

Elicitability of Return Risk Measures*

Mücahit Aygün

Dept. of Quantitative Economics
University of Amsterdam
and Tinbergen Institute
M.Aygun@uva.nl

Fabio Bellini

Dept. of Statistics and Quantitative Methods
University of Milano Bicocca
Fabio.Bellini@unimib.it

Roger J. A. Laeven

Dept. of Quantitative Economics
University of Amsterdam, CentER
and EURANDOM
R.J.A.Laeven@uva.nl

March 17, 2023

Abstract

Informally, a risk measure is said to be elicitable if there exists a suitable scoring function such that minimizing its expected value recovers the risk measure. In this paper, we analyze the elicibility properties of the class of return risk measures (i.e., normalized, monotone and positively homogeneous risk measures). First, we provide dual representation results for convex and geometrically convex return risk measures. Next, we establish new axiomatic characterizations of Orlicz premia (i.e., Luxemburg norms). More specifically, we prove, under different sets of conditions, that Orlicz premia naturally arise as the only elicitable return risk measures. Finally, we provide a general family of strictly consistent scoring functions for Orlicz premia, a myriad of specific examples and a mixture representation suitable for constructing Murphy diagrams.

Keywords Return risk measures, elicibility, Orlicz premia, consistent scoring functions, geometric convexity.

*We are very grateful to seminar participants at the University of Vienna and the University of Amsterdam for their comments and suggestions. This research was funded in part by the Netherlands Organization for Scientific Research under an NWO-Vici grant 2020–2025 (Aygün and Laeven).

1 Introduction

Since the seminal work of Savage ([41]) and Osband ([37]), an expanding and increasingly sophisticated literature has studied elicibility properties of risk measures. Classes of risk measures may, or may not, admit families of strictly consistent scoring functions, and hence be elicitable, with important implications for evaluating model performance and competing forecasts (see e.g., [26], [38], [2], [45], [36], [12], and the references therein). For example, Average Value-at-Risk per se is not elicitable, but it is jointly elicitable with Value-at-Risk since it admits bivariate strictly consistent scoring functions ([26], [21]).

Recently, [4] introduced the class of return risk measures, consisting of normalized, monotone and positively homogeneous risk measures. Return risk measures provide relative (or geometric) assessments of risk. They evaluate how much additional riskless log-return makes a financial position acceptable—whence their name. They constitute the relative counterparts of the class of monetary risk measures ([23], [16]), reminiscent of how relative risk aversion relates to absolute risk aversion. Their dynamic extensions, dynamic return risk measures, have been studied in [5].¹

Whereas elicibility properties of monetary risk measures are by now quite well understood, little is known about the elicibility properties of return risk measures. This paper aims to fill this gap by analyzing the elicibility properties of return risk measures, with a particular emphasis on Orlicz premia, also known as Luxemburg norms, which as we will see play a central role in the theory of elicitable return risk measures. Orlicz premia, and the links between risk measures and Orlicz space theory, have been extensively studied in the financial and actuarial mathematics literature (see e.g., [28], [8], [10], [11], [16], [32], [4], [5] and the references therein); however, their connection to statistical decision theory in general, and elicibility in particular, has not been uncovered to our best knowledge.

This paper makes three main contributions. We start by providing dual representation results for convex and geometrically convex return risk measures and clarify their precise relationship. In full generality, the dual representation takes the form of a supremum of discounted logarithmic certainty equivalents, where the discount factor can be interpreted as an index of model plausibility under ambiguity. We show that convex return risk measures occur as a special case in the richer class of geometrically convex return risk measures, and we also analyze their law-invariant representations. Furthermore, we introduce and analyze the class of optimized return risk measures and derive their dual representation.

Second, we establish new characterization results for Orlicz premia. We prove that Orlicz premia naturally arise as the only return risk measures that are elicitable. It has been shown in [37] that an elicitable risk measure must satisfy the convex level sets (CxLS) property. We establish that a law-invariant geometrically convex return risk measure with the CxLS property is necessarily an Orlicz

¹Return risk measures that allow for probability distortion were recently analyzed in [42], whereas applications of return risk measures to capital allocation can be found in [33] and [9].

premium. We also show that requiring identifiability for return risk measures singles out the class of Orlicz risk measures: under weak regularity conditions, they are the only identifiable, law-invariant, monotone and positively homogeneous measures of risk. These are our central results, the preparations and mathematical details of which are somewhat involved.

Third, we provide a general, rich family of scoring functions that we prove to be strictly consistent with Orlicz premia. A plethora of examples illustrates the generality of our new family of scoring functions. Special attention is devoted to scoring functions of the relative error form in view of their appealing properties in forecast evaluation. We also provide a mixture representation of the general family of scoring functions in terms of elementary scoring functions, depending on a low-dimensional parameter. This enables the use of so-called Murphy diagrams to compare competing forecasts simultaneously with respect to a full class of strictly consistent scoring functions, and we illustrate this in two examples.

Statistical decision theory demonstrates that some classes of functionals do not allow for meaningful point forecast evaluation by means of expected scores. Functionals that admit a strictly consistent scoring function, guaranteeing that accurate forecasts are rewarded more than inaccurate forecasts, are referred to as elicitable (see Definition 38 for a formal definition). Our characterization results reveal the important place of the class of Orlicz premia in the extensive literature on risk measures. This is graphically illustrated in Figure 1. We know from [43], [2] and [17] that convex shortfall risk measures occur as the subclass of monetary risk measures that are elicitable. Furthermore, the only elicitable law-invariant coherent risk measures are given by expectiles ([45], [17]). We establish in Theorems 32 and 36 that Orlicz premia naturally arise as elicitable return risk measures. Furthermore, in Theorem 26, we provide a direct proof of the result that the only convex Orlicz premia that are translation invariant (and, hence, coherent risk measures) are the expectiles.

The rest of this paper is organized as follows. In Section 2, we recall the general properties of return risk measures and derive some useful continuity properties. In Section 3, we provide dual representation results for geometrically convex and convex return risk measures, explicate their connection and analyze optimized return risk measures. In Section 4, we establish our characterization results for Orlicz premia. Section 5 presents our results on families of scoring functions strictly consistent with Orlicz premia including many examples.

2 Return risk measures

Let (Ω, \mathcal{F}, P) be a nonatomic probability space. In the present paper, random variables $X: \Omega \rightarrow \mathbb{R}$ represent financial losses. We will consider finite-valued risk measures defined on $L^\infty(\Omega, \mathcal{F}, P)$ or on its subsets $L_+^\infty(\Omega, \mathcal{F}, P) := \{X \in L^\infty \mid X \geq 0 \text{ } P\text{-a.s.}\}$ and $L_{++}^\infty(\Omega, \mathcal{F}, P) := \{X \in L^\infty \mid X > 0 \text{ } P\text{-a.s.}\}$. Equalities and inequalities between random variables are meant to hold P -almost surely without further explicit mentioning.

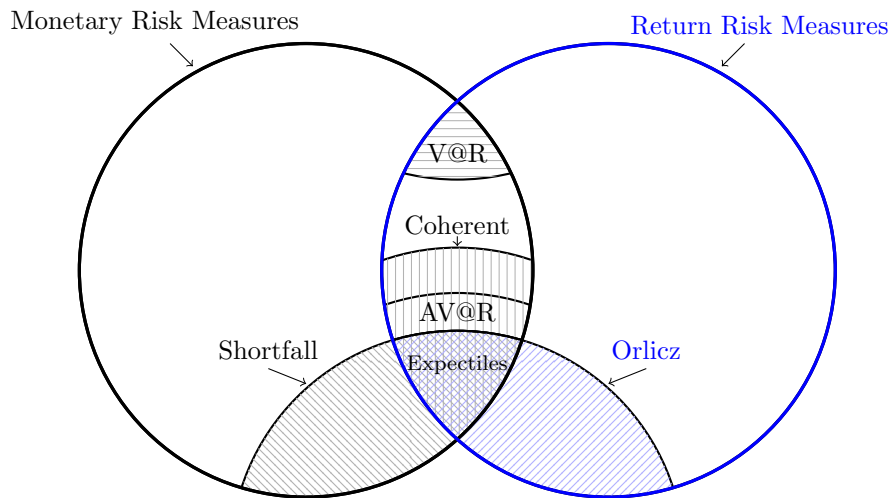


Figure 1: Venn Diagram of Classes of Risk Measures.

Notes: This figure graphically illustrates the relationships between the classes of monetary risk measures and return risk measures and some of their prominent subclasses. The intersection between monetary and return risk measures includes Value-at-Risk ($V@R$), the class of coherent risk measures in which Average Value-at-Risk ($AV@R$) occurs as a special case, and the class of expectiles. Convex shortfall risk measures occur as the subclass of monetary risk measures that are elicitable. The only coherent shortfall risk measures are the expectiles. As we establish in this paper, convex Orlicz premia arise as elicitable return risk measures. Their intersection with the class of coherent risk measures is again given by the expectiles.

Definition 1 We say that a functional $\rho: L^\infty(\Omega, \mathcal{F}, P) \rightarrow \mathbb{R}$ is:

- a) translation invariant if $\rho(X + h) = \rho(X) + h, \forall h \in \mathbb{R}, \forall X \in L^\infty$
- b) monotone if $X \leq Y \Rightarrow \rho(X) \leq \rho(Y)$
- c) monetary if ρ is translation invariant, monotone and satisfies $\rho(0) = 0$
- d) positively homogeneous if $\rho(\lambda X) = \lambda\rho(X), \forall \lambda \geq 0, \forall X \in L^\infty$
- e) convex if $\rho(\alpha X + (1 - \alpha)Y) \leq \alpha\rho(X) + (1 - \alpha)\rho(Y), \forall X, Y \in L^\infty, \forall \alpha \in (0, 1)$
- f) coherent if it is monetary, convex and positively homogeneous
- g) law invariant if $X \sim Y \Rightarrow \rho(X) = \rho(Y)$, where $X \sim Y$ means that X and Y have the same distribution.

A law-invariant functional on $L^\infty(\Omega, \mathcal{F}, P)$ induces a functional on $\mathcal{M}_{1,c}(\mathbb{R})$, the set of probability measures with compact support in \mathbb{R} , by means of

$$\rho(F) := \rho(X), \text{ if } X \sim F,$$

where each probability measure $\mu \in \mathcal{M}_{1,c}(\mathbb{R})$ is identified with its distribution function $F(x) := \mu(-\infty, x]$.

We recall from [4] the notions of return risk measure and of its associated multiplicative acceptance set.

Definition 2 A return risk measure $\tilde{\rho}: L_+^\infty \rightarrow [0, +\infty)$ is a positively homogeneous and monotone functional satisfying $\tilde{\rho}(1) = 1$. Its corresponding multiplicative acceptance set (at the level of random variables) is $B_{\tilde{\rho}} = \{X \in L_+^\infty \mid \tilde{\rho}(X) \leq 1\}$.

For return risk measures the notion of geometric convexity—also known as multiplicative convexity or GG-convexity for functions on the positive real line (see e.g., [35])—will be of interest in what follows.

Definition 3 A functional $\tilde{\rho}: L_+^\infty \rightarrow [0, +\infty)$ is geometrically convex if for each $X, Y \in L_+^\infty$ and $\alpha \in (0, 1)$ it holds that

$$\tilde{\rho}(X^\alpha Y^{1-\alpha}) \leq \tilde{\rho}^\alpha(X) \tilde{\rho}^{1-\alpha}(Y).$$

We will show in Lemma 16 that convex return risk measures are also geometrically convex. The class of geometrically convex risk measures is strictly larger than the class of convex return risk measures, a nonconvex example of the former being the logarithmic certainty equivalent $\tilde{\rho}(X) = \exp \mathbb{E}[\log X]$.

A one-to-one correspondence between return risk measures and monetary risk measures has been outlined in [4] as follows: given a monetary risk measure $\rho: L^\infty \rightarrow \mathbb{R}$, the associated return risk measure $\tilde{\rho}: L_{++}^\infty \rightarrow (0, +\infty)$ is given by

$$\tilde{\rho}(X) := \exp(\rho(\log(X))), \quad (1)$$

and, *vice versa*, given a return risk measure $\tilde{\rho}: L_{++}^{\infty} \rightarrow (0, +\infty)$, the associated monetary risk measure $\rho: L^{\infty} \rightarrow \mathbb{R}$ is

$$\rho(Y) := \log(\tilde{\rho}(\exp(Y))). \quad (2)$$

The main properties of this correspondence are recalled in the following lemma.

Lemma 4 *Let $\rho: L^{\infty} \rightarrow \mathbb{R}$ and $\tilde{\rho}: L_{++}^{\infty} \rightarrow (0, +\infty)$ be as in (1) and (2). Then:*

- a) $\rho(0) = 0 \iff \tilde{\rho}(1) = 1$
- b) ρ is translation invariant $\iff \tilde{\rho}$ is positively homogeneous
- c) ρ is monotone $\iff \tilde{\rho}$ is monotone
- d) ρ is convex $\iff \tilde{\rho}$ is geometrically convex
- e) ρ is law invariant $\iff \tilde{\rho}$ is law invariant
- f) if ρ is law invariant, then, for $F \in \mathcal{M}_{1,c}(0, +\infty)$,

$$\tilde{\rho}(F) = \exp(\rho(F(e^t))). \quad (3)$$

In Section 3, we will establish dual representations for geometrically convex return risk measures. It will turn out that for return risk measures the definitions of Fatou and Lebesgue properties have to be slightly modified. We introduce both the usual and the modified versions in the definition below.

