Conditional expectiles, time consistency and mixture convexity properties

Fabio Bellini[1], Valeria Bignozzi[1,*], and Giovanni Puccetti[2]

[1] *Department of Statistics and Quantitative Methods, University of Milano-Bicocca, Italy*
[2] *Department of Economics, Management and Quantitative Methods, University of Milano, Italy*
[*] *Corresponding author, email: valeria.bignozzi@unimib.it*

July 11, 2018

Abstract

We study conditional expectiles, defined as a natural generalisation of conditional expectations by means of the minimisation of an asymmetric quadratic loss function. We show that conditional expectiles can be equivalently characterised by a conditional first order condition and we derive their main properties. For possible applications as dynamic risk measures, we discuss their time consistency properties.

AMS 2010 Subject Classification: 91B30 (primary), 91G99 (secondary).

Jel Classification: D81.

Keywords: Conditional expectiles, dynamic risk measures, mixture concavity, time consistency, sequential consistency, supermartingale property.

1 Introduction

Expectiles have been introduced in the statistical literature by Newey and Powell (1987) as a one parameter family of statistical functionals that includes the mean as a special case. They are defined, for a random variable X in $L^2(\Omega, \mathcal{F}, \mathbb{P})$, by the following minimisation problem:

$$e_\alpha(X) = \underset{x\in\mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\left[\alpha(X-x)_+^2 + (1-\alpha)(X-x)_-^2\right],$$

where $\alpha \in (0,1)$ and $x_+ = \max(x,0)$, $x_- = \max(-x,0)$. Expectiles have many similarities with the left and right quantiles $q_\alpha^-(X)$ and $q_\alpha^+(X)$, that are defined by the minimisation of an asymmetric piecewise linear function (see e.g. Koenker, 2005):

$$[q_\alpha^-(X), q_\alpha^+(X)] = \underset{x\in\mathbb{R}}{\operatorname{argmin}}\, \mathbb{E}\left[\alpha(X-x)_+ + (1-\alpha)(X-x)_-\right].$$

Expectiles are becoming increasingly popular in the financial and actuarial literature because they have many interesting properties related to their applications as risk measures.

By definition, expectiles are elicitable; indeed, a statistical functional is said to be elicitable if it is the minimiser of a suitable expected loss function (see e.g. Osband, 1985 or Gneiting, 2011). Elicitability is considered a useful property in the field of financial risk management since it provides a natural methodology to compare different forecasting models or to test the statistical hypotheses related to them; we refer to Acerbi and Szekely (2014), Bellini and Di Bernardino (2017), Fissler et al. (2016) and the references therein.

For $\alpha \geq 1/2$, expectiles are coherent risk measures; indeed, they satisfy the well known properties of translation invariance, positive homogeneity, monotonicity and subadditivity. In the literature there are various axiomatic foundations for the expectiles, as the unique coherent risk measures satisfying additional properties. Weber (2006) proved that they are the only shortfall risk measures that are coherent; Bellini et al. (2014) proved that they are the only generalised quantiles that are coherent; Ziegel (2016), Bellini and Bignozzi (2015) and Delbaen et al. (2016) from slightly different angles proved that they are the only elicitable risk measures that are coherent.

Expectiles have a straightforward financial interpretation as capital requirements, as it was outlined e.g. in Bellini and Di Bernardino (2017). Indeed, they represent the minimum amount of capital that has to be added to a financial loss to make it acceptable, where the acceptance set $\mathcal{A}$ is defined as

$$\mathcal{A} = \left\{ X \in L^1 \text{ such that } \frac{\mathbb{E}[X_+]}{\mathbb{E}[X_-]} \leq \frac{1-\alpha}{\alpha} \right\}.$$

From an actuarial point of view, they are a special case of the zero utility premium principle, since they satisfy the equation

$$\mathbb{E}\left[\ell_\alpha(X - e_\alpha(X))\right] = 0$$

with $\ell_\alpha(x) = ax_+ - (1-\alpha)x_-$.

