^Y}(-Y) \leq (1/\lambda) \bar{\rho}^{Q^Y}(-\lambda Y) \leq \bar{\rho}^{Q^Y}(-Y)$, which entails that

$$\bar{\rho}^{Q^Y}(-\lambda Y) = \lambda \bar{\rho}^{Q^Y}(-Y),$$

$Y > 0$ and $\lambda > 0$.

Now by assumption, there exists $Q_0 \in \mathcal{Q}$ such that $0 < \theta(Q_0) < \infty$. As $\inf_{Q \in \mathcal{Q}} \theta(Q) = 0$, for all ϵ sufficiently small, there exist measures Q_ϵ such that $\theta(Q_\epsilon) = \epsilon$. By considering the mapping $\lambda \rightarrow \theta(\lambda Q_\epsilon + (1 - \lambda)Q_0)$ for $\lambda \in [0, 1]$ and using L^1 -continuity on the interior of the domain of θ , we may conclude that $(\epsilon, \theta(Q_0)) \subset \text{Image}(\theta)$. As ϵ can be chosen arbitrarily small, we have $(0, \theta(Q_0)) \subset \text{Image}(\theta)$.

Thus, for arbitrary $a \in (0, \theta(Q_0))$ there exists a $Q_a \in \mathcal{Q}$ such that $\theta(Q_a) = a$. As $\partial_\# \theta(Q_0) \neq \emptyset$, there exists a Y_a for which the maximum in (39) is attained in Q_a . For $\lambda \in \mathbb{R}^+$, define the mapping $h(\lambda) = \bar{\rho}^{Q_a}(-\lambda Y_a) = \lambda \bar{\rho}^{Q_a}(-Y_a)$. If we take the derivative of h with respect to λ , then by the definition of $\bar{\rho}_a^Q$ and the chain rule we get that

$$\frac{\varphi'(\varphi^{-1}(\mathbb{E}_{Q_a}[\lambda Y_a]) - a)}{\varphi'(\varphi^{-1}(\mathbb{E}_{Q_a}[\lambda Y_a]))} \mathbb{E}_{Q_a}[Y_a] = h'(\lambda) = \bar{\rho}^{Q_a}(-Y_a),$$

where we applied in the last equation that, by positive homogeneity, $h(\lambda) = \lambda \bar{\rho}^{Q_a}(-Y_a)$. Hence,

$$\frac{\varphi'(\varphi^{-1}(\mathbb{E}_{Q_a}[\lambda Y_a]) - a)}{\varphi'(\varphi^{-1}(\mathbb{E}_{Q_a}[\lambda Y_a]))} = \frac{\bar{\rho}^{Q_a}(-Y_a)}{\mathbb{E}_{Q_a}[Y_a]} =: g(-a). \quad (40)$$

The last definition is justified since Q_a and Y_a only depend on a . Denote $f^a(\lambda) = \varphi^{-1}(\mathbb{E}_{Q_a}[\lambda Y_a])$. Note that f^a is a continuously differentiable and strictly increasing function on $\mathbb{R}^+$ with $\text{Image}(f^a) = \mathbb{R}$. (The latter statement holds as $P[Y_a > 0] = 1$ so that $Q_a[Y_a > 0] = 1$.)

Now, we have shown in (40) that for any $\lambda > 0$,

$$\varphi'(f^a(\lambda) - a) = \varphi'(f^a(\lambda))g(-a).$$

Choosing $\lambda = (f^a)^{-1}(0)$ and using that $\varphi'(0) = 1$, we obtain $g = \varphi'$. Thus,

$$\varphi'(f^a(\lambda) - a) = \varphi'(f^a(\lambda))\varphi'(-a).$$

As $\text{Image}(f^a) = \mathbb{R}$, it follows in particular that for any $h \in \mathbb{R}$,

$$\varphi'(h - a) = \varphi'(h)\varphi'(-a). \quad (41)$$

As φ' is continuous, (41) holds for all $a \in [0, \theta(Q_0)]$ (and not only on the open interval). Now, since φ' is differentiable, it is well known that (41) yields that there exists $\gamma > 0$ such that $\varphi'(-a) = \exp(-a/\gamma)$ for any $a \in [0, \theta(Q_0)]$. (Just look at the differential quotient of $\varphi(-a)$ and apply (41).) Finally, by using a decomposition $x = -b_1\theta(Q_0) - b_2$ for $x \in (-\infty, 0)$ and $x = b_1\theta(Q_0) + b_2$, for $x \in [0, \infty)$ with $b_1 \in \mathbb{N}$ and $b_2 \in [0, \theta(Q_0))$, and applying (41) (or $\varphi'(h - a)/\varphi'(-a) = \varphi'(h)$) multiple times, it may be seen that (41) entails that $\varphi'(x) = \exp(x/\gamma)$ for any $x \in \mathbb{R}$ with $\gamma \in \mathbb{R}^+$. Consequently, $\varphi(x) = \gamma \exp(x/\gamma) + b$ for a constant b . Thus, we may conclude that ρ is γ -entropy convex. This proves the theorem. $\square$

REMARK 4.19. In the setting of Theorem 4.18, $\varphi^{-1}(\bar{\rho}(-\varphi(-X)))$ is a negative certainty equivalent in the sense of (6), under homothetic preferences.