Definition 5 *A risk measure ρ has the Fatou property if*

$$X_n \xrightarrow{P} X, \|X_n\|_{\infty} \leq k \implies \rho(X) \leq \liminf_{n \rightarrow +\infty} \rho(X_n),$$

whereas it has the Lebesgue property if

$$X_n \xrightarrow{P} X, \|X_n\|_{\infty} \leq k \implies \rho(X_n) \rightarrow \rho(X).$$

A return risk measure $\tilde{\rho}$ has the lower-bounded Fatou property if

$$X_n \xrightarrow{P} X, \|X_n\|_{\infty} \leq k, X_n \geq c > 0 \implies \tilde{\rho}(X) \leq \liminf_{n \rightarrow +\infty} \tilde{\rho}(X_n),$$

whereas it has the lower-bounded Lebesgue property if

$$X_n \xrightarrow{P} X, \|X_n\|_{\infty} \leq k, X_n \geq c > 0 \implies \tilde{\rho}(X_n) \rightarrow \tilde{\rho}(X).$$

Clearly the lower-bounded Lebesgue property implies the lower-bounded Fatou property. Both properties are weaker than the usual ones, requiring respectively lower semicontinuity and continuity under more restrictive assumptions.

Lemma 6 *Let $\rho: L^{\infty} \rightarrow \mathbb{R}$ and $\tilde{\rho}: L_{++}^{\infty} \rightarrow (0, +\infty)$ be as in (1) and (2). Then:*

- (i) $\tilde{\rho}$ has the lower-bounded Fatou property if and only if ρ has the Fatou property

(ii) $\tilde{\rho}$ has the lower-bounded Lebesgue property if and only if ρ has the Lebesgue property.

Proof. (i) Let $X_n \in L_{++}^\infty$ satisfy $X_n \xrightarrow{P} X$, $\|X_n\|_\infty \leq k$, $X_n \geq c > 0$. By the continuous mapping theorem, it follows that $\log(X_n) \xrightarrow{P} \log(X)$, and $\|\log(X_n)\|_\infty \leq \max(\log k, -\log c)$, so from the Fatou property of ρ it follows that

$$\rho(\log X) \leq \liminf_{n \rightarrow +\infty} \rho(\log(X_n)),$$

and exponentiating both sides we get $\tilde{\rho}(X) \leq \liminf_{n \rightarrow +\infty} \tilde{\rho}(\log(X_n))$. The proof of the ‘only if’ part and the proof of (ii) are similar. ■

Since a law-invariant, monetary and convex risk measure automatically satisfies the Fatou property (see [29] and [24] for recent developments on the automatic validity of the Fatou property on general spaces), it follows from Lemma 6 that a law-invariant geometrically convex return risk measure automatically has the lower-bounded Fatou property. As a consequence, we have the following mixture continuity result, in which, as usual, we denote by δ_x a probability measure supported at x .

Proposition 7 *Let $\tilde{\rho}: \mathcal{M}_{1,c}(0, +\infty) \rightarrow (0, +\infty)$ be a law-invariant geometrically convex return risk measure. Let $0 < x < y$. Then the mapping*

$$\lambda \mapsto \tilde{\rho}(\lambda\delta_x + (1 - \lambda)\delta_y)$$

is continuous at each $\lambda \in [0, 1)$.

Proof. From Lemma 4 it follows that the corresponding $\rho: L^\infty \rightarrow \mathbb{R}$ given by equation (2) is a convex law-invariant monetary risk measure. From Proposition 2.1 in [17] suitably adapted to our sign conventions it follows that

$$\lambda \mapsto \rho(\lambda\delta_u + (1 - \lambda)\delta_v)$$

is continuous at each $\lambda \in [0, 1)$, for fixed $u, v \in \mathbb{R}$ with $u < v$. Therefore, from the representation (3), we obtain

$$\tilde{\rho}(\lambda\delta_x + (1 - \lambda)\delta_y) = \exp(\rho(\lambda\delta_{\log(x)} + (1 - \lambda)\delta_{\log(y)})),$$

and the thesis follows by the continuity of compositions with exp and log. ■

3 Dual representations

In this section, we will denote by \mathbf{P} the set of probability measures on (Ω, \mathcal{F}, P) that are absolutely continuous with respect to the reference measure P . In the next theorem, we derive a dual representation of geometrically convex return risk measure as suprema of suitably weighted, or discounted, logarithmic certainty equivalents; the less plausible the probabilistic model, the lower is the corresponding discount factor.

Theorem 8 Let $\tilde{\rho}: L_{++}^{\infty} \rightarrow (0, +\infty)$ be a geometrically convex return risk measure satisfying the lower-bounded Fatou property. Then there exists a multiplicative weighting function $\tilde{\alpha}: \mathbf{P} \rightarrow [0, 1]$ satisfying $\sup_{Q \in \mathbf{P}} \tilde{\alpha}(Q) = 1$ such that

$$\tilde{\rho}(X) = \sup_{Q \in \mathbf{P}} \{ \tilde{\alpha}(Q) \exp(\mathbb{E}_Q[\log X]) \}. \quad (4)$$

Furthermore, if $\tilde{\rho}$ satisfies the lower-bounded Lebesgue property, the supremum in (4) is attained.

Proof. From Lemma 4 and Lemma 6, it follows that $\rho(X) = \log(\tilde{\rho}(\exp(X)))$ is a convex monetary risk measure with the Fatou property, so as is well-known (see e.g., [23]) it has the dual representation

$$\rho(X) = \sup_{Q \in \mathbf{P}} \{ \mathbb{E}_Q[X] - \alpha(Q) \}, \quad (5)$$

where $\alpha: \mathbf{P} \rightarrow [0, +\infty]$. Since $\tilde{\rho}(X) = \exp(\rho(\log(X)))$, it follows that

$$\begin{aligned} \tilde{\rho}(X) &= \exp \left(\sup_{Q \in \mathbf{P}} \{ \mathbb{E}_Q[\log X] - \alpha(Q) \} \right) = \sup_{Q \in \mathbf{P}} \{ \exp(\mathbb{E}_Q[\log X] - \alpha(Q)) \} \\ &= \sup_{Q \in \mathbf{P}} \{ \tilde{\alpha}(Q) \exp(\mathbb{E}_Q[\log X]) \}, \end{aligned}$$

where $\tilde{\alpha}: \mathbf{P} \rightarrow [0, 1]$ is given by $\tilde{\alpha}(Q) = \exp(-\alpha(Q))$. From $\tilde{\rho}(1) = 1$, it follows that $\sup_{Q \in \mathbf{P}} \tilde{\alpha}(Q) = 1$. By Theorem 4.22 and Exercise 4.2.2 in [23], the supremum in (5) is attained if ρ has the Lebesgue property. In view of Lemma 4 and Lemma 6, it then follows that the supremum in (4) is attained provided that $\tilde{\rho}$ satisfies the lower-bounded Lebesgue property. ■

Remark 9 The logarithmic certainty equivalent $\exp(\mathbb{E}[\log X])$ arising in the dual representation (4) can already be viewed as an example of an Orlicz premium corresponding to the unbounded Orlicz function $\Phi(x) = 1 + \log(x)$, since

$$\exp(\mathbb{E}_Q[\log X]) = \inf \left\{ k > 0 \mid \mathbb{E}_Q \left[1 + \log \left(\frac{X}{k} \right) \right] \leq 1 \right\} = H_{1+\log x, Q}(X).$$

We refer to Definition 23 in Section 4 for details on terminology and notation. As a consequence, every geometrically convex return risk measure satisfying the lower-bounded Fatou property can be seen as the supremum of a suitable family of multiplicatively weighted Orlicz premia.

We now derive a Kusuoka representation for law-invariant geometrically convex return risk measures that parallels the usual one for law-invariant convex risk measures. Recall first the definition of Average Value-at-Risk.

Definition 10 Let $X \in L^1(\Omega, \mathcal{F}, P)$. For $\lambda \in [0, 1)$, the Average Value-at-Risk of X at level λ is given by

$$AV@R_{\lambda}(X) = \frac{1}{1-\lambda} \int_{\lambda}^1 q_{\alpha}(X) d\alpha,$$

where

$$q_\alpha(X) = \inf\{x \in \mathbb{R} \mid F(x) \geq \alpha\},$$

and for $\lambda = 1$ we set by definition $AV@R_1(X) = \text{ess sup}(X)$.

Denote by $\mathcal{M}_1([0, 1])$ the set of probability measures with support in $[0, 1]$.

Theorem 11 *Let $\tilde{\rho}: L_{++}^\infty \rightarrow (0, +\infty)$ be a law-invariant geometrically convex return risk measure. Then there exists $\tilde{\beta}: \mathcal{M}_1([0, 1]) \rightarrow [0, 1]$ such that*

$$\tilde{\rho}(X) = \sup_{\mu \in \mathcal{M}_1([0, 1])} \left\{ \tilde{\beta}(\mu) \exp \left(\int_{[0, 1]} AV@R_\lambda(\log X) \mu(d\lambda) \right) \right\}. \quad (6)$$

If $\tilde{\rho}$ has the lower-bounded Lebesgue property, then $\mu(1) > 0 \Rightarrow \tilde{\beta}(\mu) = 0$.

Proof. As in the proof of the previous theorem, if $\tilde{\rho}$ is also law invariant, then from Lemma 4 the associated convex risk measure ρ given by (2) is also law invariant, and hence has the Kusuoka representation (see e.g., [23], [16])

$$\rho(X) = \sup_{\mu \in \mathcal{M}_1([0, 1])} \left(\int_{[0, 1]} AV@R_\lambda(X) \mu(d\lambda) - \beta(\mu) \right),$$

for a suitable $\beta: \mathcal{M}_1([0, 1]) \rightarrow [0, +\infty]$, from which it follows that

$$\begin{aligned} \tilde{\rho}(X) &= \exp \left(\sup_{\mu \in \mathcal{M}_1([0, 1])} \left(\int_{[0, 1]} AV@R_\lambda(\log X) \mu(d\lambda) - \beta(\mu) \right) \right) \\ &= \sup_{\mu \in \mathcal{M}_1([0, 1])} \tilde{\beta}(\mu) \exp \left(\int_{[0, 1]} AV@R_\lambda(\log X) \mu(d\lambda) \right), \end{aligned}$$

with $\tilde{\beta}(\mu) = \exp(-\beta(\mu))$. From Lemma 6, if $\tilde{\rho}$ has the lower-bounded Lebesgue property, then ρ has the Lebesgue property, and Theorem 35 in [16] implies that $\mu(1) > 0 \Rightarrow \beta(\mu) = +\infty$, from which the thesis follows. ■

Remark 12 *In the same spirit of Remark 9, the Kusuoka representation of a law-invariant geometrically convex return risk measure given in equation (6) can be written in terms of Orlicz premia as follows:*

$$\tilde{\rho}(X) = \sup_{Q \in \mathcal{M}_1(P)} \{ \tilde{\alpha}(Q) \sup_{\tilde{Q} \sim Q} H_{1+\log x, \tilde{Q}}(X) \},$$

where $\tilde{Q} \sim Q$ indicates that $\frac{dQ}{dP}$ and $\frac{d\tilde{Q}}{dP}$ have the same distribution.

As for convex risk measures, the validity of the lower-bounded Lebesgue property is linked to a suitable weak compactness property of the level sets of the weighting function $\tilde{\alpha}$ and in the law-invariant case to the so-called ψ -weak continuity. Before stating the main result we recall two basic definitions.

Definition 13 We say that a monetary convex risk measure ρ with dual representation (5) has the WC property if for each $m \in \mathbb{R}$ the lower level sets $\{Q \in \mathbf{P} \mid \alpha(Q) \leq m\}$ are compact in the $\sigma(L_1, L_\infty)$ topology. Similarly, a geometrically convex return risk measure $\tilde{\rho}$ with dual representation (4) has the \widetilde{WC} property if for each $m > 0$ the upper level sets $\{Q \in \mathbf{P} \mid \tilde{\alpha}(Q) \geq m\}$ are compact in the $\sigma(L_1, L_\infty)$ topology.

Definition 14 Let $\psi: \mathbb{R} \rightarrow [1, +\infty)$ be continuous. The ψ -weak topology on $\mathcal{M}_{1,c}(\mathbb{R})$ is the weakest topology that makes all mappings $F \mapsto \int f dF$ continuous, for each continuous f satisfying $|f| \leq c\psi$, with $c > 0$. It holds that

$$F_n \xrightarrow{\psi} F \text{ if } F_n \xrightarrow{\text{weakly}} F \text{ and } \int \psi dF_n \rightarrow \int \psi dF.$$

A functional $\rho: \mathcal{M}_{1,c} \rightarrow \mathbb{R}$ is ψ -weakly continuous if

$$F_n \xrightarrow{\psi} F \implies \rho(F_n) \rightarrow \rho(F).$$

Proposition 15 Let $\tilde{\rho}: \mathcal{M}_{1,c} \rightarrow (0, +\infty)$ be a law-invariant geometrically convex return risk measure. The following are equivalent:

- a) $\tilde{\rho}$ has the \widetilde{WC} property
- b) $\tilde{\rho}$ has the lower-bounded Lebesgue property
- c) $\tilde{\rho}$ is $\tilde{\psi}$ -weakly continuous for some $\tilde{\psi}: (0, +\infty) \rightarrow \mathbb{R}$
- d) For each $x, y > 0$ with $x < y$ and $\lambda \in [0, 1]$, the function

$$\lambda \mapsto \tilde{\rho}(\lambda\delta_x + (1-\lambda)\delta_y)$$

is continuous.

Proof. If ρ and $\tilde{\rho}$ are related by the correspondence given in (1) and (2), then the WC property of ρ is equivalent to the \widetilde{WC} property of $\tilde{\rho}$, since

$$\{Q \in \mathbf{P} \mid \tilde{\alpha}(Q) \geq m\} = \{Q \in \mathbf{P} \mid \exp(-\alpha(Q)) \geq m\} = \{Q \in \mathbf{P} \mid \alpha(Q) \leq -\log(m)\}.$$

So (a) holds if and only if the associated convex risk measure ρ has the WC property. As is well-known (see e.g., [16]), for convex risk measures the WC property is equivalent to the Lebesgue property, so from Lemma 6 it follows that (a) is equivalent to (b). From Proposition 2.7 in [17] adapted to our sign conventions, it follows that the WC property of ρ is equivalent to ψ -weak continuity with respect to some gauge function ψ . From Lemma 4 of [4], it holds that ρ is ψ -weakly continuous if and only if $\tilde{\rho}$ is $\tilde{\psi}$ -weakly continuous with $\tilde{\psi}(t) = \psi(\log(t))$, which shows the equivalence between (a) and (c). Further, Proposition 2.7 in [17] shows that the WC property of ρ is equivalent to its mixture continuity for $\lambda \rightarrow 1^-$, which combined with Proposition 7 shows that (a) is equivalent to (d). ■

As we anticipated after Definition 3, geometrically convex return risk measures are a generalization of convex risk measures, as the following shows.