In the present paper we focus on the notion of conditional expectile, that we define as a natural asymmetric generalisation of the notion of conditional expectation:

$$e_\alpha^{\mathcal{G}}(X) = \underset{Z \in L^2(\Omega,\mathcal{G},\mathbb{P})}{\operatorname{argmin}} \mathbb{E}\left[\alpha(X-Z)_+^2 + (1-\alpha)(X-Z)_-^2\right],$$

where $\mathcal{G} \subseteq \mathcal{F}$ is a σ-algebra on $(\Omega, \mathcal{F})$. We show in Theorems 2.1 and 2.2 that the definition is well-posed since the minimiser is $\mathbb{P}$-a.s. unique and can be characterised by a conditional first order condition. Moreover, it turns out that, for bounded random variables, conditional expectiles belong to the class of conditional shortfall risk measures introduced in Weber (2006). Conditional expectiles are conditional coherent risk measures and satisfy a number of properties that we collect in Proposition 1 and Theorem 2.3.

We then consider the dynamic risk measure $(e_\alpha, e_\alpha^{\mathcal{G}})$ and study its time consistency properties. Recall that a dynamic risk measure $(\rho, \rho^{\mathcal{G}})$ is said to be *sequentially consistent* (see Roorda and Schumacher, 2007) if it satisfies both

$$\rho^{\mathcal{G}}(X) \leq 0 \, \mathbb{P}\text{-a.s.} \Rightarrow \rho_0(X) \leq 0 \text{ (acceptance consistency) and}$$
$$\rho^{\mathcal{G}}(X) \geq 0 \, \mathbb{P}\text{-a.s.} \Rightarrow \rho_0(X) \geq 0 \text{ (rejection consistency),}$$

while it is said to have the *supermartingale property* (see e.g. Detlefsen and Scandolo, 2005), if

$$\rho_0(X) \geq \mathbb{E}[\rho^{\mathcal{G}}(X)].$$

These properties are extremely relevant in a multiperiod setting, when the risk assessment of X might depend on new information becoming available at intermediate times. Sequential consistency guarantees that no loss will be accepted (rejected) at time 0, if it will be accepted (rejected) for sure in the future; the supermartingale property postulates that the riskiness of X should decrease on average when new information is available, see e.g. Detlefsen and Scandolo (2005). It is straightforward to see that dynamic expectiles are sequentially consistent (see Theorem 2.3, f)). The supermartingale property does not hold in general, as it is shown in Example 2. However, in Corollary 3.1 we are able to provide two sufficient conditions that are related to mixture convexity properties of unconditional expectiles.

2 Conditional expectiles

Let X be a real valued random variable defined on a nonatomic probability space $(\Omega, \mathcal{F}, \mathbb{P})$. In the following we adopt the standard actuarial notation where a positive (negative) value of X corresponds to a financial loss (profit). All equalities and inequalities between random variables are meant to hold $\mathbb{P}$-a.s.. For the sake of notational simplicity, $L^p(\Omega, \mathcal{F}, \mathbb{P})$ will be denoted with $L^p(\mathcal{F})$, $p \in [0, +\infty]$. For $X \in L^2(\mathcal{F})$ and $\alpha \in (0,1)$, the expectile $e_\alpha(X)$ is defined as

$$e_\alpha(X) = \operatorname*{argmin}_{x \in \mathbb{R}} \mathbb{E}\left[\alpha(X-x)_+^2 + (1-\alpha)(X-x)_-^2\right], \tag{1}$$

where $x_+ = \max(x, 0)$, $x_- = \max(-x, 0)$. Expectiles can also be expressed as the unique solution of the first order condition

$$\mathbb{E}\left[\ell_\alpha(X-x)\right] = 0, \tag{2}$$

where

$$\ell_\alpha(x) = \alpha x_+ - (1-\alpha)x_-. \tag{3}$$

Indeed, equation (2) is a better definition of e_α since it is valid for each $X \in L^1(\mathcal{F})$. Our first definition of conditional expectiles is a natural extension of (1).