REMARK 4.20. If a φ -convex measure of risk ρ is a negative certainty equivalent under variational preferences, then we must have that

$$\sup_{Q \in \mathcal{Q}} \{\varphi^{-1}(\mathbb{E}_Q[\varphi(-X)]) - \theta(Q)\} = \sup_{Q \in \mathcal{Q}} \{\varphi^{-1}(\mathbb{E}_Q[\varphi(-X)] - \alpha(Q))\}.$$

If we exclude the risk-neutral case that φ is linear, it may be seen by a similar argument as above and under appropriate conditions that this entails that θ is an indicator function. On the contrary, if we exclude the case that θ is an indicator function, then φ must be linear, in which case ρ is a convex risk measure.

5. The dual conjugate. In this section we study the dual conjugate function, defined in (8), for entropy coherent and entropy convex measures of risk. Quite unusually, some explicit results on the dual conjugate function can be obtained. Let $\gamma \in \mathbb{R}^+$. We state the following proposition:

PROPOSITION 5.1. Suppose that ρ is γ -entropy convex. Then

$$\rho^*(Q) = \sup_{\bar{P} \ll P} \{\alpha(\bar{P}) - \gamma H(\bar{P} | Q)\}. \quad (42)$$

PROOF. We write

$$\begin{aligned} \rho^*(Q) &= \sup_{X \in L^\infty} \{e_{\gamma, Q}(X) - \rho(X)\} = \sup_{X \in L^\infty} \sup_{\bar{P} \ll P} \{\mathbb{E}_{\bar{P}}[-X] - \gamma H(\bar{P} | Q) - \rho(X)\} \\ &= \sup_{\bar{P} \ll P} \sup_{X \in L^\infty} \{\mathbb{E}_{\bar{P}}[-X] - \rho(X) - \gamma H(\bar{P} | Q)\} = \sup_{\bar{P} \ll P} \{\alpha(\bar{P}) - \gamma H(\bar{P} | Q)\}. \quad \square \end{aligned}$$

Notice that (42) yields that $\alpha(\bar{P}) \leq \rho^*(Q) + \gamma H(\bar{P} | Q)$. Hence,

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

The next penalty function duality theorem will show that this inequality under additional assumptions is actually an equality. It also establishes the explicit relationship between the dual conjugate α and the penalty function θ for γ -entropy convex measures of risk.

THEOREM 5.2. Suppose that ρ is γ -entropy convex with penalty function θ . Then:

(i) The dual conjugate of ρ , defined in (8), is given by the largest convex and lower-semicontinuous function α being dominated by $\inf_{Q \in \mathcal{Q}} \{\gamma H(\bar{P} | Q) + \theta(Q)\}$.

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

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

$$\alpha(\bar{P}) = \min_{Q \in \mathcal{Q}} \{\gamma H(\bar{P} | Q) + \theta(Q)\}. \quad (43)$$

PROOF. (i): We write

$$\begin{aligned}\rho(X) &= \sup_{Q \in \mathcal{Q}} \left\{ \gamma \log \left(E_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) - \theta(Q) \right\} = \sup_{Q \in \mathcal{Q}} \sup_{\bar{P} \ll P} \{ E_{\bar{P}}[-X] - \gamma H(\bar{P} | Q) - \theta(Q) \} \\ &= \sup_{\bar{P} \ll P} \sup_{Q \in \mathcal{Q}} \{ E_{\bar{P}}[-X] - \gamma H(\bar{P} | Q) - \theta(Q) \} = \sup_{\bar{P} \ll P} \left\{ E_{\bar{P}}[-X] - \inf_{Q \in \mathcal{Q}} \{ \gamma H(\bar{P} | Q) + \theta(Q) \} \right\},\end{aligned}$$

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 (7), statement (i) follows.

(ii): Now assume that θ 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 $X \in L^\infty$, $-\gamma \log(E_Q[\exp\{-X/\gamma\}])$ is convex in Q , and $E_{\bar{P}}[-X]$ is convex in $\bar{P}$. Hence, $E_{\bar{P}}[-X] - \gamma \log(E_Q[\exp\{-X/\gamma\}])$ is jointly convex in $(\bar{P}, Q)$, and therefore

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

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

To see (b), note that 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 $E_{Q_n}[-X] \xrightarrow{n \rightarrow \infty} E_Q[-X]$ and $E_{\bar{P}_n}[-X] \xrightarrow{n \rightarrow \infty} E_{\bar{P}}[-X]$. Since

$$\begin{aligned}E_{\bar{P}}[-X] - \gamma \log \left(E_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) &= \lim_n \left\{ E_{\bar{P}_n}[-X] - \gamma \log \left(E_{Q_n} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\} \\ &\leq \liminf_n \sup_{X \in L^\infty} \left\{ E_{\bar{P}_n}[-X] - \gamma \log \left(E_{Q_n} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\},\end{aligned}$$

it follows that

$$\begin{aligned}\gamma H(\bar{P} | Q) &= \sup_{X \in L^\infty} \left\{ E_{\bar{P}}[-X] - \gamma \log \left(E_Q \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\} \\ &\leq \liminf_n \sup_{X \in L^\infty} \left\{ E_{\bar{P}_n}[-X] - \gamma \log \left(E_{Q_n} \left[\exp \left\{ \frac{-X}{\gamma} \right\} \right] \right) \right\} \\ &= \liminf_n \gamma H(\bar{P}_n | Q_n).\end{aligned}$$

This proves (b).

(a) and (b) imply that $\gamma H(\bar{P} | Q)$ is jointly convex and lower-semicontinuous in $(\bar{P}, Q)$. Furthermore, $\theta(Q)$ is convex and lower-semicontinuous. Therefore $\gamma H(\bar{P} | Q) + \theta(Q)$ is jointly convex and lower-semicontinuous as well. By Theorem 2.1.3(v) of Zalinescu [45] it follows that $\inf_{Q \in \mathcal{Q}} \{ \gamma H(\bar{P} | Q) + \theta(Q) \}$ is convex in $\bar{P}$. Now (ii) follows since α is the minimal lower-semicontinuous and convex function satisfying (7).