Lemma 16 *Let $\tilde{\rho}: L_+^\infty \rightarrow [0, +\infty)$ be a convex return risk measure. Then $\tilde{\rho}$ is geometrically convex.*

Proof. If $\tilde{\rho}(X) = 0$ or $\tilde{\rho}(Y) = 0$ the thesis is trivial. Let $X, Y \in L_+^\infty$ and $\lambda \in (0, 1)$. By using the AM-GM inequality and the monotonicity and convexity of $\tilde{\rho}$, it follows that

$$\begin{aligned} \tilde{\rho} \left(\left(\frac{X}{\tilde{\rho}(X)} \right)^\lambda \left(\frac{Y}{\tilde{\rho}(Y)} \right)^{1-\lambda} \right) &\leq \tilde{\rho} \left(\lambda \frac{X}{\tilde{\rho}(X)} + (1-\lambda) \frac{Y}{\tilde{\rho}(Y)} \right) \\ &\leq \lambda \tilde{\rho} \left(\frac{X}{\tilde{\rho}(X)} \right) + (1-\lambda) \tilde{\rho} \left(\frac{Y}{\tilde{\rho}(Y)} \right) = 1. \end{aligned}$$

Next, from positive homogeneity, it follows that

$$\tilde{\rho}(X^\lambda Y^{1-\lambda}) \leq \tilde{\rho}(X)^\lambda \tilde{\rho}(Y)^{1-\lambda},$$

which completes the proof. ■

It is then interesting to compare the dual representation of geometrically convex return risk measures given in equation (4) with the following dual representation for convex return risk measures.

Proposition 17 *Let $\tilde{\rho}: L_+^\infty \rightarrow [0, +\infty)$ be a convex return risk measure satisfying the Fatou property. Then there exists $\beta: \mathbf{P} \rightarrow [0, 1]$ with $\sup_{Q \in \mathbf{P}} \beta(Q) = 1$ such that*

$$\tilde{\rho}(X) = \sup_{Q \in \mathbf{P}} \{ \beta(Q) \mathbb{E}_Q[X] \}. \quad (7)$$

Furthermore, if $\tilde{\rho}$ satisfies the Lebesgue property, the supremum in (7) is attained.

Proof. The first part of the statement is easily derived from Proposition 4.3 in [31]. For the second part, it follows from the proof of Proposition 4.3 in [31] that

$$\tilde{\rho}(X) = \sup_{Z \in H} \mathbb{E}[XZ],$$

where $H = \{ Z \in L_+^1 : \mathbb{E}[ZY] \leq \tilde{\rho}(Y) \text{ for any } Y \in L_+^\infty \}$. If we take $Y = 1$, then $\mathbb{E}[Z] \leq 1$ for any $Z \in H$, which gives the norm-boundedness of the set H . Furthermore, H is weakly closed, since it is an intersection of weakly closed sets. Let us take a decreasing sequence $(A_n)_n \in \mathcal{F}$ of which the intersection is the empty set. For any $Z \in H$, we have $\mathbb{E}[Z1_{A_n}] \leq \tilde{\rho}(1_{A_n})$ for every n . Therefore, by using the Lebesgue property of $\tilde{\rho}$, we have

$$\lim_{n \rightarrow +\infty} \sup_{Z \in H} \mathbb{E}[Z1_{A_n}] \leq \lim_{n \rightarrow +\infty} \tilde{\rho}(1_{A_n}) = 0,$$

which gives that H is uniformly integrable. Because H is bounded, weakly closed and uniformly integrable, it is weakly compact as a consequence of the Dunford-Pettis theorem (see, e.g., Theorem A.67 in [23]). Therefore, the supremum is

attained as a result of the Weierstrass Theorem (see, e.g., Corollary 2.35 in [1]). Suppose the supremum is attained for $\tilde{Z} \in H$. Then, the supremum is attained for \tilde{Q} such that $\frac{d\tilde{Q}}{dP} = \frac{\tilde{Z}}{\mathbb{E}[\tilde{Z}]}$. ■

Since from Lemma 16 a convex return risk measure is geometrically convex, it follows that the dual representation (4) is implied by the dual representation (7). This can be seen starting from the well-known dual representation of the exponential certainty equivalent (see e.g., [18]), given by

$$\log \mathbb{E}_Q[\exp(Y)] = \sup_{R \ll Q} \{\mathbb{E}_R[Y] - H(R, Q)\}, \quad (8)$$

where $H(R, Q)$ is the relative entropy or Kullback-Leibler divergence defined by ([13])

$$H(R, Q) := \begin{cases} \mathbb{E}_Q \left[\frac{dR}{dQ} \log \frac{dR}{dQ} \right] & \text{if } R \ll Q \\ +\infty & \text{otherwise.} \end{cases}$$

Letting $X = \exp(Y)$ and exponentiating both sides of (8), we obtain

$$\mathbb{E}_Q[X] = \sup_{R \ll Q} \{\tilde{\alpha}(R) \exp(\mathbb{E}_R[\log X])\}, \quad (9)$$

where

$$\tilde{\alpha}(R) = \exp(-H(R, Q)).$$

Now note that $\tilde{\alpha}(R) = 0$ when R is not absolutely continuous with respect to Q . Using this fact, we can rewrite expression (9) for $Q \in \mathbf{P}$, as follows:

$$\mathbb{E}_Q[X] = \sup_{R \ll Q} \{\tilde{\alpha}(R) \exp(\mathbb{E}_R[\log(X)])\} = \sup_{R \in \mathbf{P}} \{\tilde{\alpha}(R) \exp(\mathbb{E}_R[\log(X)])\}, \quad (10)$$

since we take a supremum of nonnegative numbers, $\tilde{\alpha}(R) = 0$ when $R \notin \mathbf{Q}$ and $\mathbf{Q} \subset \mathbf{P}$, where \mathbf{Q} denotes the set of probability measures absolutely continuous with respect to Q . Upon substituting the expression for $\mathbb{E}_Q[X]$ derived in (10) in (7), we have clarified the connection between (4) and (7), as follows:

$$\begin{aligned} \tilde{\rho}(X) &= \sup_{Q \in \mathbf{P}} \{\beta(Q) \mathbb{E}_Q[X]\} \\ &= \sup_{Q \in \mathbf{P}} \left\{ \beta(Q) \sup_{R \ll Q} \{\tilde{\alpha}(R) \exp(\mathbb{E}_R[\log X])\} \right\} \\ &= \sup_{Q \in \mathbf{P}} \left\{ \beta(Q) \sup_{R \in \mathbf{P}} \{\tilde{\alpha}(R) \exp(\mathbb{E}_R[\log X])\} \right\} \\ &= \sup_{Q \in \mathbf{P}} \sup_{R \in \mathbf{P}} \{\beta(Q) \{\exp(-H(R, Q)) \exp(\mathbb{E}_R[\log X])\}\} \\ &= \sup_{R \in \mathbf{P}} \sup_{Q \in \mathbf{P}} \{\beta(Q) \{\exp(-H(R, Q)) \exp(\mathbb{E}_R[\log X])\}\} \\ &= \sup_{R \in \mathbf{P}} \{c(R) \exp(\mathbb{E}_R[\log X])\}, \end{aligned}$$

where

$$c(R) = \sup_{Q \in \mathbf{P}} \beta(Q) \exp(-H(R, Q)).$$

Following the construction outlined in [6, 7] and [3], return risk measures may be optimized and become translation invariant, hence monetary, as follows:

Definition 18 *An optimized return risk measure (henceforth, OR risk measure) $\rho : L_+^\infty \rightarrow \mathbb{R}$ is defined as*

$$\rho(X) = \inf_{x \in \mathbb{R}} \left\{ x + \tilde{\rho} \left((X - x)^+ \right) \right\}, \quad (11)$$

for a corresponding return risk measure $\tilde{\rho} : L_+^\infty \rightarrow [0, +\infty)$.

Lemma 19 *An OR risk measure satisfies the following properties:*

- a) *monotonicity*
- b) *positive homogeneity*
- c) *translation invariance*
- d) *if $\tilde{\rho}$ is convex, then ρ is convex.*

Proof. Take $X, Y \in L_+^\infty$ such that $X \leq Y$. For an arbitrary $x \in \mathbb{R}$, we have $(Y - x)^+ \geq (X - x)^+$, which implies $x + \tilde{\rho}((Y - x)^+) \geq x + \tilde{\rho}((X - x)^+)$ due to the monotonicity of $\tilde{\rho}$. Since this is valid for any $x \in \mathbb{R}$, by taking the infimum on both sides, we obtain $\rho(X) \leq \rho(Y)$. For (b), by using the positive homogeneity of $\tilde{\rho}$ and of the positive part function, we have, for any $\lambda > 0$,

$$\begin{aligned} \rho(\lambda X) &= \inf_{x \in \mathbb{R}} \left\{ x + \tilde{\rho} \left((\lambda X - x)^+ \right) \right\} = \inf_{x \in \mathbb{R}} \left\{ x + \lambda \tilde{\rho} \left(\left(X - \frac{x}{\lambda} \right)^+ \right) \right\} \\ &= \inf_{\tilde{x} \in \mathbb{R}} \left\{ \lambda \tilde{x} + \lambda \tilde{\rho} \left((X - \tilde{x})^+ \right) \right\} = \lambda \inf_{\tilde{x} \in \mathbb{R}} \left\{ \tilde{x} + \tilde{\rho} \left((X - \tilde{x})^+ \right) \right\} = \lambda \rho(X). \end{aligned}$$

For (c), we have, for any $h \in \mathbb{R}$,

$$\begin{aligned} \rho(X + h) &= \inf_{x \in \mathbb{R}} \left\{ x + \tilde{\rho} \left((X + h - x)^+ \right) \right\} = \inf_{x \in \mathbb{R}} \left\{ x + \tilde{\rho} \left((X - (x - h))^+ \right) \right\} \\ &= \inf_{\tilde{x} \in \mathbb{R}} \left\{ \tilde{x} + h + \tilde{\rho} \left((X - \tilde{x})^+ \right) \right\} = h + \inf_{\tilde{x} \in \mathbb{R}} \left\{ \tilde{x} + \tilde{\rho} \left((X - \tilde{x})^+ \right) \right\} \\ &= \rho(X) + h. \end{aligned}$$

Finally, let us assume that $\tilde{\rho}$ is convex and take $X, Y \in L_+^\infty$. Because ρ is positively homogeneous, it is sufficient for (d) to prove that ρ is subadditive. We have

$$\begin{aligned} \rho(X + Y) &= \inf_{x, y \in \mathbb{R}} \left\{ x + y + \tilde{\rho} \left((X + Y - x - y)^+ \right) \right\} \\ &\leq \inf_{x, y \in \mathbb{R}} \left\{ x + y + \tilde{\rho} \left((X - x)^+ \right) + \tilde{\rho} \left((Y - y)^+ \right) \right\} \\ &= \inf_{x \in \mathbb{R}} \left\{ x + \tilde{\rho} \left((X - x)^+ \right) \right\} + \inf_{y \in \mathbb{R}} \left\{ y + \tilde{\rho} \left((Y - y)^+ \right) \right\} \\ &= \rho(X) + \rho(Y), \end{aligned}$$

where we have used the convexity and positive homogeneity of $\tilde{\rho}$ and of the positive part function in the second line. ■

The class of OR risk measures encompasses as special cases the Rockafellar-Uryasev [40] construction of Average-Value-at-Risk as well as its generalization given by the (robust) HG risk measure ([4]). We now establish that the OR risk measure admits an inf-convolution and a dual representation.

Definition 20 *The inf-convolution $(f \square g)$ of two convex functionals $f : L^\infty \rightarrow \overline{\mathbb{R}}$ and $g : L^\infty \rightarrow \overline{\mathbb{R}}$ is defined as follows:*

$$(f \square g)(X) = \inf_{Y \in L^\infty} \{f(X - Y) + g(Y)\}.$$

Lemma 21 *An OR risk measure ρ can be written as*

$$\rho(X) = (f \square g)(X),$$

where $f(X) = \tilde{\rho}(X^+)$ and

$$g(Y) = \begin{cases} x & \text{if } Y = x, \\ +\infty & \text{otherwise,} \end{cases}$$

when the corresponding return risk measure $\tilde{\rho}$ is convex.

Proof. Note that the functional f is convex since $\tilde{\rho}$ is convex and monotone and the positive part function is convex, and g is convex, too. Then, the inf-convolution of the functionals f and g agrees with the definition of ρ in (11). ■

Recall that the dual space of L^∞ can be identified with L^1 w.r.t. the $\sigma(L^\infty, L^1)$ -topology. The convex conjugate $h^* : L^1 \rightarrow \overline{\mathbb{R}}$ of a function $h : L^\infty \rightarrow \overline{\mathbb{R}}$ is defined as:

$$h^*(\varphi) = \sup_{X \in L^\infty} \{\mathbb{E}[\varphi X] - h(X)\}.$$

Proposition 22 *An OR risk measure ρ , with a corresponding convex return risk measure $\tilde{\rho}$, admits the following dual representation:*

$$\rho(X) = \sup_{Q \in A_P} \mathbb{E}_Q[X], \tag{12}$$

where $A_P = \{Q \in \mathbf{P} : \mathbb{E}_Q[Y] \leq \tilde{\rho}(Y) \text{ for any } Y \in L_+^\infty\}$. Furthermore, if $\tilde{\rho}$ satisfies the Lebesgue property, then the supremum in (12) is attained.