Definition 1. *Let $X \in L^2(\mathcal{F})$, $\mathcal{G} \subseteq \mathcal{F}$ be a σ-algebra and $\alpha \in (0,1)$. We define the conditional expectile $e_\alpha^{\mathcal{G}}(X)$ as:*

$$e_\alpha^{\mathcal{G}}(X) = \operatorname*{argmin}_{Z \in L^2(\mathcal{G})} \mathbb{E}\left[\alpha(X-Z)_+^2 + (1-\alpha)(X-Z)_-^2\right]. \tag{4}$$

Clearly, if $\mathcal{G} = \{\emptyset, \Omega\}$ then $e_\alpha^{\mathcal{G}}(X) = e_\alpha(X)$, while if $\mathcal{G} = \mathcal{F}$ then $e_\alpha^{\mathcal{G}}(X) = X$. If $\alpha = 1/2$ then $e_\alpha^{\mathcal{G}}(X) = \mathbb{E}[X \mid \mathcal{G}]$. In order to show that Definition 1 is well posed, in the following theorem we prove that problem (4) has always a $\mathbb{P}$-a.s. unique solution, characterised by means of a conditional first order condition.

Theorem 2.1. *Let $X \in L^2(\mathcal{F})$, $\mathcal{G} \subseteq \mathcal{F}$ and $\alpha \in (0,1)$. The minimisation problem*

$$\min_{Z \in L^2(\mathcal{G})} \mathbb{E}\left[\alpha(X-Z)_+^2 + (1-\alpha)(X-Z)_-^2\right] \tag{5}$$

admits a $\mathbb{P}$-a.s. unique solution Z^* that satisfies

$$\mathbb{E}\left[\ell_\alpha(X-Z^*) \mid \mathcal{G}\right] = 0. \tag{6}$$

Proof. To prove existence and uniqueness of the minimiser, we show that the objective function is strictly convex, continuous and coercive; the thesis then follows from standard results in convex optimisation (see e.g. Kurdila and Zabarankin, 2005). Let

$$h_X(Z) := \alpha\mathbb{E}[(X-Z)_+^2] + (1-\alpha)\mathbb{E}[(X-Z)_-^2].$$

Since $g(x,z) := \alpha(x-z)_+^2 + (1-\alpha)(x-z)_-^2$ is strictly convex in z for each $x \in \mathbb{R}$, it follows that h_X is strictly convex in Z. If $Z_n \xrightarrow{L^2} Z$, then $(X-Z_n)_+ \xrightarrow{L^2} (X-Z)_+$ and $(X-Z_n)_- \xrightarrow{L^2} (X-Z)_-$, from which continuity of h_X in L^2 follows. Since $(x-y)_+ \geq x_+ - y_+$ and $(x-y)_- \geq x_- - y_-$, it holds that

$$\begin{aligned} h_X(Z) &\geq \alpha\mathbb{E}[(X_+ - Z_+)^2] + (1-\alpha)\mathbb{E}[(X_- - Z_-)^2] \\ &= \alpha\|(X_+ - Z_+)\|_2^2 + (1-\alpha)\|(X_- - Z_-)\|_2^2 \\ &\geq \alpha\left(\|X_+\|_2 - \|Z_+\|_2\right)^2 + (1-\alpha)\left(\|X_-\|_2 - \|Z_-\|_2\right)^2 \to +\infty \end{aligned}$$

when $\|Z\|_2 \to +\infty$, that shows coercivity of $h_X(Z)$. To prove the second part of the thesis, let Z^* be the minimiser of (5). For $A \in \mathcal{G}$, define

$$f_A(t) := \alpha\mathbb{E}[(X-(Z^* + t\mathbb{1}_A))_+^2] + (1-\alpha)\mathbb{E}[(X-(Z^* + t\mathbb{1}_A))_-^2].$$

From the dominated convergence theorem, f_A is differentiable in a neighborhood of 0 and

$$f_A'(t) = 2\alpha\mathbb{E}[\mathbb{1}_A(X-(Z^* + t\mathbb{1}_A))_+] - 2(1-\alpha)\mathbb{E}[\mathbb{1}_A(X-(Z^* + t\mathbb{1}_A))_-].$$

Since Z^* is a minimiser it must hold that $f_A'(0) = 0$, that gives for each $A \in \mathcal{G}$

$$\alpha\mathbb{E}[\mathbb{1}_A\,(X-Z^*)_+] - (1-\alpha)\mathbb{E}[\mathbb{1}_A\,(X-Z^*)_-] = 0,$$

from which the thesis follows. □

The natural domain of conditional expectiles is $L^1(\mathcal{F})$. It is therefore convenient to define them directly from the conditional first order condition (6).