(iii): If we could show that

$$\beta(\bar{P}) = \inf_{Q \in \mathcal{Q}} \{ \gamma H(\bar{P} | Q) + \theta(Q) \} \quad (44)$$

is also lower-semicontinuous and that the infimum is attained, then (43) would follow from the uniqueness of α . First, let us show that the infimum in (44) is attained. Let $Q_k \ll P$ be the minimizing sequence. Since $\theta \neq \infty$ we have for all $\bar{P}$ that $\beta(\bar{P}) < \infty$. Thus,

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

In particular, $(\theta(Q_k))_k$ 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 [15]), the sequence Q_k is weakly relatively compact. Hence, for fixed $\bar{P}$ we may take the infimum in (44) over the weakly compact set $\{Q_1, Q_2, \dots\}$. As by (b) above, $Q \rightarrow \gamma H(\bar{P} | Q) + \theta(Q)$ 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

$$\beta(\bar{P}) \leq \liminf_n \beta(\bar{P}_n). \quad (45)$$

If $\liminf_n \beta(\bar{P}_n) = \infty$, then clearly (45) 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

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

As $\limsup_j \theta(Q_{n_j}) \leq \lim_j \gamma H(\bar{P}_{n_j} | Q_{n_j}) + \theta(Q_{n_j}) = r$, 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 $(\bar{P}, Q) \rightarrow H(\bar{P} | Q)$ proved in (b),

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

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 $\beta = \alpha$. $\square$

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

$$\alpha(\bar{P}) = \min_{Q \in M} \gamma H(\bar{P} | Q). \quad (46)$$

PROOF. The first part of the corollary is precisely (ii) of Theorem 5.2 with $\theta = \bar{I}_M$. The second part follows as for all $r \in \mathbb{R}^+$ we have $\{Q \in \mathcal{Q} \mid \theta(Q) \leq r\} = M$. Equation (46) 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

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

Then ρ is γ -entropy coherent if and only if $\rho(X) = e_{\gamma}(X)$.

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 (7), indeed

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

COROLLARY 5.5. Let ρ be a convex risk measure. Then the following statements are equivalent:

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

$$\alpha(\bar{P}) = \min_{Q \in \mathcal{Q}} \{\gamma H(\bar{P} | Q) + \theta(Q)\}. \quad (47)$$

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

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

$$\begin{aligned} \rho(X) &= \sup_{\bar{P} \in \mathcal{Q}} \{E_{\bar{P}}[-X] - \alpha(\bar{P})\} = \sup_{\bar{P} \in \mathcal{Q}} \left\{ E_{\bar{P}}[-X] - \min_{Q \in \mathcal{Q}} [\gamma H(\bar{P} | Q) + \theta(Q)] \right\} \\ &= \sup_{Q \in \mathcal{Q}} \sup_{\bar{P} \in \mathcal{Q}} \{E_{\bar{P}}[-X] - \gamma H(\bar{P} | Q) + \theta(Q)\} = \sup_{Q \in \mathcal{Q}} \{e_{\gamma, Q}(X) - \theta(Q)\}. \quad \square \end{aligned}$$

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

- (i) ρ is γ -entropy convex with ρ^* having uniformly integrable sublevel sets.
- (ii) θ is convex and lower-semicontinuous with $\inf_{Q \in \mathcal{Q}} \theta(Q) = 0$ and uniformly integrable sublevel sets, and for every $\bar{P} \in \mathcal{Q}$,

$$\begin{aligned} & \inf_{Q \in \mathcal{Q}} \sup_{\hat{P} \in \mathcal{Q}} \{\gamma H(\bar{P} | Q) + \alpha(\hat{P}) - \gamma H(\hat{P} | Q)\} \\ &= \sup_{\hat{P} \in \mathcal{Q}} \inf_{Q \in \mathcal{Q}} \{\gamma H(\bar{P} | Q) + \alpha(\hat{P}) - \gamma H(\hat{P} | Q)\}. \end{aligned} \quad (48)$$

PROOF. We can write the right-hand side of (48) 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 $d\bar{P}/d\hat{P} \neq 1$ on a nonzero set we have that $\log(d\bar{P}/d\hat{P}) < 0$ on a nonzero 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 (48) is equal to $\alpha(\bar{P})$. Moreover, for the left-hand side we have that

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

Now the theorem follows from Proposition 5.1 and Corollary 5.5. $\square$

6. Acceptance sets. Throughout this section we assume that the probability space is rich. We further denote by F_X^Q the cumulative distribution function (cdf) of X under the probabilistic model Q . In the theory of convex risk measures, the notion of acceptability plays an important role. Specifically, for a convex risk measure ρ , the set of all acceptable positions $\mathcal{A}_\rho$ is defined by

$$\mathcal{A}_\rho = \{X \in L^\infty \mid \rho(X) \leq 0\}. \quad (49)$$

Every acceptance set induces a rejection set $\mathcal{R}_\rho = L^\infty \setminus \mathcal{A}_\rho$. As is well known (see, for instance, Föllmer and Schied [18]),

$$\rho(X) = \inf \{m \in \mathbb{R} \mid X + m \in \mathcal{A}_\rho\}, \quad (50)$$

which means that a convex risk measure $\rho(X)$ can be interpreted as the minimal amount of capital needed to make the financial position X acceptable.