Proof. From, e.g., Theorem 2.3.1 in [44], it is known that

$$(f \square g)^* = f^* + g^*. \tag{13}$$

Let us consider the conjugates of the functionals f and g in Lemma 21. The conjugate f^* can be calculated as follows:

$$\begin{aligned} f^*(\varphi) &= \sup_{Y \in L^\infty} \{\mathbb{E}[\varphi Y] - f(Y)\} = \sup_{Y \in L^\infty} \{\mathbb{E}[\varphi Y] - \tilde{\rho}(Y^+)\} \\ &= \begin{cases} 0 & \text{if } \varphi \in L_+^1 \text{ and } \mathbb{E}[\varphi Y] \leq \tilde{\rho}(Y) \text{ for any } Y \in L_+^\infty, \\ +\infty & \text{otherwise,} \end{cases} \end{aligned}$$

by using the positive homogeneity of $\tilde{\rho}$. The conjugate g^* can be calculated as follows:

$$\begin{aligned} g^*(\varphi) &= \sup_{Y \in L^\infty} \{\mathbb{E}[\varphi Y] - g(Y)\} = \sup_{x \in \mathbb{R}} \{\mathbb{E}[\varphi x] - x\} = \sup_{x \in \mathbb{R}} x(\mathbb{E}[\varphi] - 1) \\ &= \begin{cases} 0 & \text{if } \mathbb{E}[\varphi] = 1, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned}$$

By using (13) and Lemma 21, we have the following:

$$\rho^*(\varphi) = \begin{cases} 0 & \text{if } \varphi \in L_+^1, \mathbb{E}[\varphi] = 1 \text{ and } \mathbb{E}[\varphi Y] \leq \tilde{\rho}(Y) \text{ for any } Y \in L_+^\infty, \\ +\infty & \text{otherwise.} \end{cases}$$

Therefore, ρ^* is the indicator function of the set

$$A = \{\varphi \in L_+^1 : \mathbb{E}[\varphi] = 1 \text{ and } \mathbb{E}[\varphi Y] \leq \tilde{\rho}(Y) \text{ for any } Y \in L_+^\infty\}.$$

Since the functional ρ is the inf-convolution of the functionals f and g , as a consequence of the Fenchel-Moreau theorem, we have

$$\begin{aligned} \rho(X) &= \sup_{\varphi \in L_1} \{\mathbb{E}[\varphi X] - \rho^*(\varphi)\} = \sup_{\varphi \in A} \mathbb{E}[\varphi X] \\ &= \sup_{Q \in A_P} \mathbb{E}_Q[X], \end{aligned}$$

where $A_P = \{Q \in \mathbf{P} : \mathbb{E}_Q[Y] \leq \tilde{\rho}(Y) \text{ for any } Y \in L_+^\infty\}$. Hence,

$$\rho(X) = \sup_{Q \in A_P} \mathbb{E}_Q[X]. \quad (14)$$

Following the same argument used in the proof of Proposition 17, it follows that if $\tilde{\rho}$ has the Lebesgue property, then the set A_P is weakly compact, from which the attainment of the maximum follows. Indeed, the set A_P is norm-bounded by definition. It is weakly closed, since it is the intersection of weakly closed sets. Now let us take a decreasing sequence $(A_n)_n \in \mathcal{F}$ of which the intersection is the empty set. For any $Q \in A_P$, we have $\mathbb{E}_Q[1_{A_n}] \leq \tilde{\rho}(1_{A_n})$ for every n . Hence, by using the Lebesgue property of $\tilde{\rho}$, we have

$$\lim_{n \rightarrow +\infty} \sup_{Q \in A_P} \mathbb{E}_Q[1_{A_n}] \leq \lim_{n \rightarrow +\infty} \tilde{\rho}(1_{A_n}) = 0,$$

which gives that A_P is uniformly integrable. Therefore, the supremum is attained due to the Dunford-Pettis and Weierstrass theorems; cf. also Theorem 3.6 in [15] and Theorem 8 in [14]. ■

4 Axiomatizations of Orlicz premia

In this section, we first define Orlicz premia and derive some properties that are relevant in this paper; and next we establish new axiomatizations of Orlicz premia and compare them with the one given in [4].

4.1 Orlicz premia: Definition and properties

The mathematical definition of the Orlicz premium corresponds to the Luxemburg norm on the Orlicz space

$$L^\Phi := \left\{ X \in L^0(\Omega, \mathcal{F}, P) : \mathbb{E} \left[\frac{|X|}{k} \right] < +\infty \text{ for some } k > 0 \right\},$$

given by

$$H_\Phi(X) := \inf\{k > 0 \mid \mathbb{E}[\Phi(X/k)] \leq 1\},$$

where the Orlicz function $\Phi: [0, +\infty) \rightarrow [0, +\infty]$ satisfies $\Phi(0) = 0$, is nondecreasing, left-continuous, convex, and nontrivial in the sense that $\Phi(x) > 0$ for some $x > 0$ and $\Phi(x) < +\infty$ for some $x > 0$. We refer to [19] for the basic properties of Luxemburg norms under these assumptions. Notice that in the actuarial and financial mathematics literature (e.g., [28], [10], [11], [4], [5]) there are small differences in the set of properties required to Φ . In this paper, we are interested in possibly nonconvex Orlicz functions that may not satisfy $\Phi(0) = 0$; conversely, we will limit the domain of H_Φ to L_+^∞ . (When the function $\Phi(\cdot)$ is convex and satisfies several additional properties, it is often referred to as a Young function; as these conditions are not assumed in this paper, we refer to $\Phi(\cdot)$ as an Orlicz function.) This leads to the following definition.

Definition 23 *Let $\Phi: [0, +\infty) \rightarrow \mathbb{R}$ satisfy:*

- a) $\Phi(1) = 1$, $\lim_{x \rightarrow +\infty} \Phi(x) = +\infty$
- b) Φ is nondecreasing
- c) Φ is left-continuous

For $X \in L_+^\infty$, the Orlicz premium is defined by

$$H_\Phi(X) = \inf\{k > 0 \mid \mathbb{E}[\Phi(X/k)] \leq 1\}.$$

We recall the relevant properties of Orlicz premia in the following proposition.

Proposition 24 *Let $\Phi: [0, +\infty) \rightarrow \mathbb{R}$ and $H_\Phi(X)$ be as in Definition 23. Then,*

- a) H_Φ is monotone, positively homogeneous and satisfies $H_\Phi(1) = 1$
- b) for each $X \in L_{++}^\infty$, it holds that $\mathbb{E}[\Phi(X/H_\Phi(X))] = 1$
- c) if Φ is increasing, then

$$\begin{aligned} \mathbb{E}[\Phi(X/k)] = 1 &\iff k = H_\Phi(X) \\ \mathbb{E}[\Phi(X/k)] > 1 &\iff k < H_\Phi(X) \end{aligned}$$

d) H_{Φ} is convex if and only if Φ is convex.

Proof. (a) The proof is standard. (b) Let $g(k) := \mathbb{E}[\Phi(X/k)]$. If $k_n \downarrow k$ then $\Phi(X/k_n) \uparrow \Phi(X/k)$, so from the monotone convergence theorem it follows that g is right-continuous. Since $H_{\Phi}(X) = \inf\{k \mid g(k) \leq 1\}$, it follows that $g(H_{\Phi}(X)) = 1$, that is, $\mathbb{E}[\Phi(X/H_{\Phi}(X))] = 1$. (c) The ‘only if’ part of the first implication follows by strict monotonicity of Φ . The ‘if’ part of the second implication is just the definition of H_{Φ} , while the ‘only if’ part follows from (c). (d) The proof of the ‘if’ part is standard. To prove the ‘only if’ part, assume first by contradiction that Φ is not midconvex, i.e., there exist $x_1, x_2 \geq 0$ such that $\Phi((x_1 + x_2)/2) > (\Phi(x_1) + \Phi(x_2))/2$. Then, there exists $z \in [0, +\infty)$ and $\lambda \in (0, 1)$ such that

$$\lambda\Phi(z) + (1 - \lambda)\Phi((x_1 + x_2)/2) > 1 > \lambda\Phi(z) + (1 - \lambda)\frac{\Phi(x_1) + \Phi(x_2)}{2}. \quad (15)$$

Let $A, B, C \in \mathcal{F}$ be disjoint sets with $P(A) = \lambda$, $P(B) = P(C) = \frac{1-\lambda}{2}$ and let

$$\begin{aligned} X &= z1_A + x_11_B + x_21_C \\ Y &= z1_A + x_21_B + x_11_C \\ Z &= z1_A + \frac{x_1 + x_2}{2}1_{B \cup C} = \frac{X + Y}{2}. \end{aligned}$$

From (15), we have $\mathbb{E}[\Phi(Z)] > 1$ and $\mathbb{E}[\Phi(X)] = \mathbb{E}[\Phi(Y)] < 1$, which contradicts the convexity of H_{Φ} . As a consequence, Φ is midconvex and since it is nondecreasing it is also convex. ■

A remarkable example in which $\Phi(0) \neq 0$ and Φ is not differentiable is the following.

Example 25 (Expectiles) Let $0 < q < 1$ and let

$$\Phi_q(x) = 1 + q(x - 1)^+ - (1 - q)(x - 1)^-.$$

Then, $\Phi_q(0) = q$, $\Phi_q(1) = 1$ and Φ_q is convex if $1/2 \leq q < 1$ and concave if $0 < q \leq 1/2$. The corresponding Orlicz premium H_{Φ_q} satisfies

$$\mathbb{E}[\Phi_q(X/H_{\Phi_q}(X)) - 1] = 0,$$

which gives

$$\mathbb{E}[q(X - H_{\Phi_q}(X))^+ - (1 - q)(X - H_{\Phi_q}(X))^-] = 0,$$

so $H_{\Phi_q}(X)$ coincides with the q -expectile of X ([34, 30]), denoted by $e_q(X)$ and defined for $X \in L^1(\Omega, \mathcal{F}, P)$ by the condition

$$q \mathbb{E}[(X - e_q(X))^+] = (1 - q) \mathbb{E}[(X - e_q(X))^-].$$

The following theorem shows that expectiles are the most general translation invariant convex Orlicz premia.

Theorem 26 *If Φ is increasing and H_Φ is translation invariant and convex, then*

$$\Phi(x) = 1 + a(x-1)^+ - b(x-1)^-,$$

with $a > b$ and $b < 1$.

Proof. Let $h(x) := \Phi(x+1) - 1$. Then $h(x) = 0 \iff x = 0$ and $h(x) > 0 \iff x > 0$, and

$$\begin{aligned} \mathbb{E}[h((Y/k) - 1)] = 0 &\iff H_\Phi(Y) = k \\ \mathbb{E}[h((Y/k) - 1)] < 0 &\iff H_\Phi(Y) < k. \end{aligned}$$

Fix $x < 0$ and $z > 0$. Let p be a solution of the equation

$$ph(x) + (1-p)h(z) = 0, \tag{16}$$

and let $Y = p\delta_{x+1} + (1-p)\delta_{z+1}$. It follows that $H_\Phi(Y) = 1$, and from translation invariance it follows that $H_\Phi(Y+c) = c+1$ for each $c \in \mathbb{R}_+$, which implies

$$0 = \mathbb{E} \left[h \left(\frac{Y+c}{c+1} - 1 \right) \right] = ph \left(\frac{x}{c+1} \right) + (1-p)h \left(\frac{z}{c+1} \right). \tag{17}$$

Let $\lambda = \frac{1}{c+1}$ and note that $0 < \lambda \leq 1$. By combining (17) with (16), we have

$$\lambda ph(x) + \lambda(1-p)h(z) = 0 = ph(\lambda x) + (1-p)h(\lambda z),$$

which gives

$$p(\lambda h(x) - h(\lambda x)) + (1-p)(\lambda h(z) - h(\lambda z)) = 0.$$

From the convexity of h it follows that $\lambda h(x) - h(\lambda x) \geq 0$ and $\lambda h(z) - h(\lambda z) \geq 0$, so from the last equality we get $h(\lambda z) = \lambda h(z)$ and $h(\lambda x) = \lambda h(x)$ for every $0 \leq \lambda \leq 1$ and for each $x < 0$ and $z > 0$, from which the thesis follows. ■

It is immediate to check that $H_\Phi(X) = e_q(X)$, with $q = a/(a+b)$. Further, the same argument also shows that expectiles with $0 < q \leq 1/2$ are the only concave translation invariant Orlicz premia. It is interesting to compare the theorem above with [28] and [27], where it is shown that a translation invariant Orlicz premium must be equal to the mean, but in their result actually also the differentiability of the Orlicz function Φ is assumed.

Definition 27 *A function $f: [0, +\infty) \rightarrow \mathbb{R}$ is called GA-convex if, for each $\lambda \in (0, 1)$ and $x, y > 0$, it holds that*

$$f(x^\lambda y^{1-\lambda}) \leq \lambda f(x) + (1-\lambda)f(y).$$

It is not difficult to verify that a nondecreasing and convex function on $(0, +\infty)$ is GA-convex. For completeness we report the proof in Lemma 49 in the Appendix. The converse does not hold, an example being $f(x) = \log x$ that is increasing and GA-convex but not convex. We refer to [35] for further properties of GA-convex functions.