Definition 2. *Let $X \in L^1(\mathcal{F})$, $\mathcal{G} \subseteq \mathcal{F}$ and $\alpha \in (0,1)$. We define the conditional expectile $e_\alpha^{\mathcal{G}}(X)$ as the $\mathbb{P}$-a.s. unique solution of the equation*

$$\mathbb{E}\left[\ell_\alpha(X-Z) \mid \mathcal{G}\right] = 0. \tag{7}$$

From Theorem 2.1, Definition 1 and Definition 2 coincide on $L^2(\mathcal{F})$. In the following theorem we show that Definition 2 is well posed for any $X \in L^1(\mathcal{F})$.

Theorem 2.2. *Let $X \in L^1(\mathcal{F})$, $\mathcal{G} \subseteq \mathcal{F}$ and $\alpha \in (0,1)$. There exists a $\mathbb{P}$-a.s. unique $Z^*_\alpha \in L^1(\mathcal{G})$ such that*

$$\mathbb{E}\left[\ell_\alpha(X - Z^*_\alpha) \mid \mathcal{G}\right] = 0. \tag{8}$$

Moreover,

$$e^{\mathcal{G}}_\alpha(X) = \operatorname{ess\,inf}\left\{Z \in L^1(\mathcal{G}) \mid \mathbb{E}\left[\ell_\alpha(X - Z) \mid \mathcal{G}\right] \le 0\right\} \tag{9}$$

$$= \operatorname{ess\,sup}\{Z \in L^1(\mathcal{G}) \mid \mathbb{E}\left[\ell_\alpha(X - Z) \mid \mathcal{G}\right] \ge 0\}. \tag{10}$$

Proof. Let $F_{\mathcal{G}}$ be a regular conditional distribution of X given $\mathcal{G}$, whose existence is always guaranteed (see e.g. Theorem 6.6.2 in Ash, 1972). Equation (8) can then be written as

$$\int \ell_\alpha(x - z)\mathrm{d}F_{\mathcal{G}}(x,\omega) = 0 \tag{11}$$

(see e.g. Theorem 10.2.5 in Dudley, 2002). Since $F_{\mathcal{G}}(x,\omega)$ is a distribution function for almost every $\omega \in \Omega$, it follows that equation (11) has a unique solution that we denote $Z^*_\alpha(\omega)$. In order to show its measurability, we notice that

$$Z^*_\alpha(\omega) = \operatorname*{argmin}_{z\in\mathbb{R}} S(\omega, z),$$

where

$$S(\omega, z) = \int \left[\alpha(x - z)^2_+ + (1-\alpha)(x - z)^2_-\right] \mathrm{d}F_{\mathcal{G}}(x,\omega).$$

Since $S(\omega, z)$ is continuous in z and measurable in ω, the measurability of Z^*_α follows from the measurable maximum theorem (see e.g. Theorem 18.19 in Aliprantis and Border, 2006). To show that $Z^*_\alpha \in L^1(\mathcal{G})$, assume first that $\alpha \le 1/2$. In this case $Z^*_\alpha \le \mathbb{E}[X \mid \mathcal{G}]$, so

$$\mathbb{E}[(X - Z^*_\alpha)_-] \le \mathbb{E}[(X - \mathbb{E}[X \mid \mathcal{G}])_-] < +\infty$$

and

$$\mathbb{E}[(X - Z^*_\alpha)_+] = \frac{1-\alpha}{\alpha}\,\mathbb{E}[(X - Z^*_\alpha)_-] < +\infty,$$