One may verify that ρ is a convex risk measure if and only if $\mathcal{A}_\rho$ defined in (49) has 0 as its smallest constant element, is closed, monotone (if $X \leq Y$ and X is acceptable, then Y is acceptable as well), and convex (if X and Y are both acceptable, then, for every $\lambda \in [0, 1]$, the “diversified” portfolio $\lambda X + (1 - \lambda)Y$ is acceptable as well). Convexity captures the notion of diversification. Coherent risk measures correspond to $\mathcal{A}_\rho$ being additionally assumed to be positively homogeneous, i.e., a cone ($X \in \mathcal{A}$ implies $\lambda X \in \mathcal{A}$ for any $\lambda > 0$). Rather than starting with a convex risk measure ρ and defining $\mathcal{A}_\rho$ through (49), one can also start with a set $\mathcal{A} \subset L^\infty$, satisfying the appropriate conditions, and define a convex risk measure ρ through (50). In particular, ρ is uniquely defined through its acceptance set.

In this section, we characterize entropy coherent and entropy convex measures of risk in terms of properties of their acceptance sets. Before turning to the general case of entropy convex measures of risk, we first consider the case of entropy coherent measures of risk. To highlight the difference with coherent risk measures we also provide an alternative characterization of coherent risk measures in terms of their acceptance sets.

For a fixed probabilistic model Q and a nonempty, monotone set $\mathcal{A}^Q$ with $\mathcal{R}^Q = L^\infty \setminus \mathcal{A}^Q \neq \emptyset$, we consider the following properties:

- (RN) Risk neutrality: If $X \in \mathcal{A}^Q$, then, for any $\lambda \geq 1$, $\lambda X \in \mathcal{A}^Q$.
- (NRN) Nonrisk neutrality: If $X \in \mathcal{A}^Q$ and $Q[X < 0] > 0$, then there exists a $\lambda \geq 1$ such that $\lambda X \notin \mathcal{A}^Q$.
 - (i) Acceptance additivity: If $X, Y \in \mathcal{A}^Q$ are independent under Q , then also $X + Y \in \mathcal{A}^Q$.
 - (ii) Rejection additivity: If $X, Y \in \mathcal{R}^Q$ are independent under Q , then also $X + Y \in \mathcal{R}^Q$.
 - (iii) Mixing: If $X, Y \in \mathcal{A}^Q$, then, for any $\lambda \in [0, 1]$, the random variable Y with cdf $\lambda F_X^Q + (1 - \lambda)F_Y^Q$ under Q is in $\mathcal{A}^Q$ as well. Furthermore, if $X, Y \in \mathcal{R}^Q$, then, for any $\lambda \in [0, 1]$, the random variable Y with cdf $\lambda F_X^Q + (1 - \lambda)F_Y^Q$ under Q is in $\mathcal{R}^Q$ as well.
 - (iv) Continuity: $\mathcal{A}^Q$ is closed with respect to convergence in distribution under Q if the limiting distribution has bounded support. Furthermore, there exists $x \in \mathcal{A}^Q$ such that for any $y \in \mathcal{R}^Q$, $(1 - \alpha)\delta_x + \alpha\delta_y \in \mathcal{A}^Q$ for α close enough to zero.
 - (v) φ -convexity: There exist constants $x_1 \in \text{int}(\mathcal{A}^Q)$ and $x_2 \in \mathcal{R}^Q$ such that the function

$$\bar{\varphi}(x) := \begin{cases} \frac{z - \bar{\alpha}(x)}{1 - \bar{\alpha}(x)}, & \text{if } x \geq r, \\ \frac{z}{\bar{\alpha}(x)}, & \text{if } x < r, \end{cases} \quad (51)$$

is convex, where $r = \inf\{m \mid m \in \mathcal{A}^Q\}$, $z := \sup\{0 \leq \alpha \leq 1 \mid \alpha\delta_{x_2} + (1 - \alpha)\delta_{x_1} \in \mathcal{A}^Q\}$, $\bar{\alpha}(x) = \sup\{0 \leq \alpha \leq 1 \mid \alpha\delta_{x_2} + (1 - \alpha)\delta_x \in \mathcal{A}^Q\}$ for $x \geq 0$, and $\bar{\alpha}(x) = \sup\{0 \leq \alpha \leq 1 \mid \alpha\delta_x + (1 - \alpha)\delta_{x_1} \in \mathcal{A}^Q\}$ for $x < 0$. Here, δ_x is a Dirac point mass in x . We define $\varphi(x) = \bar{\varphi}(-x)$.

Note that in (iv)–(v) random variables are identified with their cdf's under Q . (By (iii), if $F_X^Q = F_Y^Q$, then $X \in \mathcal{A}^Q$ if and only if $Y \in \mathcal{A}^Q$, and similarly for $\mathcal{R}^Q$.)

Now we can consider a whole family of sets $(\mathcal{A}^Q)_Q$ satisfying properties (i)–(v), where the upper index Q should express that these properties are satisfied for every given probabilistic model Q , i.e., model-wise. For a family of sets $(\mathcal{A}^Q)_Q$, we introduce an additional property:

- (vi) Acceptance neutrality: If X with cdf F_X^Q under Q and Y with cdf $F_Y^{Q'}$ under Q' satisfy $F_X^Q = F_Y^{Q'}$, then $X \in \mathcal{A}^Q$ if and only if $Y \in \mathcal{A}^{Q'}$.

The following proposition provides an alternative characterization of coherent risk measures in terms of their acceptance sets. It features properties (RN), (i)–(vi).