Proposition 28 *Let $\Phi: [0, +\infty) \rightarrow \mathbb{R}$ and $H_\Phi(X)$ be as in Definition 23. Then H_Φ is geometrically convex if and only if Φ is GA-convex.*

Proof. We first prove the ‘if’ part. Notice first that, for each $X \in L_+^\infty$ and each $k > H_\Phi(X)$, it holds by definition that $\mathbb{E}[\Phi(X/k)] \leq 1$. Since Φ is nondecreasing and left-continuous an application of the monotone convergence theorem shows that $\mathbb{E}[\Phi(X/H_\Phi(X))] \leq 1$. Let now $X, Y \in L_+^\infty$ and $\lambda \in (0, 1)$. From the GA-convexity of Φ it follows that

$$\begin{aligned} & \mathbb{E} \left[\Phi \left(\left(\frac{X}{H_\Phi(X)} \right)^\lambda \left(\frac{Y}{H_\Phi(Y)} \right)^{1-\lambda} \right) \right] \\ & \leq \lambda \mathbb{E} \left[\Phi \left(\frac{X}{H_\Phi(X)} \right) \right] + (1 - \lambda) \mathbb{E} \left[\Phi \left(\frac{Y}{H_\Phi(Y)} \right) \right] \leq 1, \end{aligned}$$

which implies

$$H_\Phi \left(\frac{X^\lambda Y^{1-\lambda}}{H_\Phi(X)^\lambda H_\Phi(Y)^{1-\lambda}} \right) \leq 1,$$

which from positive homogeneity gives

$$H_\Phi(X^\lambda Y^{1-\lambda}) \leq H_\Phi(X)^\lambda H_\Phi(Y)^{1-\lambda}.$$

To prove the ‘only if’ part, we first assume by contradiction that Φ is not GA-midconvex, i.e., there exist $x_1, x_2 \geq 0$ such that $\Phi(\sqrt{x_1 x_2}) > (\Phi(x_1) + \Phi(x_2))/2$. Then, there exist $z \in [0, +\infty)$ and $\lambda \in (0, 1)$ such that

$$\lambda \Phi(z) + (1 - \lambda) \Phi(\sqrt{x_1 x_2}) > 1 > \lambda \Phi(z) + (1 - \lambda) \frac{\Phi(x_1) + \Phi(x_2)}{2}. \quad (18)$$

Take disjoint sets $A, B, C \in \mathcal{F}$ with $P(A) = \lambda$, $P(B) = P(C) = \frac{1-\lambda}{2}$ and let

$$\begin{aligned} X &= z1_A + x_1 1_B + x_2 1_C \\ Y &= z1_A + x_2 1_B + x_1 1_C \\ Z &= z1_A + \sqrt{x_1 x_2} 1_{B \cup C} = \sqrt{XY}. \end{aligned}$$

From (18), we have $\mathbb{E}[\Phi(Z)] > 1$ and $\mathbb{E}[\Phi(X)] = \mathbb{E}[\Phi(Y)] < 1$, which contradicts with the geometric convexity of H_Φ . As a consequence, Φ is GA-midconvex. Since Φ is also nondecreasing the thesis follows. Indeed, this is seen as follows. By the induction hypothesis, Φ is rationally GA-convex. Now let us take $x, y \geq 0$ and $\lambda \in (0, 1)$. Without loss of generality, assume that $x > y$. Take a $q \in \mathbb{Q} \cap [0, 1]$ such that $q > \lambda$. By monotonicity and rational GA-convexity of Φ , we have

$$\Phi(x^\lambda y^{1-\lambda}) \leq \Phi(x^q y^{1-q}) \leq q\Phi(x) + (1 - q)\Phi(y).$$

Since this inequality is valid for any $q \in \mathbb{Q} \cap [0, 1]$ such that $q > \lambda$, we can take the infimum over the set $\Lambda_Q := \{q \in \mathbb{Q} : 1 \geq q > \lambda\}$ and obtain

$$\Phi(x^\lambda y^{1-\lambda}) \leq \inf_{q \in \Lambda_Q} q\Phi(x) + (1 - q)\Phi(y) = \lambda\Phi(x) + (1 - \lambda)\Phi(y),$$

which gives the GA-convexity of Φ . ■

Since a nondecreasing and convex function is GA-convex, it follows that a convex Orlicz premium is also geometrically convex. The converse does not hold, an example being the logarithmic certainty equivalent, which is also the Orlicz premium corresponding to $\Phi(x) = 1 + \log(x)$.

4.2 Axiomatization based on the properties of the multiplicative acceptance set

This is Theorem 2 in [4] that we report below for convenience.

Theorem 29 *Let $\tilde{\rho}: \mathcal{M}_{1,c}(0, +\infty) \rightarrow \mathbb{R}$ be a law-invariant return risk measure and let $\mathcal{B}_{\tilde{\rho}}$ be the corresponding multiplicative acceptance set as in Definition 2 (now at the level of distributions). Assume that*

- a) $\mathcal{B}_{\tilde{\rho}}$ and $\mathcal{B}_{\tilde{\rho}}^c$ are convex with respect to mixtures, i.e., for each $\lambda \in (0, 1)$, $F, G \in \mathcal{B}_{\tilde{\rho}} \Rightarrow \lambda F + (1 - \lambda)G \in \mathcal{B}_{\tilde{\rho}}$, and similarly for $\mathcal{B}_{\tilde{\rho}}^c$
- b) $\mathcal{B}_{\tilde{\rho}}$ is $\tilde{\psi}$ -weakly closed for some gauge function $\tilde{\psi}$
- c) for each $0 < \tilde{x} < 1$ and $\tilde{y} > 1$, there exists $\alpha \in (0, 1)$ such that

$$\alpha \delta_{\tilde{x}} + (1 - \alpha) \delta_{\tilde{y}} \in \mathcal{B}_{\tilde{\rho}}.$$

Then there exists a nondecreasing function Φ that satisfies $\Phi(0) < 1 < \Phi(+\infty)$ such that $\tilde{\rho}(F) = H_{\Phi}(F)$.

As we will see, the convexity with respect to mixtures (at the level of distributions) of the multiplicative acceptance set and its complement assumed in item (a) in the theorem above, is implied by the CxLS property.

4.3 Axiomatizations based on CxLS

Definition 30 *A law-invariant functional ρ has the CxLS property if*

$$\rho(F) = \rho(G) = \gamma \Rightarrow \rho(\lambda F + (1 - \lambda)G) = \gamma,$$

for each $\gamma \in \mathbb{R}$, $F, G \in \mathcal{M}_{1,c}$ and $\lambda \in (0, 1)$.

Lemma 31 *Let $\tilde{\rho}$ be a law-invariant return risk measure with CxLS. Then $\mathcal{B}_{\tilde{\rho}}$ and $\mathcal{B}_{\tilde{\rho}}^c$ are convex with respect to mixtures.*

Proof. Let us take $F, G \in \mathcal{B}_{\tilde{\rho}}^c$ and $\lambda \in (0, 1)$. Let $X, Y \in L_{++}^{\infty}$ such that the distributions of X and Y are F and G . Take $A \in \mathcal{F}$ such that $P(A) = \lambda$ and X, Y and A are independent. Choosing such X, Y and A is possible because we are working in an atomless probability space, see Lemma 3.1 in [17]. Without loss of generality, assume that $\tilde{\rho}(X) = k\tilde{\rho}(Y)$ for some $k \geq 1$. Define $X' = X/k$ and denote its distribution by F' . By positive homogeneity, we have $\tilde{\rho}(X') = \tilde{\rho}(Y)$.

Then, $X1_A + Y1_{A^c}$ has distribution $\lambda F + (1 - \lambda)G$ and $X'1_A + Y1_{A^c}$ has distribution $\lambda F' + (1 - \lambda)G$. Since $k \geq 1$ and $X \in L_{++}^\infty$, we have

$$X1_A + Y1_{A^c} \geq X'1_A + Y1_{A^c}.$$

By using the monotonicity and the CxLS property, we have

$$\tilde{\rho}(X1_A + Y1_{A^c}) \geq \tilde{\rho}(X'1_A + Y1_{A^c}) = \tilde{\rho}(Y) > 1,$$

which gives the convexity of \mathcal{B}_ρ^c with respect to mixtures. Similarly, it can be proved that \mathcal{B}_ρ is convex with respect to mixtures. ■

Theorem 32 *Let $\tilde{\rho}: L_{++}^\infty \rightarrow (0, +\infty)$ be a law-invariant geometrically convex return risk measure with CxLS. Then there exists a nondecreasing GA-convex Orlicz function $\Phi: [0, +\infty) \rightarrow \mathbb{R} \cup \{+\infty\}$ such that $\tilde{\rho}(X) = H_\Phi(X)$.*

Proof. From the hypotheses and Lemma 4, it follows that the corresponding ρ given by (2) is a convex law-invariant monetary risk measure with CxLS. From Theorem 3.10 in [17], there exists a convex function $\varphi: \mathbb{R} \rightarrow \mathbb{R} \cup \{+\infty\}$ such that $\rho(X) \leq 0$ if and only if $\mathbb{E}[\varphi(X)] \leq 0$. Letting $\Phi(x) := 1 + \varphi(\log(x))$, it follows that

$$\begin{aligned} \tilde{\rho}(X/k) \leq 1 &\iff \rho(\log(X/k)) \leq 0 \iff \mathbb{E}[\varphi(\log(X/k))] \leq 0 \\ &\iff \mathbb{E}[\Phi(X/k)] \leq 1. \end{aligned}$$

From the convexity of φ , it follows that for each $x, y > 0$ and $\lambda \in (0, 1)$,

$$\begin{aligned} \Phi(x^\lambda y^{1-\lambda}) &= 1 + \varphi(\log(x^\lambda y^{1-\lambda})) = 1 + \varphi(\lambda \log(x) + (1 - \lambda) \log(y)) \\ &\leq 1 + \lambda \varphi(\log(x)) + (1 - \lambda) \varphi(\log(y)) = \lambda \Phi(x) + (1 - \lambda) \Phi(y), \end{aligned}$$

which shows the GA-convexity of Φ . ■

Notice the consistency between Proposition 28 and Theorem 32. Notice also that under the assumptions of Theorem 32, the function Φ is not necessarily convex as the example of logarithmic certainty equivalents shows.

Theorem 33 *Let $\tilde{\rho}: L_{++}^\infty \rightarrow (0, +\infty)$ be a law-invariant convex return risk measure with CxLS. Then there exists a nondecreasing convex Orlicz function $\Phi: [0, +\infty) \rightarrow \mathbb{R} \cup \{+\infty\}$ such that $\tilde{\rho}(X) = H_\Phi(X)$.*

Proof. Since a positively homogeneous, monotone convex functional defined on L_{++}^∞ is geometrically convex, it follows from Theorem 32 that there exists a nondecreasing GA-convex Orlicz function $\Phi: [0, +\infty) \rightarrow \mathbb{R} \cup \{+\infty\}$ such that $\tilde{\rho}(X) = H_\Phi(X)$. Since H_Φ is convex only if Φ is convex, as has been shown in Proposition 24, the thesis follows. ■

4.4 Axiomatization based on identifiability

Definition 34 We say that a functional $\rho: \mathcal{M}_{1,c}(0, +\infty) \rightarrow (0, +\infty)$ is identifiable if there exists at least one identification function $I: (0, +\infty) \times (0, +\infty) \rightarrow \mathbb{R}$ that satisfies, for each $F \in \mathcal{M}_{1,c}(0, +\infty)$,

$$\begin{aligned} \int I(x, y) dF(y) = 0 &\iff x = \rho(F), \\ \int I(x, y) dF(y) > 0 &\iff x < \rho(F). \end{aligned}$$

Example 35 Let $\rho(F) = \mathbb{E}[F]$. Then ρ is identifiable and two possible identification functions are $I_1(x, y) = y - x$ and $I_2(x, y) = y/x - 1$.

An identifiable functional has the CxLS property, since

$$\begin{aligned} \rho(F) = \rho(G) = \gamma &\Rightarrow \int I(\gamma, y) dF(y) = 0, \int I(\gamma, y) dG(y) = 0 \\ &\Rightarrow \int I(\gamma, y) d(\lambda F + (1 - \lambda)G) = 0 \Rightarrow \rho(\lambda F + (1 - \lambda)G) = \gamma, \end{aligned}$$

for each $\lambda \in (0, 1)$.

Theorem 36 Let $\tilde{\rho}: \mathcal{M}_{1,c}(0, +\infty) \rightarrow (0, +\infty)$ be an identifiable and positively homogeneous functional satisfying $\tilde{\rho}(1) = 1$. Then there exists $\Phi: [0, +\infty) \rightarrow \mathbb{R}$ satisfying $\Phi(1) = 1$ and $\Phi(x) > 1 \iff x > 1$ such that

$$\begin{aligned} \int \Phi(y/x) dF(y) = 1 &\iff x = \tilde{\rho}(F), \\ \int \Phi(y/x) dF(y) > 1 &\iff x < \tilde{\rho}(F). \end{aligned} \tag{19}$$

Furthermore, under these assumptions, if $\tilde{\rho}$ is monotone then Φ is nondecreasing and if $\tilde{\rho}$ is monotone and convex then Φ is nondecreasing and convex.

The proof of Theorem 36 is based on the following lemma, the proof of which is postponed to the Appendix.