from which the thesis follows. A similar argument applies for $\alpha \ge 1/2$. To prove the last statement, let $\bar{Z}_\alpha = \operatorname{ess\,inf}\left\{Z \in L^1(\mathcal{G}) \mid \mathbb{E}\left[\ell_\alpha(X - Z) \mid \mathcal{G}\right] \le 0\right\}$. Since $e^{\mathcal{G}}_\alpha(X)$ satisfies (8), it holds that $\bar{Z}_\alpha \le e^{\mathcal{G}}_\alpha(X)$. Assume by contradiction that there exists $A \in \mathcal{G}$ with $\mathbb{P}(A) > 0$ such that $\bar{Z}_\alpha > e^{\mathcal{G}}_\alpha(X)$ on A. From the strict monotonicity of ℓ_α

$$\mathbb{E}\left[\ell_\alpha(X - \bar{Z}_\alpha)\mathbb{1}_A\right] > \mathbb{E}\left[\ell_\alpha(X - e^{\mathcal{G}}_\alpha(X))\mathbb{1}_A\right] = 0,$$

which contradicts the definition of $\bar{Z}_\alpha$. A similar argument proves the second equality. □

We see from (9) that for $X \in L^\infty(\mathcal{F})$ conditional expectiles are a special case of conditional shortfall risk measures introduced by Weber (2006) and defined as

$$\rho^{\mathcal{G}}_\ell(X) := \operatorname{ess\,inf}\{Z \in L^\infty(\mathcal{G}) \mid \mathbb{E}[\ell(X - Z) \mid \mathcal{G}] \le 0\},$$

for a generic $\ell\colon \mathbb{R} \to \mathbb{R}$ non decreasing, non constant and with 0 in the interior of its range.

In the static case, robust shortfall risk measures on L^∞ have been studied in Bellini et al. (2018), as an extension of robust entropic risk measures considered in Laeven and Stadje (2013), while the properties of convex shortfall risk measures on L^1 have been studied in Kaina and Rüschendorf (2009).

The following proposition collects elementary properties of conditional expectiles that can be proved directly via the corresponding properties of conditional shortfalls on L^∞ by noting that ℓ_α is non decreasing with respect to α and convex for $\alpha \geq \frac{1}{2}$. From these properties, it follows that, for $\alpha \geq \frac{1}{2}$, conditional expectiles belong to the class of conditional coherent risk measures as introduced, up to a sign change, by Detlefsen and Scandolo (2005).

Proposition 1. *Let $X \in L^1(\mathcal{F})$ and let $e^{\mathcal{G}}_\alpha \colon L^1(\mathcal{F}) \to L^1(\mathcal{G})$ as in Definition 2. Then:*

a) $\alpha_1 \leq \alpha_2 \Rightarrow e^{\mathcal{G}}_{\alpha_1}(X) \leq e^{\mathcal{G}}_{\alpha_2}(X)$;

b) $X \leq Y \Rightarrow e^{\mathcal{G}}_\alpha(X) \leq e^{\mathcal{G}}_\alpha(Y)$;

c) for any $H \in L^1(\mathcal{G})$, $e^{\mathcal{G}}_\alpha(X+H) = e^{\mathcal{G}}_\alpha(X) + H$;

d) for any non negative $\Lambda \in L^\infty(\mathcal{G})$, $e^{\mathcal{G}}_\alpha(\Lambda X) = \Lambda e^{\mathcal{G}}_\alpha(X)$;

e) for any $X, Y \in L^1(\mathcal{F})$ *, if* $\alpha \geq \frac{1}{2}$

$$e^{\mathcal{G}}_\alpha(X+Y) \leq e^{\mathcal{G}}_\alpha(X) + e^{\mathcal{G}}_\alpha(Y);$$

if $\alpha \leq \frac{1}{2}$,

$$e^{\mathcal{G}}_\alpha(X+Y) \geq e^{\mathcal{G}}_\alpha(X) + e^{\mathcal{G}}_\alpha(Y).$$

Further results on conditional expectiles are gathered in the following Theorem 2.3. In particular, *a*) states that the conditional expectile is the static expectile applied to the conditional distribution of the underlying position – given the information in $\mathcal{G}$. This property is to be interpreted as a conditional law-invariance, as the conditional risk measurement $e^{\mathcal{G}}_\alpha(X)$ solely depends on the conditional distribution of X. Property *b*) instead is a generalised version of Jensen's inequality for the asymmetric conditional expectation. Additional continuity properties are given in *d*) and *e*), while *f*) and *g*) will be used in the next section to establish time consistency properties of dynamic expectiles.