THEOREM 6.1. *A mapping $\rho: L^\infty \rightarrow \mathbb{R}$ defined by (50) is a coherent risk measure if and only if there exists a closed and convex set $M \subset \mathcal{Q}$ such that the acceptance set $\mathcal{A}_\rho$ can be written as*

$$\mathcal{A}_\rho = \bigcap_{Q \in M} \mathcal{A}^Q, \quad (52)$$

where $(\mathcal{A}^Q)_{Q \in M}$ satisfy (RN), (i)–(vi).

PROOF. Recall that $\rho(X)$ is coherent if and only if there exists a set $M \subset \mathcal{Q}$ such that $\rho(X) = \sup_{Q \in M} E_Q[-X]$. Moreover, M can be chosen as a closed and convex subset of $\mathcal{Q}$. To prove “ $\Rightarrow$,” define $\mathcal{A}^Q := \{X \in L^\infty \mid E_Q[-X] \leq 0\}$. Then it is straightforward to see that the $(\mathcal{A}^Q)_{Q \in M}$ satisfy (RN), (i)–(vi) (with $x_1 = 1$, $x_2 = -1$, and $\bar{\varphi}(x) = 1/2(1 - x)$ in (v)) and that $\mathcal{A}_\rho = \bigcap_{Q \in M} \mathcal{A}^Q$.

Next, let us show that “ $\Leftarrow$ ” holds. Suppose that

$$\mathcal{A}_\rho = \bigcap_{Q \in M} \mathcal{A}^Q,$$

where $(\mathcal{A}^Q)_{Q \in M}$ satisfy (RN), (i)–(vi). From Weber [44], Theorem 3.3, it follows that conditions (iii)–(iv) imply that there exists an increasing function $\varphi(x) = \bar{\varphi}(-x)$, with $\bar{\varphi}$ as defined in (51), and a constant w such that $\mathcal{A}^Q = \{X \in L^\infty \mid \varphi^{-1}(E_Q[\varphi(-X)]) \leq w\}$. By condition (v), φ is convex. Define

$$\rho^Q(X) = \inf\{m \in \mathbb{R} \mid X + m \in \mathcal{A}^Q\}. \quad (53)$$

Note that if we could show that ρ^Q is additive for independent random variables, then we must have $\rho^Q(0) = 0$. Furthermore, it would then follow from Gerber [23] that φ is linear or exponential. But condition (RN) entails that φ cannot be exponential, so φ would have to be linear.

Let X and Y be independent under Q . Note that, by (50),

$$\begin{aligned}\rho^Q(X+Y) &= \inf\{m \in \mathbb{R} \mid X+Y+m \in \mathcal{A}^Q\} \\ &= \inf\{m = m_X + m_Y \in \mathbb{R} \mid X+m_X + Y+m_Y \in \mathcal{A}^Q\} \\ &\leq \inf\{m_X \in \mathbb{R} \mid X+m_X \in \mathcal{A}^Q\} + \inf\{m_Y \in \mathbb{R} \mid Y+m_Y \in \mathcal{A}^Q\} \\ &= \rho^Q(X) + \rho^Q(Y),\end{aligned}\tag{54}$$

where the inequality holds because, by property (ii), if $X+m_X, Y+m_Y \in \mathcal{A}^Q$, then (since $X+m_X$ and $Y+m_Y$ are independent under Q) also $X+m_X + Y+m_Y \in \mathcal{A}^Q$. On the other hand, using translation invariance of ρ^Q , it is straightforward to prove that

$$\rho^Q(X) = \sup\{m \in \mathbb{R} \mid X+m \in \mathcal{R}^Q\}.$$

Therefore,

$$\begin{aligned}\rho^Q(X+Y) &= \sup\{m \in \mathbb{R} \mid X+Y+m \in \mathcal{R}^Q\} \\ &= \sup\{m = m_X + m_Y \in \mathbb{R} \mid X+m_X + Y+m_Y \in \mathcal{R}^Q\} \\ &\geq \sup\{m_X \in \mathbb{R} \mid X+m_X \in \mathcal{R}^Q\} + \sup\{m_Y \in \mathbb{R} \mid Y+m_Y \in \mathcal{R}^Q\} \\ &= \rho^Q(X) + \rho^Q(Y).\end{aligned}\tag{55}$$

It follows from (54)–(55) that ρ^Q is additive for independent random variables. Therefore, φ is linear, and hence ρ is a coherent risk measure. $\square$

Note that the probabilistic models $Q \in M$ may be viewed as stress test measures. By Theorem 6.1, with $\mathcal{A}_\rho = \bigcap_{Q \in M} \mathcal{A}^Q$, a financial position X is acceptable if and only if it is acceptable under every stress test measure $Q \in M$. In this context, $\mathcal{A}^Q$ may be referred to as a stress test set. The following theorem provides a characterization of entropy coherent measures of risk in terms of their stress test sets. It features properties (NRN), (i)–(vi). It shows that entropy coherent measures of risk can be obtained just as coherent risk measures in Theorem 6.1, but by moving from risk neutrality (RN) of the stress test sets to assuming nonrisk neutrality (NRN).