Lemma 37 Let $\tilde{\rho}: \mathcal{M}_{1,c}(0, +\infty) \rightarrow (0, +\infty)$ be positively homogeneous with $\tilde{\rho}(1) = 1$ and identifiable by the function $I(x, y): (0, +\infty) \times (0, +\infty) \rightarrow \mathbb{R}$. Then there exists $g: (0, +\infty) \rightarrow \mathbb{R}$ with $g(1) = 0$ and $g(t) > 0 \iff t > 1$ with

$$\begin{aligned} \int g(y/x) dF(y) = 0 &\iff x = \tilde{\rho}(F), \\ \int g(y/x) dF(y) > 0 &\iff x < \tilde{\rho}(F). \end{aligned}$$

Proof of Theorem 36. By applying Lemma 37 and letting $\Phi(t) = 1 + g(t)$, equations (19) follow. To prove monotonicity of Φ , take $x_1, x_2 \in [0, +\infty)$ with $x_1 < x_2$. If $x_1 < 1 < x_2$, then $\Phi(x_1) < 1 < \Phi(x_2)$ so monotonicity is trivial. Assume that $1 < x_1 < x_2$. Take any $z < 1$ and set

$$\lambda := \frac{\Phi(x_2) - 1}{\Phi(x_2) - \Phi(z)} \in (0, 1).$$

Take $A \in \mathcal{F}$ with $P(A) = \lambda$ and let $X_1 := z1_A + x_11_{A^c}$ and $X_2 := z1_A + x_21_{A^c}$. By construction, $X_1 \leq X_2$ and $\mathbb{E}[\Phi(X_2)] = 1$. From the monotonicity of $\tilde{\rho}$ it follows that $\tilde{\rho}(X_1) \leq \tilde{\rho}(X_2) = 1$, so $\mathbb{E}[\Phi(X_1)] \leq 1 = \mathbb{E}[\Phi(X_2)]$, from which the thesis follows. The proof in the case $x_1 < x_2 < 1$ is similar. To prove convexity of Φ , we first show mid-convexity. Assume by contradiction that there exist $x_1, x_2 \in [0, +\infty)$ such that $\Phi((x_1 + x_2)/2) > (\Phi(x_1) + \Phi(x_2))/2$. Then there exist $z \in [0, +\infty)$ and $\lambda \in (0, 1)$ such that

$$\lambda\Phi(z) + (1 - \lambda)\Phi\left(\frac{x_1 + x_2}{2}\right) > 1 > \lambda\Phi(z) + (1 - \lambda)\frac{\Phi(x_1) + \Phi(x_2)}{2}.$$

Take disjoint sets A, B, C such that $P(A) = \lambda$, $P(B) = \frac{1-\lambda}{2}$ and $P(C) = \frac{1-\lambda}{2}$ and let $X_1 = z1_A + x_11_B + x_21_C$, $X_2 = z1_A + x_11_C + x_21_B$, and

$$X = z1_A + \frac{x_1 + x_2}{2}1_B + \frac{x_1 + x_2}{2}1_C = \frac{X_1 + X_2}{2}.$$

It holds that

$$\mathbb{E}[\Phi(X_1)] = \mathbb{E}[\Phi(X_2)] = \lambda\Phi(z) + (1 - \lambda)\left(\frac{\Phi(x_1) + \Phi(x_2)}{2}\right) < 1,$$

which implies $\tilde{\rho}(Y_1) < 1$ and $\tilde{\rho}(Y_2) < 1$. Similarly,

$$\mathbb{E}[\Phi(X)] = \lambda\Phi(z) + (1 - \lambda)\Phi\left(\frac{x_1 + x_2}{2}\right) > 1,$$

which implies $\tilde{\rho}(X) > 1$, contradicting convexity of $\tilde{\rho}$. Therefore, Φ is mid-convex. Since a nondecreasing mid-convex function is convex, the thesis follows. ■

5 Consistent scoring functions for Orlicz premia

In this section, we show that Orlicz premia are elicitable and we study general families as well as specific examples of strictly consistent scoring functions.

5.1 Elicitability and strict consistency

We start by recalling a few standard definitions adapted to the class of strictly positive return risk measures.

Definition 38 (Elicitability and strictly consistent scoring functions)

A functional $\rho: L_{++}^\infty \rightarrow (0, +\infty)$ is elicitable if there exists a strictly consistent scoring function $S: (0, +\infty) \times (0, +\infty) \rightarrow [0, +\infty)$ satisfying $S(x, y) \geq 0$ and $S(x, y) = 0$ if and only if $x = y$ such that, for each $Y \in L_{++}^\infty$, it holds that

$$\rho(Y) = \arg \min_{x \in (0, +\infty)} \mathbb{E}[S(x, Y)]. \quad (20)$$

The strictly consistent scoring function S is said to be of the prediction error form if $S(x, y) = f(x - y)$ and of the relative error form if $S(x, y) = g(y/x)$, where f, g are functions of a single variable.

The following theorem provides a general, rich family of scoring functions that are strictly consistent with Orlicz premia.

Theorem 39 Let Φ be as in Definition 23, with Φ increasing. Let $h: (0, +\infty) \rightarrow (0, +\infty)$ be any integrable function. Then,

$$S_\Phi(x, y) = \int_x^y h(z)(\Phi(y/z) - 1) dz \quad (21)$$

is a strictly consistent scoring function for the Orlicz premium H_Φ .

Proof. Let $x > H_\Phi(Y)$. We compute

$$\begin{aligned} & \mathbb{E}[S_\Phi(x, Y)] - \mathbb{E}[S_\Phi(H_\Phi(Y), Y)] \\ &= \mathbb{E} \left[\int_x^Y h(z)(\Phi(Y/z) - 1) dz \right] - \mathbb{E} \left[\int_{H_\Phi(Y)}^Y h(z)(\Phi(Y/z) - 1) dz \right] \\ &= -\mathbb{E} \left[\int_{H_\Phi(Y)}^x h(z)(\Phi(Y/z) - 1) dz \right] = -\int_{H_\Phi(Y)}^x h(z) \mathbb{E}[\Phi(Y/z) - 1] dz > 0, \end{aligned}$$

where in the last line we have used Fubini's Theorem and item (c) of Proposition 24. The same argument shows that if $x < H_\Phi(Y)$, then again

$$\mathbb{E}[S_\Phi(x, Y)] - \mathbb{E}[S_\Phi(H_\Phi(Y), Y)] = \int_x^{H_\Phi(Y)} h(z) \mathbb{E}[\Phi(Y/z) - 1] dz > 0,$$

so S_Φ is a strictly consistent scoring function for the Orlicz premium. ■

Two particularly interesting cases arise by taking $h(z) = 1/z$ and $h(z) = 1/z^2$. In the first case, with the change of variable $y/z = e^t$, we obtain

$$\begin{aligned} S_\Phi(x, y) &= \int_x^y \frac{1}{z} (\Phi(y/z) - 1) dz = \int_0^{\log(y/x)} (\Phi(e^t) - 1) dt \\ &= \varphi(\log(y/x)), \end{aligned} \quad (22)$$

where

$$\varphi(t) = \int_0^t [\Phi(e^s) - 1] ds. \quad (23)$$

In the second case, with the change of variable $y/z = t$, we obtain

$$S_{\Phi}(x, y) = \int_1^{y/x} \frac{\Phi(t) - 1}{y} dt. \quad (24)$$

We now present a collection of examples of strictly consistent scoring functions for Orlicz premia, using Orlicz functions that are commonly adopted in the literature. As we will see, in some cases the general approach based on Theorem 39 or the specific forms in equations (22) and (24) recover scoring functions that are already known in the literature, whereas in many other cases new families of scoring functions are obtained. In several cases the corresponding Orlicz premium H_{Φ} admits an explicit expression, whereas in some other cases it has to be computed numerically. All our examples, including some not discussed below, are summarized in Table 1.

Example 40 (Mean) Let $\Phi(x) = x$. Then, $H_{\Phi}(Y) = \mathbb{E}[Y]$ and, from (22) and (23), we obtain

$$S_{\Phi}(x, y) = \frac{y}{x} - \log\left(\frac{y}{x}\right) - 1, \quad (25)$$

which is an alternative scoring function for the mean. From the classical result of [41], as recalled e.g., in Theorem 7 of [26], the most general class of strictly consistent scoring functions for the mean belongs to the family of Bregman functions, of the form

$$S(x, y) = \phi(y) - \phi(x) - \phi'(x)(y - x),$$

where ϕ is a convex function with subgradient ϕ' . The scoring function given in equation (25) arises if $\phi(x) = -\log x$. This scoring function is known as the quasi-likelihood (QLIKE) scoring function in the econometrics literature, and it is of common use in assessing forecasts of nonnegative quantities such as volatility (see e.g., [38] and the references therein).

Example 41 (Expectiles) Let $\Phi_q(x) = 1 + q(x - 1)^+ - (1 - q)(x - 1)^-$ with $0 < q < 1$, as in Example 25. The corresponding Orlicz premium is the q -expectile and, from (22) and (23), the corresponding scoring function is given by

$$S_{\Phi}(x, y) = q(y/x - \log(y/x) - 1)^+ + (1 - q)(y/x - \log(y/x) - 1)^-. \quad (26)$$

The class of strictly consistent scoring functions for expectiles has been characterized in Theorem 10 of [26], as an asymmetric extension of Bregman functions, defined by

$$S(x, y) = |1_{x \geq y} - q| \{ \phi(y) - \phi(x) - \phi'(x)(y - x) \},$$

and as in the case of the mean our scoring function (26) corresponds to the case $\phi(x) = -\log x$.

Example 42 (p -norms) Let $\Phi(x) = x^p$ with $p \geq 1$. Then, the Orlicz premium H_Φ is given by

$$H_\Phi(Y) = \|Y\|_p,$$

and the corresponding scoring function from (22) and (23) takes the form

$$S_\Phi(x, y) = \frac{1}{p} \left(\frac{y^p}{x^p} - 1 \right) - \log \left(\frac{y}{x} \right).$$

In view of (25), which arises as a special case when $p \equiv 1$, we will refer to this novel scoring function as the *PQLIKE* scoring function.

When $0 < p < 1$, the resulting Orlicz premium is no longer a norm, but the scoring function is still valid. Taking a (suitably normalized and scaled) limit for $p \rightarrow 0$, such that $\Phi(x) = 1 + \log(x)$ arises, the Orlicz premium H_Φ is the logarithmic certainty equivalent given by

$$H_\Phi(Y) = \exp(\mathbb{E}[\log Y]).$$

The corresponding scoring function from (22) and (23) takes the form

$$S_\Phi(x, y) = \left(\log \left(\frac{y}{x} \right) \right)^2.$$

Example 43 (Mean-variance) Let $\Phi(x) = \lambda x + (1 - \lambda)x^2$ with $0 \leq \lambda \leq 1$ as in Section 5.1 of [28]. Then, the Orlicz premium H_Φ is given by

$$H_\Phi(Y) = \mathbb{E}[Y] \left(\frac{\lambda}{2} + \sqrt{\left(\frac{\lambda}{2} - 1 \right)^2 + (1 - \lambda) \frac{\text{Var}[Y]}{\mathbb{E}^2[Y]}} \right),$$

and the corresponding scoring function from (22) and (23) is

$$S_\Phi(x, y) = \frac{\lambda y}{x} + \frac{(1 - \lambda) y^2}{2 x^2} - \log \left(\frac{y}{x} \right) - \frac{\lambda + 1}{2}.$$

Example 44 Let $\Phi(x) = \frac{e^{\alpha x} - 1}{e^\alpha - 1}$ with $\alpha > 0$ as in Section 5.2 of [28]. The corresponding Orlicz premium H_Φ does not in general admit an explicit expression. It is the solution of the equation $f_Y(\alpha/H_\Phi) = e^\alpha$, where $f_Y(t) = \mathbb{E}[\exp(tY)]$ is the moment generating function of Y . For example, if Y has a Gamma distribution with shape parameter $\theta > 0$ and rate parameter $\gamma > 0$, we have

$$H_\Phi(Y) = \frac{\alpha}{\gamma(1 - \exp(-\alpha/\theta))} = \frac{\alpha e^{\alpha/\theta}}{\theta(e^{\alpha/\theta} - 1)} \mathbb{E}[Y].$$

The corresponding scoring function from (24) is given by

$$S_\Phi(x, y) = \frac{e^\alpha}{(e^\alpha - 1)} \left(\frac{1}{\alpha y} \left(e^{\alpha \left(\frac{y}{x} - 1 \right)} - 1 \right) + \frac{1}{y} - \frac{1}{x} \right).$$

Orlicz Function $\Phi(x)$	Orlicz Premium $H_\Phi(Y)$	Scoring Function $S_\Phi(x, y)$
x	$\mathbb{E}[Y]$	$\frac{y}{x} - \log\left(\frac{y}{x}\right) - 1$
$\alpha + \mathbb{1}_{\{x>1\}}, 0 < \alpha < 1$	$q_\alpha(Y)$	$(\mathbb{1}_{\{x \geq y\}} - \alpha) \log\left(\frac{x}{y}\right)$
$1 + q(x-1)^+ - (1-q)(x-1)^-, 0 < q < 1$	$e_q(Y)$	$q\left(\frac{y}{x} - \log\left(\frac{y}{x}\right) - 1\right)^+ + (1-q)\left(\frac{y}{x} - \log\left(\frac{y}{x}\right) - 1\right)^-$
$1 + \log(x)$	$\exp(\mathbb{E}[\log Y])$	$(\log\left(\frac{y}{x}\right))^2$
$x^p, p \geq 1$	$\ Y\ _p$	$\frac{1}{p}\left(\frac{y^p}{x^p} - 1\right) - \log\left(\frac{y}{x}\right)$
$\lambda x^p + (1-\lambda)x^{2p}, 0 \leq \lambda \leq 1, p \geq 1$	$\left(\frac{1}{2}\left(\lambda \mathbb{E}[Y^p] + \sqrt{\lambda^2 \mathbb{E}[Y^p]^2 + 4(1-\lambda)\mathbb{E}[Y^{2p}]}\right)\right)^{\frac{1}{p}}$	$\frac{1-\lambda}{2p}\left(\frac{y}{x}\right)^{2p} + \frac{\lambda}{p}\left(\frac{y}{x}\right)^p - \log\left(\frac{y}{x}\right) - \frac{\lambda+1}{2p}$
$x \log(e-1+x)$	no explicit form	$\left(\frac{y}{x} + e - 1\right) \log\left(\frac{y}{x} + e - 1\right) - \log\left(\frac{y}{x}\right)$
$x^p + \mathbb{1}_{\{x>1\}}x^p \log(x), p \geq 1$	no explicit form	$-\log\left(\frac{y}{x}\right) + \frac{1}{p}\left(\frac{y^p}{x^p} - 1\right) + \mathbb{1}_{\{y>x\}}\left(\frac{1}{p}\frac{y^p}{x^p} \log\left(\frac{y}{x}\right) - \frac{1}{p^2}\left(\frac{y^p}{x^p} - 1\right)\right)$
$\frac{e^{\alpha x} - 1}{e^{\alpha} - 1}, \alpha > 0$	$\frac{\alpha e^{\alpha/\theta}}{\theta(e^{\alpha/\theta} - 1)} \mathbb{E}[Y]$	$\frac{e^{-\alpha}}{(e^{\alpha} - 1)}\left(\frac{1}{\alpha y}\left(e^{\alpha\left(\frac{y}{x^2} - 1\right)} - 1\right) + \frac{1}{y} - \frac{1}{x}\right)$
$x e^{\alpha(x^2-1)}, \alpha > 0$	$\sqrt{\frac{2}{\pi}}\left(\frac{1}{2}e^{-\alpha} + \sqrt{\frac{1}{4}e^{-2\alpha} + \pi\alpha}\right)\sigma$	$\frac{1}{2\alpha y}\left(e^{\alpha\left(\frac{y^2}{x^2} - 1\right)} - 1\right) + \frac{1}{y} - \frac{1}{x}$
$\frac{e^x - x - 1}{e^{-2}}$	no explicit form	$\frac{1}{e^{-2}}\left(\frac{e^x - 1}{y} + \frac{(2-2e)^{x-y}}{2x^2}\right)$