Theorem 2.3. *Let $X \in L^1(\mathcal{F})$ and let $e^{\mathcal{G}}_\alpha \colon L^1(\mathcal{F}) \to L^1(\mathcal{G})$ as in Definition 2. Then:*

a) $e^{\mathcal{G}}_\alpha$ *satisfies*

$$e^{\mathcal{G}}_\alpha(X)(\omega) = e_\alpha(F_{\mathcal{G}}(\cdot, \omega)),$$

where $F_{\mathcal{G}}(\cdot, \omega)$ is a regular conditional distribution of X on $\mathcal{G}$;

b) if $\phi \colon \mathbb{R} \to \mathbb{R}$ is increasing and convex, then

$$\phi(e^{\mathcal{G}}_\alpha(X)) \leq e^{\mathcal{G}}_\alpha(\phi(X));$$

if $\phi \colon \mathbb{R} \to \mathbb{R}$ is convex, then

$$\phi(e^{\mathcal{G}}_\alpha(X)) \leq \max(e^{\mathcal{G}}_\alpha(\phi(X)), e^{\mathcal{G}}_{1-\alpha}(\phi(X)));$$

c) $X \in L^p(\mathcal{F}) \Rightarrow e_\alpha^\mathcal{G}(X) \in L^p(\mathcal{G})$;

d) $X_n \uparrow X$ $\mathbb{P}$*-a.s.* $\Rightarrow e_\alpha^\mathcal{G}(X_n) \uparrow e_\alpha^\mathcal{G}(X)$ $\mathbb{P}$*-a.s.*;

e) *for any* $X, Y \in L^1(\mathcal{F})$,

$$\|e_\alpha^\mathcal{G}(X) - e_\alpha^\mathcal{G}(Y)\|_1 \leq \beta \|X - Y\|_1,$$

where $\beta = \max\{\alpha/(1-\alpha), (1-\alpha)/\alpha\}$;

f) *Let* $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$. *Then* $e_\alpha^{\mathcal{G}_2}(X) \leq 0 \Rightarrow e_\alpha^{\mathcal{G}_1}(X) \leq 0$ *and* $e_\alpha^{\mathcal{G}_2}(X) \geq 0 \Rightarrow e_\alpha^{\mathcal{G}_1}(X) \geq 0$.

g) *Let* $\mathcal{G}_1 \subseteq \mathcal{G}_2 \subseteq \mathcal{F}$. *If* $\alpha \in [1/2, 1)$, *then*

$$e_\alpha^{\mathcal{G}_1}(e_\alpha^{\mathcal{G}_2}(X)) \geq e_\alpha^{\mathcal{G}_1}(X),$$

while if $\alpha \in (0, 1/2]$, *we have*

$$e_\alpha^{\mathcal{G}_1}(e_\alpha^{\mathcal{G}_2}(X)) \leq e_\alpha^{\mathcal{G}_1}(X).$$

Proof. a) Recall that from the definition of regular conditional distribution $F_\mathcal{G}(\cdot, \omega)$ is a distribution function for almost every $\omega \in \Omega$, and from (8) it holds $\mathbb{P}$-a.s. that

$$0 = \mathbb{E}[\ell_\alpha(X - e_\alpha^\mathcal{G}(X)) \mid \mathcal{G}] = \int \ell_\alpha(x - e_\alpha^\mathcal{G}(X)) \mathrm{d}F_\mathcal{G}(x, \omega) = \mathbb{E}[\ell_\alpha(Y - e_\alpha^\mathcal{G}(X))],$$

where Y is a random variable with distribution $F_\mathcal{G}(\cdot, \omega)$.