THEOREM 6.2. *A mapping $\rho: L^\infty \rightarrow \mathbb{R}$ defined by (50) is an entropy coherent measure of risk if and only if there exists a closed and convex set $M \subset \mathcal{Q}$ such that the acceptance set $\mathcal{A}_\rho$ can be written as (52) where $(\mathcal{A}^Q)_{Q \in M}$ satisfy (NRN), (i)–(vi).*

PROOF. Entropy coherence means that there exists a set $M \subset \mathcal{Q}$ such that $\rho(X) = \sup_{Q \in M} e_{\gamma, Q}(X)$. Again, M can be chosen as a closed and convex subset of $\mathcal{Q}$. One easily verifies that, if X and Y are independent under Q ,

$$e_{\gamma, Q}(X+Y) = e_{\gamma, Q}(X) + e_{\gamma, Q}(Y).\tag{56}$$

To prove “ $\Rightarrow$,” define $\mathcal{A}^Q := \{X \in L^\infty \mid e_{\gamma, Q}(X) \leq 0\}$. Then it is straightforward to see that the $(\mathcal{A}^Q)_{Q \in M}$ satisfy (NRN), (i)–(vi) (with $\bar{\varphi}(x) = ae^{-x/\gamma} + b$, $\gamma \in \mathbb{R}^+$, and a and b chosen such that $\bar{\varphi}(1) = 0$ and $\bar{\varphi}(-1) = 1$ in (v)). In particular, the fact that (i) and (ii) hold for every $\mathcal{A}^Q$ follows from (56). Furthermore, by definition of ρ , clearly, $\mathcal{A}_\rho = \bigcap_{Q \in M} \mathcal{A}^Q$.

To see the direction “ $\Leftarrow$,” note that one can show exactly as in the proof of Theorem 6.1 that there exists a convex function $\varphi(x) = \bar{\varphi}(-x)$, with $\bar{\varphi}$ given by (51), and a constant w such that $\mathcal{A}^Q = \{X \in L^\infty \mid \varphi^{-1}(E_Q[\varphi(-X)]) \leq w\}$, and that ρ^Q defined in (53) is additive for independent random variables. But then Gerber [23] yields that φ is linear or exponential, and (NRN) now implies that φ has to be exponential. Furthermore, $0 \in \mathcal{A}^Q$ entails that $w = 0$. It follows that there exists $\gamma \in \mathbb{R}^+$ such that $\rho^Q(X) = e_{\gamma, Q}(X)$. As $\mathcal{A}_\rho = \bigcap_{Q \in M} \mathcal{A}^Q$, we have that $X \in \mathcal{A}_\rho$ if and only if $\sup_{Q \in M} e_{\gamma, Q}(X) \leq 0$. Now it follows from (50) and the translation invariance of $e_{\gamma, Q}$ that indeed $\rho(X) = \sup_{Q \in M} e_{\gamma, Q}(X)$. This proves that ρ is an entropy coherent measure of risk. $\square$

Theorem 6.2 shows that moving from coherent to entropy coherent measures of risk exactly and solely means moving from scale risk-neutral to nonrisk-neutral stress test sets, where scale risk neutrality is expressed by scale invariance of the stress test sets. In fact, it follows from Gerber [23] that, under acceptance and rejection additivity, scale risk neutrality implies full risk neutrality. It makes explicit that entropy coherent measures of risk may be viewed as the nonrisk-neutral counterparts of the risk-neutral coherent risk measures.

Now let us turn to the more general entropy convex case. Let $\theta: \mathcal{Q} \rightarrow \mathbb{R} \cup \{\infty\}$ be a penalty function. For a fixed probabilistic model Q with $\theta(Q) < \infty$, and a monotone set $\mathcal{A}^Q$ with $\mathcal{A}^Q \neq \emptyset$, $\mathcal{R}^Q \neq \emptyset$, we consider the following properties:

(NRN') Nonrisk neutrality with penalty: If $X \in \mathcal{A}^Q$ and $Q[X < 0] > 0$, then there exists a $\lambda \geq 1$ such that $\lambda X + (\lambda - 1)\theta(Q) \notin \mathcal{A}^Q$.

(i') Acceptance additivity with penalty: If $X, Y \in \mathcal{A}^Q$ are independent under Q , then also $X + Y + \theta(Q) \in \mathcal{A}^Q$.

(ii') Rejection additivity with penalty: If $X, Y \in \mathcal{R}^Q$ are independent under Q , then also $X + Y + \theta(Q) \in \mathcal{R}^Q$.

Notice that since $\theta(Q) \geq 0$, $\theta(Q)$ in properties (NRN'), (i'), and (ii') can be interpreted as an additional regulator's charge for increasing the scale of a financial position. We let axioms (iii)–(vi) remain unchanged.

It may then be proved similarly as in Theorem 6.1 that convex risk measures are equivalent to acceptance sets of the form (52) with $(\mathcal{A}^Q)_{Q \in \mathcal{Q}}$ satisfying properties (i'), (ii'), (iii)–(vi). Furthermore, entropy convex measures of risk are equivalent to $(\mathcal{A}^Q)_{Q \in \mathcal{Q}}$ satisfying properties (NRN'), (i'), (ii'), (iii)–(vi). In particular, this may be seen by noting that $\tilde{\rho}^Q(X) := \rho^Q(X) + \theta(Q) = \inf\{m \in \mathbb{R} \mid X + m \in \mathcal{A}^Q\}$, where ρ^Q is as defined in (53), and $\tilde{\mathcal{A}}^Q := \mathcal{A}^Q + \theta(Q)$ satisfies (NRN), (i), (ii) whenever $\mathcal{A}$ satisfies (NRN'), (i'), (ii'). To characterize convex risk measures, an analog of axiom (RN) is not needed since even if φ is exponential, ρ is still a convex risk measure. This means that entropy convex measures of risk arise from convex risk measures exactly and solely by additionally requiring the property of nonrisk neutrality (NRN') of the stress test sets. It makes explicit that entropy convex measures of risk may be viewed as those convex risk measures that are nonrisk neutral.