Table 1: Examples of Orlicz functions with the corresponding Orlicz premia and the corresponding strictly consistent scoring functions. Eqn. (22) is used to obtain the scoring functions of the first eight examples. For the remaining examples, (24) is used. Although $\Phi(x) = \alpha + \mathbb{1}_{\{x>1\}}$ does not satisfy the conditions of Theorem 39, it still gives rise to a strictly consistent scoring function. When $\Phi(x) = \frac{e^{\alpha x} - 1}{e^{\alpha} - 1}$, the corresponding Orlicz premium does not in general admit an explicit expression. The expression given in the table corresponds to the case in which Y follows a Gamma distribution with shape parameter $\theta > 0$ and rate parameter $\gamma > 0$. When $\Phi(x) = x e^{\alpha(x^2-1)}$, the corresponding Orlicz premium also does not in general admit an explicit expression. The given expression corresponds to the case in which $Y = |Z|$, where Z follows a normal distribution with mean 0 and standard deviation σ .

5.2 Mixture representations

It is clear that a scoring function of the form (21) depends on the choice of the function h . Hence, the ranking of competing forecasts may depend on this choice, in particular in finite samples and under model misspecification (see e.g., [39]). To remedy the dependence of the ranking on the specific choice of the strictly consistent scoring function, [20] develop a method to compare forecasts simultaneously with respect to a class of strictly consistent scoring functions by considering so-called Murphy diagrams. This method relies on the availability of a mixture representation of the strictly consistent scoring functions under consideration, in terms of elementary scoring functions depending on a low-dimensional parameter. Mixture representations for the class of strictly consistent scoring functions for quantiles and expectiles have been given in [20]. A mixture representation of strictly consistent scoring functions for the triplet of Range Value-at-Risk and its two associated Value-at-Risks has been given in [22]. In the following theorem, we provide such a mixture representation for our new family of scoring functions in (21).

Theorem 45 *Any strictly consistent scoring function for the Orlicz premium H_Φ of the form (21) admits a representation of the form*

$$S_\Phi(x, y) = \int_0^{+\infty} S_z(x, y) dH(z), \quad (27)$$

for a positive measure H , where $S_z(x, y) = |\Phi(\frac{y}{z}) - 1| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}}$. Conversely, for any choice of the positive measure H , we obtain a strictly consistent scoring function of the form (21) for the Orlicz premium H_Φ .

Proof. By using (21), we have

$$\begin{aligned} S_\Phi(x, y) &= \int_x^y h(z) (\Phi(y/z) - 1) dz \\ &= \int_x^y (\Phi(y/z) - 1) dH(z) \\ &= \int_0^{+\infty} S_z(x, y) dH(z), \end{aligned}$$

where $dH(z) = h(z) dz$ and $S_z(x, y) = |\Phi(\frac{y}{z}) - 1| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}}$. Note that, since the function h in (21) is strictly positive, the Riemann integral $\int_0^{+\infty} S_z(x, y) dH(z)$ is well-defined. ■

As a corollary, we provide a mixture representation for the scoring functions of p -norms.

Corollary 46 *Any strictly consistent scoring function of the form (21) with $\Phi(x) = x^p$, $p \geq 1$, admits a representation of the form:*

$$S_\Phi(x, y) = \int_0^{+\infty} S_z^p(x, y) dH(z), \quad (28)$$

for a positive measure H , where $S_z^p(x, y) = |y^p - z^p| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}}$.

Proof. From Theorem 45 with $\Phi(x) = x^p$, we obtain

$$\begin{aligned} S(x, y) &= \int_0^{+\infty} \left| \frac{y^p}{z^p} - 1 \right| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}} dH(z) \\ &= \int_0^{+\infty} |y^p - z^p| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}} d\tilde{H}(z), \end{aligned}$$

where $d\tilde{H}(z) := \frac{1}{z^p} dH(z)$. ■

We conduct two simulation experiments to illustrate how one can use Theorem 45 to rank competing forecasts. In particular, we will generate Murphy diagrams for logarithmic certainty equivalents, p -norms and expectiles by using the corresponding elementary scoring functions.

Example 47 We first suppose that the true distribution of the outcome variable Y is given by $\log(Y)|\mu \sim \mathcal{N}(\mu, \sigma_Y^2)$ where $\mu \sim \mathcal{N}(0, \sigma_\mu^2)$. We take $\sigma_Y = \sigma_\mu = 0.2$. We consider four different forecasters, who will be referred to as perfect, unconditional, unfocused and sign-reversed, similar to [25] and [20], suitably modified to the current setting. The perfect forecaster issues the true distribution of the outcome Y as predictive distribution. Therefore, his/her point forecasts of the logarithmic certainty equivalent (LCE) and p -norm are $\exp(\mu)$ and $\exp(\mu + \sigma_Y^2 p/2)$ with $\sigma_Y = 0.2$. The unconditional forecaster does not have knowledge of μ and issues the unconditional distribution of $\log(Y)$ as predictive distribution: $\mathcal{N}(0, \sigma_\mu^2 + \sigma_Y^2)$. Therefore, his/her point forecasts of the LCE and p -norm are $\exp(0)$ and $\exp((\sigma_\mu^2 + \sigma_Y^2)p/2)$ with $\sigma_Y = \sigma_\mu = 0.2$. The remaining two forecasters, unfocused and sign-reversed, have knowledge of μ , but their predictive distributions fail to be ideal. The unfocused forecaster issues a mixture distribution as predictive distribution of $\log(Y)$, involving an independent random variable τ that takes the values 0.2 and -0.2 with probability 1/2, leading to $\frac{1}{2}(\mathcal{N}(\mu, \sigma_Y^2) + \mathcal{N}(\mu + \tau, \sigma_Y^2))$ yielding $\exp(\mu + \tau/2)$ and $\exp(\mu + \tau/2 + \sigma_Y^2 p/4)$ with $\sigma_Y = 0.2$ as the corresponding forecasts of the LCE and p -norm. The sign-reversed forecaster issues a predictive distribution of $\log(Y)$ with the sign of μ flipped: $\mathcal{N}(-\mu, \sigma_Y^2)$. Therefore, his/her point forecasts are $\exp(-\mu)$ and $\exp(-\mu + \sigma_Y^2 p/2)$. The point forecasts generated by the four predictive distributions are summarized in Table 2. Using 10,000 simulations of sample size 1,000 each, we obtain the Murphy diagrams displayed in Figure 2 for the LCE and p -norm with $p = 1, 2, 3$. As can be seen in Figure 2, the perfect forecaster dominates the other forecasters for the LCE and all p -norms considered, as expected. Although not clearly visible, the expected scores for the other three forecasters intersect in all four cases, such that none of these forecasters dominates the other.

Example 48 For q -expectiles, from Theorem 45 with $\Phi_q(x) = 1 + q(x - 1)^+ -$

Forecaster	Predictive distribution of $\log(Y)$	Point forecast of p -norm
Perfect	$\mathcal{N}(\mu, \sigma_Y^2)$	$\exp(\mu + \sigma_Y^2 p/2)$
Unconditional	$\mathcal{N}(0, \sigma_\mu^2 + \sigma_Y^2)$	$\exp((\sigma_\mu^2 + \sigma_Y^2)p/2)$
Unfocused	$\frac{1}{2}(\mathcal{N}(\mu, \sigma_Y^2) + \mathcal{N}(\mu + \tau, \sigma_Y^2))$	$\exp(\mu + \tau/2 + \sigma_Y^2 p/4)$
Sign-reversed	$\mathcal{N}(-\mu, \sigma_Y^2)$	$\exp(-\mu + \sigma_Y^2 p/2)$

Table 2: Predictive distributions and point forecasts. The point forecasts for the LCE arise by taking $p \equiv 0$.

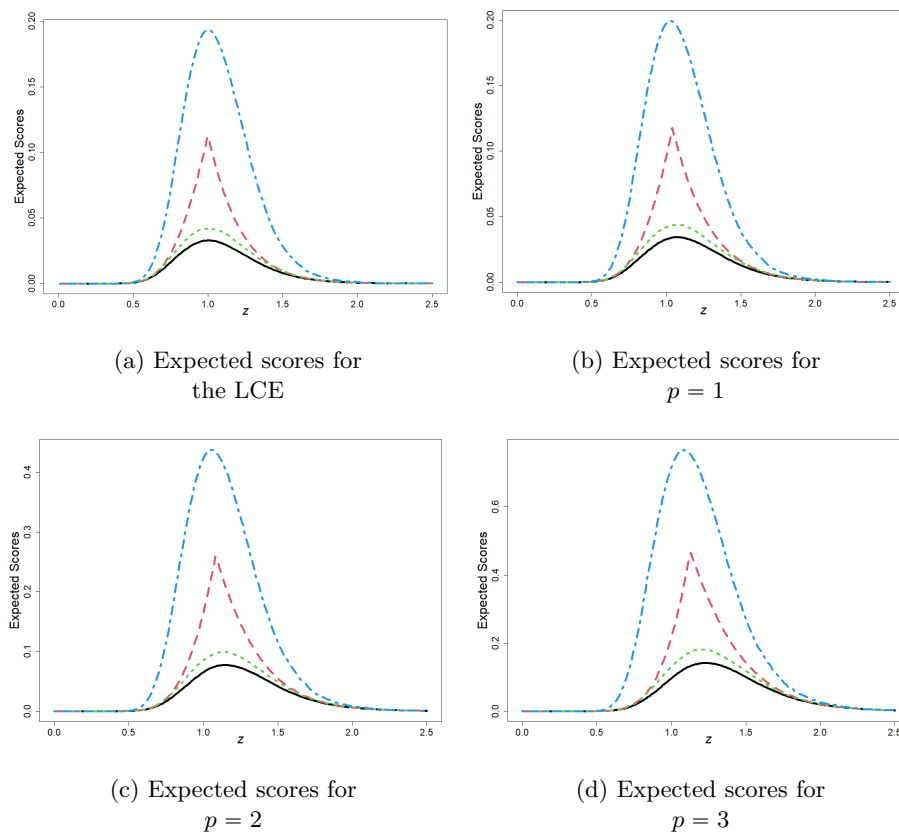


Figure 2: Murphy diagrams of the four forecasters for the LCE and different p -norms (perfect: black; unconditional: red; unfocused: green; sign-reversed: blue)

$(1 - q)(x - 1)^-$, $0 < q < 1$, we have

$$\begin{aligned}
S(x, y) &= \int_0^{+\infty} \left| q \left(\frac{y}{z} - 1 \right)^+ - (1 - q) \left(\frac{y}{z} - 1 \right)^- \right| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}} dH(z) \\
&= \int_0^{+\infty} \left| q(y - z)^+ - (1 - q)(y - z)^- \right| 1_{\{x \leq z < y\} \cup \{y \leq z < x\}} d\tilde{H}(z),
\end{aligned}$$

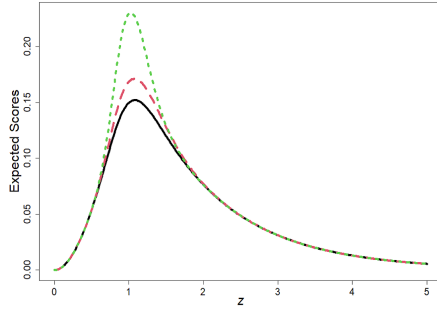
where $d\tilde{H}(z) := \frac{1}{z}dH(z)$. Therefore, the elementary scoring function for expectiles can be expressed as

$$S_z^q(x, y) = q(y - z)1_{\{x \leq z < y\}} + (1 - q)(z - y)1_{\{y \leq z < x\}}. \quad (29)$$

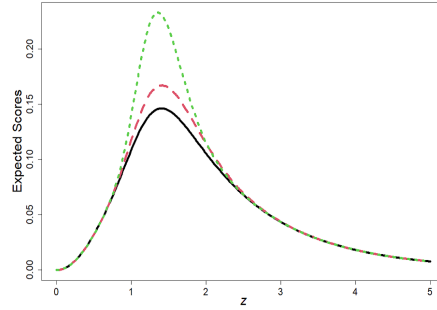
Suppose now that the true distribution of the outcome variable Y is given by $Y|\lambda \sim \exp(\lambda)$ where $\log(\lambda) \sim \mathcal{N}(0, \sigma_\lambda^2)$. We take $\sigma_\lambda = 0.2$. We consider three different forecasters: perfect, unfocused and mean-reversed, similar to Example 47. The perfect forecaster issues the true distribution of Y as predictive distribution. The unfocused forecaster issues $\exp(\tau\lambda)$ as predictive distribution, involving an independent random variable τ that takes the values $5/4$ and $4/5$ each with probability $1/2$. The mean-reversed forecaster issues a predictive distribution with the mean reversed: $\exp(1/\lambda)$. The point forecasts generated by the three predictive distributions are displayed in Table 3. Using 10,000 simulations of sample size 1,000 each, we obtain the Murphy diagrams displayed in Figure 3 for $q = 0.5, 0.7, 0.9, 0.95$. As we see from the figure, the perfect forecaster dominates the other forecasters, as expected. There is no ordering relationship between the unfocused and mean-reversed forecasters, because their expected scores intersect.