b) Since ϕ is convex and increasing, $\phi(x) = \sup_n\{a_n x + b_n\}$, with $a_n \geq 0$. Then

$$\phi(e_\alpha^\mathcal{G}(X)) = \sup_n\{a_n e_\alpha^\mathcal{G}(X) + b_n\} = \sup_n\{e_\alpha^\mathcal{G}(a_n X + b_n)\} \leq e_\alpha^\mathcal{G}(\phi(X)).$$

For the general case notice that if $a_n \leq 0$ the preceding argument becomes

$$\phi(e_\alpha^\mathcal{G}(X)) = \sup_n\{a_n e_\alpha^\mathcal{G}(X) + b_n\} = \sup_n\{e_{1-\alpha}^\mathcal{G}(a_n X + b_n)\} \leq e_{1-\alpha}^\mathcal{G}(\phi(X)).$$

c) Follows immediately from b).

d) Since $X_n \uparrow X$ $\mathbb{P}$-a.s., it follows that $e_\alpha^\mathcal{G}(X_n) \leq e_\alpha^\mathcal{G}(X_{n+1})$, so $e_\alpha^\mathcal{G}(X_n) \uparrow Z$ $\mathbb{P}$-a.s., with $Z \leq e_\alpha^\mathcal{G}(X)$. Since

$$|\ell_\alpha(X_n - e_\alpha^\mathcal{G}(X_n))| \leq (X - e_\alpha^\mathcal{G}(X_1))_+ + (X_1 - e_\alpha^\mathcal{G}(X))_-,$$

from the dominated convergence theorem it follows that

$$0 = \mathbb{E}[\ell_\alpha(X_n - e_\alpha^\mathcal{G}(X_n)) \mid \mathcal{G}] \to \mathbb{E}[\ell_\alpha(X - Z) \mid \mathcal{G}] = 0,$$

that shows $Z = e_\alpha^\mathcal{G}(X)$.

e) As in the scalar case, the conditional first order condition can be written as

$$e_\alpha^{\mathcal{G}}(X) = \mathbb{E}[X \mid \mathcal{G}] + \frac{2\alpha-1}{1-\alpha}\,\mathbb{E}\left[\left(X - e_\alpha^{\mathcal{G}}(X)\right)_+ \mid \mathcal{G}\right].$$

It follows that

$$\left\|e_\alpha^{\mathcal{G}}(Y) - e_\alpha^{\mathcal{G}}(X)\right\|_1 \le \|Y-X\|_1 + \frac{2\alpha-1}{1-\alpha}\left\|\left(Y - e_\alpha^{\mathcal{G}}(Y)\right)_+ - \left(X - e_\alpha^{\mathcal{G}}(X)\right)_+\right\|_1.$$

Assume now that $\alpha \ge 1/2$ and let $A = \mathbb{1}_{\{e_\alpha^{\mathcal{G}}(Y) \ge e_\alpha^{\mathcal{G}}(X)\}}$. Clearly $A \in \mathcal{G}$ and

$$\begin{aligned}&\left\|\left(Y - e_\alpha^{\mathcal{G}}(Y)\right)_+ - \left(X - e_\alpha^{\mathcal{G}}(X)\right)_+\right\|_1 = \mathbb{E}\left[\left|\left(Y - e_\alpha^{\mathcal{G}}(Y)\right)_+ - \left(X - e_\alpha^{\mathcal{G}}(X)\right)_+\right| \mathbb{1}_A\right]\\ &+ \left[\left|\left(Y - e_\alpha^{\mathcal{G}}(Y)\right)_+ - \left(X - e_\alpha^{\mathcal{G}}(X)\right)_+\right| \mathbb{1}_{A^c}\right] \le \mathbb{E}\left[(Y-X)_+\mathbb{1}_A + (X-Y)_+\mathbb{1}_{A^c}\right]\\ &\le \|Y-X\|_1.\end{aligned}$$

Summing up,

$$\left\|e_\alpha^{\mathcal{G}}(Y) - e_\alpha^{\mathcal{G}}(X)\right\|_1 \le \frac{\alpha}{1-\alpha}\|Y-X\|_1.$$

A similar argument holds for $\alpha \le 1/2$.