7. Conclusions. In this paper, we have introduced two subclasses of convex risk measures: entropy coherent and the more general entropy convex measures of risk. A variety of representation and duality results as well as some examples have made explicit that entropy coherent and entropy convex measures of risk are distinct and important classes of risk measures, and satisfy many appealing properties. These results include the facts that entropy convex measures of risk are (i) the only convex risk measures among φ -convex measures of risk with nonlinear φ , thus allowing for risk aversion; (ii) the only convex risk measures among negative certainty equivalents under variational and homothetic preferences with nonlinear utility; (iii) the only φ -convex measures of risk with nonlinear φ and nontrivial penalty function among negative certainty equivalents under variational and homothetic preferences. Furthermore, we have shown that the acceptance sets generating entropy convex measures of risk satisfy the same characteristic properties as the acceptance sets generating convex risk measures, the only difference being that the former additionally satisfy a nonrisk neutrality property that the latter need not satisfy. Entropy coherent and entropy convex measures of risk are moreover the natural generalizations of the popular entropic measure of risk. The theory developed in this paper is of a static nature. In future research we intend to develop its dynamic counterpart.

Acknowledgments. The authors are very grateful to a referee for helpful comments and suggestions that significantly improved the paper. They are also grateful to Patrick Cheridito, Hans Föllmer, Dilip Madan, Alexander Schied, Hans Schumacher, and seminar and conference participants at the Fields Institute in Toronto, the AMaMeF Conference in Berlin, the EURANDOM Lecture Day on Advances in Financial Mathematics, and the Colloquium on Risk Management and Risk Measures at Leibniz Universität Hannover for their comments and suggestions. This research was funded in part by The Netherlands Organization for Scientific Research (Laeven) [Grants NWO VENI 2006, NWO VIDI 2009].

Appendix.

LEMMA 7.1. Let $U \subset \mathbb{R}$ and $M \subset \mathcal{Q}$. Let $(f_Q(\cdot))_{Q \in M}$ be a family of functions mapping from U to $\mathbb{R}$. Suppose that for every $Q \in M$, $f_Q(\cdot)$ is a differentiable function. Let

$$f(m) := \sup_{Q \in M} f_Q(m), \quad (57)$$

and fix $m_0 \in U$. Suppose that for $f(m_0)$ the supremum is attained in a $Q_0 \in M$. Assume further that f is differentiable in m_0 . Then we must have that $f'(m_0) = f'_{Q_0}(m_0)$.

PROOF. Since f is differentiable in m_0 we must have

$$\begin{aligned} f'(m_0) &= \lim_{\epsilon \rightarrow 0} \frac{f(m_0 + \epsilon) - f(m_0)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon)}{\epsilon} + \lim_{\epsilon \rightarrow 0} \frac{f_{Q_0}(m_0 + \epsilon) - f_{Q_0}(m_0)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \frac{f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon)}{\epsilon} + f'_{Q_0}(m_0). \end{aligned} \quad (58)$$

In particular, $\lim_{\epsilon \rightarrow 0} ((f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon))/\epsilon)$ must exist. However, as for every ϵ we have that $f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon) \geq 0$, it follows that

$$\begin{aligned} \lim_{\epsilon \downarrow 0} \frac{f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon)}{\epsilon} &\geq 0, \\ \lim_{\epsilon \uparrow 0} \frac{f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon)}{\epsilon} &\leq 0. \end{aligned}$$

Therefore, $\lim_{\epsilon \rightarrow 0} ((f(m_0 + \epsilon) - f_{Q_0}(m_0 + \epsilon))/\epsilon)$ must be equal to zero. Thus, it follows from (58) that indeed $f'(m_0) = f'_{Q_0}(m_0)$. $\square$

LEMMA 7.2. Let $-\infty \leq a < b \leq \infty$. Suppose that $\varphi \in C^3((a, b))$ and that there does not exist p, γ, q such that, for all $x \in (a, b)$, $\varphi(x) = \gamma \exp\{x/\gamma\} + q$ or $\varphi(x) = px + q$. Then the function $\varphi' \circ \varphi^{-1}$ is not linear on $(\varphi(a), \varphi(b))$.

PROOF. Suppose that there exists c, d such that $\varphi' \circ \varphi^{-1}(x) = cx + d$ for all $x \in (\varphi(a), \varphi(b))$. As $\varphi' \circ \varphi^{-1} = 1/(\varphi^{-1})'$, we get that