Forecaster	Predictive distribution of Y	Point forecast of q -expectile
Perfect	$\exp(\lambda)$	$\frac{1}{\lambda} \left(1 + W\left(\frac{2q-1}{(1-q)e}\right) \right)$
Unfocused	$\exp(\tau\lambda)$	$\frac{1}{\tau\lambda} \left(1 + W\left(\frac{2q-1}{(1-q)e}\right) \right)$
Mean-Reversed	$\exp(1/\lambda)$	$\lambda \left(1 + W\left(\frac{2q-1}{(1-q)e}\right) \right)$

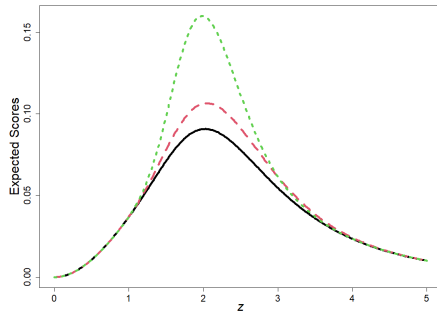
Table 3: Predictive distributions and point forecasts. W denotes the Lambert function.



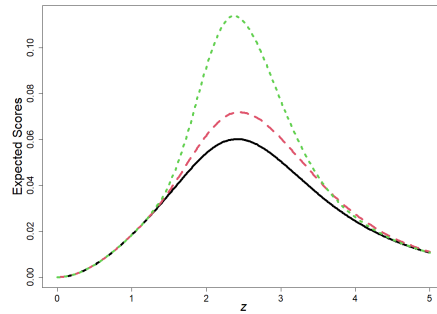
(a) Expected scores for $q = 0.5$



(b) Expected scores for $q = 0.7$



(c) Expected scores for $q = 0.9$



(d) Expected scores for $q = 0.95$

Figure 3: Murphy diagrams of the three forecasters for different q -expectiles (perfect: black; unfocused: green; mean-reversed: red)

6 Appendix

Lemma 49 *Let $f: (0, +\infty) \rightarrow \mathbb{R}$ be nondecreasing and convex. Then, f is GA-convex.*

Proof. For $x, y > 0$ and $\lambda \in (0, 1)$ from the AM-GM inequality it holds that $\lambda x + (1 - \lambda)y \geq x^\lambda y^{1-\lambda}$. Since f is nondecreasing and convex it follows that

$$f(x^\lambda y^{1-\lambda}) \leq f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y),$$

which gives the thesis. ■

Proof of Lemma 37. Since $\tilde{\rho}$ is law invariant and positively homogeneous with $\tilde{\rho}(1) = 1$, it follows that $\tilde{\rho}(\delta_y) = y$, for each $y > 0$. From identifiability, it follows that

$$\begin{aligned} I(x, y) = 0 &\iff x = y, \\ I(x, y) > 0 &\iff x < y. \end{aligned}$$

For each $0 < y_1 < x < y_2$, define

$$\begin{aligned} \bar{p} &= \frac{I(x, y_2)}{I(x, y_2) - I(x, y_1)} \in (0, 1), \\ \bar{F} &= \bar{p}\delta_{y_1} + (1 - \bar{p})\delta_{y_2}. \end{aligned}$$

Since

$$\begin{aligned} \int I(x, y) d\bar{F}(y) &= \bar{p}I(x, y_1) + (1 - \bar{p})I(x, y_2) \\ &= \frac{I(x, y_1)I(x, y_2)}{I(x, y_2) - I(x, y_1)} + \frac{-I(x, y_2)I(x, y_1)}{I(x, y_2) - I(x, y_1)} = 0, \end{aligned}$$

it follows that $\tilde{\rho}(\bar{F}) = x$. From the positive homogeneity of $\tilde{\rho}$, it follows that for each $\lambda > 0$,

$$\tilde{\rho}((1 - \bar{p})\delta_{\lambda y_1} + \bar{p}\delta_{\lambda y_2}) = \lambda x,$$

and from identifiability

$$\int I(\lambda x, y) d[(1 - \bar{p})\delta_{\lambda y_1} + \bar{p}\delta_{\lambda y_2}] = 0,$$

which gives

$$\frac{-I(x, y_2)I(\lambda x, \lambda y_1)}{I(x, y_1) - I(x, y_2)} + \frac{I(x, y_1)I(\lambda x, \lambda y_2)}{I(x, y_1) - I(x, y_2)} = 0,$$

so we can conclude that

$$\frac{I(x, y_1)}{I(x, y_2)} = \frac{I(\lambda x, \lambda y_1)}{I(\lambda x, \lambda y_2)},$$

and letting $\lambda = 1/x$ we find that

$$\frac{I(x, y_1)}{I(x, y_2)} = \frac{I(1, y_1/x)}{I(1, y_2/x)}, \quad (30)$$

for each $0 < y_1 < x < y_2$.

We now want to prove that

$$I(x, y) = g(y/x) \cdot h(x), \quad (31)$$

with $h(x) > 0$, from which the thesis follows immediately. We consider two cases.

If $y > x$, we set $y_1 = x/2$ and $y_2 = y$ in (30), obtaining

$$\frac{I(x, x/2)}{I(x, y)} = \frac{I(1, 1/2)}{I(1, y/x)},$$

which gives

$$I(x, y) = I(1, y/x) \cdot \frac{I(x, x/2)}{I(1, 1/2)}. \quad (32)$$

If instead $y < x$, we set $y_1 = y$ and $y_2 = 2x$ in (30), obtaining

$$\frac{I(x, y)}{I(x, 2x)} = \frac{I(1, y/x)}{I(1, 2)},$$

which gives

$$I(x, y) = I(1, y/x) \cdot \frac{I(x, 2x)}{I(1, 2)}. \quad (33)$$

Notice also that from (30) it follows that

$$\frac{I(x, x/2)}{I(1, 1/2)} = \frac{I(x, 2x)}{I(1, 2)},$$

so combining (32) and (33) it follows that (31) is satisfied with

$$\begin{aligned} g(t) &= I(1, t), \\ h(x) &= \frac{I(x, x/2)}{I(1, 1/2)}, \end{aligned}$$

from which the thesis follows. ■

References

- [1] ALIPRANTIS, C. D., BORDER, K. C. (2006). *Infinite Dimensional Analysis*. Third edition, Springer Verlag, Berlin.
- [2] BELLINI, F., BIGNOZZI, V. (2015). On elicitable risk measures. *Quantitative Finance* 15(5), 725-733.

- [3] BELLINI, F., ROSAZZA GIANIN, E. (2008). On Haezendonck risk measures. *Journal of Banking and Finance* 32(6), 986–994.
- [4] BELLINI, F., LAEVEN, R. J. A., ROSAZZA GIANIN, E. (2018). Robust return risk measures. *Mathematics and Financial Economics* 12(1), 5-32.
- [5] BELLINI, F., LAEVEN, R. J. A., ROSAZZA GIANIN, E. (2021). Dynamic robust Orlicz premia and Haezendonck-Goovaerts risk measures. *European Journal of Operational Research* 291(2), 438-446.
- [6] BEN-TAL, A., TEBoulLE, M. (1986). Expected utility, penalty functions, and duality in stochastic nonlinear programming. *Management Science* 32(11), 1445-1466.
- [7] BEN-TAL, A., TEBoulLE, M. (2007). An old-new concept of convex risk measures: The optimized certainty equivalent. *Mathematical Finance* 17(3), 449-476.
- [8] BIAGINI, S., FRITTELLI, M. (2008). A unified framework for utility maximization problems: An Orlicz space approach. *The Annals of Applied Probability* 18(3), 929-966.
- [9] CANNA, G., CENTRONE, F., ROSAZZA GIANIN, E. (2021). Haezendonck-Goovaerts capital allocation rules. *Insurance: Mathematics and Economics* 101, Part B, 173-185.
- [10] CHERIDITO, P., LI, T. (2008). Dual characterization of properties of risk measures on Orlicz hearts. *Mathematics and Financial Economics* 2(29), 29-55.
- [11] CHERIDITO, P., LI, T. (2009). Risk measures on Orlicz Hearts. *Mathematical Finance* 19(2), 189-214.
- [12] CONT, R., DEGUEST, R., SCANDOLO, G. (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance* 10(6), 593-606.
- [13] CSISZÁR, I. (1975). I -Divergence geometry of probability distributions and minimization problems. *Annals of Probability* 3(1), 146-158.
- [14] DELBAEN, F. (2000). Coherent Risk Measures. Lecture Notes of the Cattedra Galileiana at the Scuola Normale Superiore di Pisa.
- [15] DELBAEN, F. (2002). Coherent Risk Measures on General Probability Spaces. *Advances in Finance and Stochastics*. Springer Verlag, Berlin, 1-37.
- [16] DELBAEN, F. (2012). Monetary Utility Functions. Osaka University Press, Osaka.
- [17] DELBAEN, F., BELLINI, F., BIGNOZZI, V., ZIEGEL, J. F. (2014). Risk measures with the CxLS property. *Finance and Stochastics* 20(2), 433-453.
- [18] DRAPEAU, S., KUPPER, M. (2013). Risk preferences and their robust representation. *Mathematics of Operations Research* 38(1), 28-62.
- [19] EDGAR, G. A., SUCHESTON, L. (1992). Stopping Times and Directed Processes. Cambridge University Press, Cambridge.
- [20] EHM, W., GNEITING, T., JORDAN, A., KRÜGER, F. (2016). Of quantiles and expectiles: Consistent scoring functions, Choquet representations and forecast rankings. *Journal of the Royal Statistical Society Series B* 78(3), 505-562.
- [21] FISSLER, T., ZIEGEL, J. F. (2016). Higher order elicibility and Osband’s principle. *The Annals of Statistics* 44(4), 1680-1707.
- [22] FISSLER, T., ZIEGEL, J. F. (2021). On the elicibility of range value at risk. *Statistics and Risk Modelling* 38(1-2), 25-46.
- [23] FÖLLMER, H., SCHIED, A. (2011). Stochastic Finance. Third edition, De Gruyter, Berlin.

- [24] GAO, N., LEUNG, D. H., MUNARI, C., XANTHOS, F. (2018). Fatou property, representations, and extensions of law-invariant risk measures on general Orlicz spaces. *Finance and Stochastics* 22(2), 395–415.
- [25] GNEITING, T., BALABDAOUI, F., RAFTERY, A. E. (2007). Probabilistic forecasts, calibration, sharpness. *Journal of the Royal Statistical Society Series B* 69(2), 243-268.
- [26] GNEITING, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association* 106(494), 746-762.
- [27] GOOVAERTS, M. J., DE VYLDER, F. E. C., HAEZENDONCK, J. (1984). Insurance Premiums. North-Holland Publishing, Amsterdam.
- [28] HAEZENDONCK, J., GOOVAERTS, M. J. (1982). A new premium calculation principle based on Orlicz norms. *Insurance: Mathematics and Economics* 1(1), 41-53.
- [29] JOUINI, E., SCHACHERMAYER, W., TOUZI, N. (2006). Law invariant risk measures have the Fatou property. In: Kusuoka, S., Yamazaki, A. (Eds.). *Advances in Mathematical Economics* 9, 49–71.
- [30] KUAN, C.-M., YEH, J.-H., HSU, Y.-C. (2009). Assessing value at risk with CARE, the Conditional Autoregressive Expectile models. *Journal of Econometrics* 150(2), 261-270.
- [31] LAEVEN, R. J. A., STADJE, M. (2013). Entropy coherent and entropy convex measures of risk. *Mathematics of Operations Research* 38(2), 265-293.
- [32] LAEVEN, R. J. A., STADJE, M. (2014). Robust portfolio choice and indifference valuation. *Mathematics of Operations Research* 39(4), 1109-1141.
- [33] MOHAMMED, N., FURMAN, E., SU, J. (2021). Can a regulatory risk measure induce profit-maximizing risk capital allocations? The case of conditional tail expectation. *Insurance: Mathematics and Economics* 101, Part B, 425-436.
- [34] NEWEY, W. K., POWELL, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica* 55(4), 819-847.
- [35] NICULESCU, C. P. (2000). Convexity according to the geometric mean. *Mathematical Inequalities & Applications* 3(2), 155-167.
- [36] NOLDE, N., ZIEGEL, J. F. (2017). Elicitability and backtesting: Perspectives for banking regulation. *The Annals of Applied Statistics* 11(4), 1833-1874.
- [37] OSBAND, K. H. (1985). Providing Incentives for Better Cost Forecasting. PhD Thesis, University of California, Berkeley.
- [38] PATTON, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics* 160(1), 246-256.
- [39] PATTON, A. J. (2020). Comparing possibly misspecified forecasts. *Journal of Business & Economic Statistics* 38(4), 796–809.
- [40] ROCKAFELLAR, R. T., URYASEV, S. P. (2000). Optimization of conditional value-at-risk. *Journal of Risk* 2(3), 21-42.
- [41] SAVAGE, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association* 66(336), 783–801.
- [42] WU, Q., XU, H. (2022). Robust distorted Orlicz premium: Modelling, computational scheme and applications. Available at SSRN: <https://ssrn.com/abstract=4093580> or <http://dx.doi.org/10.2139/ssrn.4093580>
- [43] WEBER, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance* 16(2), 419-441.

- [44] ZĂLINESCU, C. (2002). *Convex Analysis in General Vector Spaces*. World Scientific, Singapore.
- [45] ZIEGEL, J. F. (2016). Coherence and elicibility. *Mathematical Finance* 26(4), 901-918.