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

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

References

- [1] Anscombe FJ, Aumann RJ (1963) A definition of subjective probability. *Ann. Math. Statist.* 34(1):199–205.
- [2] Artzner Ph, Delbaen F, Eber J-M, Heath D (1999) Coherent measures of risk. *Math. Finance* 9(3):203–228.
- [3] Barrieu P, El Karoui N (2005) Inf-convolution of risk measures and optimal risk transfer. *Finance Stochastics* 9(2):269–298.
- [4] Ben-Tal A, Teboulle M (1986) Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Sci.* 32(11):1445–1466.
- [5] Ben-Tal A, Teboulle M (1987) Penalty functions and duality in stochastic programming via φ -divergence functionals. *Math. Oper. Res.* 12(2):224–240.
- [6] Ben-Tal A, Teboulle M (2007) An old-new concept of convex risk measures: The optimized certainty equivalent. *Math. Finance* 17(3):449–476.
- [7] Borch K (1962) Equilibrium in a reinsurance market. *Econometrica* 30(3):424–444.
- [8] Cerreia-Vioglio S, Maccheroni F, Marinacci M, Montrucchio L (2011) Uncertainty averse preferences. *J. Econom. Theory* 146(4):1275–1330.
- [9] Cerreia-Vioglio S, Maccheroni F, Marinacci M, Montrucchio L (2011) Risk measures: Rationality and diversification. *Math. Finance* 21(4):743–774.
- [10] Chateauneuf A, Faro JH (2010) Ambiguity through confidence functions. *J. Math. Econom.* 45(9–10):535–558.
- [11] Cheridito P, Kupper M (2009) Recursiveness of indifference prices and translation-invariant preferences. *Math. Financial Econom.* 2(3):173–188.
- [12] Csiszár I (1975) I -divergence geometry of probability distributions and minimization problems. *Ann. Probab.* 3(1):146–158.
- [13] Dana R-A (2005) A representation result for concave Schur concave functions. *Math. Finance* 15(4):613–634.
- [14] Deprez O, Gerber HU (1985) On convex principles of premium calculation. *Insurance: Math. Econom.* 4(3):179–189.
- [15] Dunford N, Schwartz JT (1958) *Linear Operators. Part I: General Theory* (Wiley Interscience, New York).
- [16] El Karoui N, Ravanelli C (2009) Cash subadditive risk measures and interest rate ambiguity. *Math. Finance* 19(4):561–590.
- [17] Föllmer H, Schied A (2002) Convex measures of risk and trading constraints. *Finance Stochastics* 6(4):429–447.
- [18] Föllmer H, Schied A (2004) *Stochastic Finance*, 2nd ed. (De Gruyter, Berlin).
- [19] Föllmer H, Schied A, Weber S (2009) Robust preferences and robust portfolio choice. Ciarlet P, Bensoussan A, Zhang Q, eds. *Mathematical Modelling and Numerical Methods in Finance. Handbook of Numerical Analysis* 15:29–88.
- [20] Frittelli M, Gianin ER (2002) Putting order in risk measures. *J. Banking and Finance* 26(7):1473–1486.

- [21] Frittelli M, Scandolo G (2006) Risk measures and capital requirements for processes. *Math. Finance* 16(4):589–612.
- [22] Gerber HU (1979) *An Introduction to Mathematical Risk Theory*, S.S. Huebner Foundation Monograph 8 (Irwin, Homewood, IL).
- [23] Gerber HU (1985) On additive principles of zero utility. *Insurance: Math. Econom.* 4(4):249–251.
- [24] Gilboa I, Schmeidler D (1989) Maxmin expected utility with nonunique prior. *J. Math. Econom.* 18(2):141–153.
- [25] Gollier C (2001) *The Economics of Risk and Time* (MIT Press, Cambridge, MA).
- [26] Goovaerts MJ, De Vylder FEC, Haezendonck J (1984) *Insurance Premiums* (North-Holland Publishing, Amsterdam).
- [27] Goovaerts MJ, Kaas R, Laeven RJA, Tang Q (2004) A comonotonic image of independence for additive risk measures. *Insurance: Math. Econom.* 35(3):581–594.
- [28] Hansen LP, Sargent TJ (2001) Robust control and model uncertainty. *Amer. Econom. Rev.* 91(2):60–66.
- [29] Hansen LP, Sargent TJ (2007) *Robustness* (Princeton University Press, Princeton, NJ).
- [30] Hardy GH, Littlewood JE, Pólya G (1952) *Inequalities*, 2nd ed. (Cambridge University Press, Cambridge).
- [31] Heath D, Ku H (2004) Pareto equilibria with coherent measures of risk. *Math. Finance* 14(2):163–172.
- [32] Huber PJ (1981) *Robust Statistics* (John Wiley & Sons, New York).
- [33] Maccheroni F, Marinacci M, Rustichini A (2006) Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica* 74(6):1447–1498.
- [34] Mania M, Schweizer M (2005) Dynamic exponential utility indifference valuation. *Ann. Appl. Probab.* 15(3):2113–2143.
- [35] Moreau JJ (1970) Inf-convolution, sous-additivité, convexité des fonctions numériques. *Journal de Mathématiques Pures et Appliquées* 49(9):109–154.
- [36] Rockafellar RT, Uryasev SP (2000) Optimization of conditional value-at-risk. *J. Risk* 2(3):21–42.
- [37] Rockafellar RT, Uryasev SP, Zabarankin M (2008) Risk tuning with generalized linear regression. *Math. Oper. Res.* 33(3):712–729.
- [38] Ruszczyński A, Shapiro A (2006) Optimization of convex risk functions. *Math. Oper. Res.* 31(3):433–452.
- [39] Ruszczyński A, Shapiro A (2006) Conditional risk mappings. *Math. Oper. Res.* 31(3):544–561.
- [40] Savage LJ (1954) *The Foundations of Statistics*, 2nd ed. 1972, Dover, New York, (John Wiley & Sons, New York).
- [41] Strzalecki T (2011) Axiomatic foundations of multiplier preferences. *Econometrica* 79(1):47–73.
- [42] Von Neumann J, Morgenstern O (1944) *Theory of Games and Economic Behavior*, 3rd ed. 1953, (Princeton University Press, Princeton, NJ).
- [43] Wald A (1950) *Statistical Decision Functions* (John Wiley & Sons, New York).
- [44] Weber S (2006) Distribution-invariant risk measures, information, and dynamic consistency. *Math. Finance* 16(2):419–441.
- [45] Zalinescu C (2002) *Convex Analysis in General Vector Spaces* (World Scientific, Singapore).