f) From (9) it follows that

$$e_\alpha^{\mathcal{G}_2}(X) \le 0 \Rightarrow \mathbb{E}[\ell_\alpha(X) \mid \mathcal{G}_2] \le 0 \Rightarrow \mathbb{E}[\ell_\alpha(X) \mid \mathcal{G}_1] \le 0 \Rightarrow e_\alpha^{\mathcal{G}_1}(X) \le 0.$$

The other inequality follows similarly from (10).

g) For $\alpha \in [1/2, 1)$, $e_\alpha^{\mathcal{G}_1}$ is a coherent (and in particular subadditive) risk measure, therefore

$$e_\alpha^{\mathcal{G}_1}(X) = e_\alpha^{\mathcal{G}_1}(X - e_\alpha^{\mathcal{G}_2}(X) + e_\alpha^{\mathcal{G}_2}(X)) \le e_\alpha^{\mathcal{G}_1}(X - e_\alpha^{\mathcal{G}_2}(X)) + e_\alpha^{\mathcal{G}_1}(e_\alpha^{\mathcal{G}_2}(X));$$

clearly $e_\alpha^{\mathcal{G}_2}(X - e_\alpha^{\mathcal{G}_2}(X)) = 0$, thus it follows from $f)$ that $e_\alpha^{\mathcal{G}_1}(X - e_\alpha^{\mathcal{G}_2}(X)) = 0$, and we obtain

$$e_\alpha^{\mathcal{G}_1}(X) \le e_\alpha^{\mathcal{G}_1}(e_\alpha^{\mathcal{G}_2}(X)).$$

The second part of the statement follows noting that for $\alpha \in (0, 1/2]$ expectiles are superadditive.

□

Being conditional coherent risk measures, conditional expectiles can be written as suprema of conditional expectations over a set of generalised scenarios. Their dual representation may be derived as a special case of the dual representation of conditional shortfall risk measures that can be found in Föllmer and Schied (2011), although here the domain is L^1 instead of L^∞. We provide a short and direct proof, that closely resembles the static case considered in Bellini et al. (2014). In order to formulate it, we denote with $\operatorname{ess\,sup}[\varphi \mid \mathcal{G}]$ the conditional supremum of a random variable $\varphi \in L^\infty(\mathcal{F})$ with respect to a σ-algebra $\mathcal{G} \subseteq \mathcal{F}$, that is the minimal random variable $Z \in L^\infty(\mathcal{G})$ such that $\operatorname{ess\,sup}(\varphi\mathbb{1}_A) \le \operatorname{ess\,sup}(Z\mathbb{1}_A)$ for each $A \in \mathcal{G}$ (see Barron et al., 2003 for further properties of conditional essential suprema).

Theorem 2.4. *Let $X \in L^1(\mathcal{F})$. If $\alpha \in [1/2, 1)$, then*

$$e_\alpha^{\mathcal{G}}(X) = \operatorname*{ess\,max}_{\varphi \in \mathcal{M}_\alpha^{\mathcal{G}}} \mathbb{E}[\varphi X \mid \mathcal{G}], \tag{12}$$

while if $\alpha \in (0, 1/2]$, we have

$$e_\alpha^{\mathcal{G}}(X) = \operatorname*{ess\,min}_{\varphi \in \mathcal{M}_\alpha^{\mathcal{G}}} \mathbb{E}[\varphi X \mid \mathcal{G}] \tag{13}$$

where

$$\mathcal{M}_\alpha^{\mathcal{G}} = \left\{ \varphi \in L^\infty(\mathcal{F}) \;\middle|\; \varphi > 0, \mathbb{E}[\varphi \mid \mathcal{G}] = 1, \frac{\operatorname{ess\,sup}[\varphi \mid \mathcal{G}]}{\operatorname{ess\,inf}[\varphi \mid \mathcal{G}]} \le \beta \right\}, \tag{14}$$

with $\beta = \max\left(\frac{1-\alpha}{\alpha}, \frac{\alpha}{1-\alpha}\right)$. The optimum in (12) and (13) is achieved by

$$\bar{\varphi} = \frac{\alpha \mathbb{1}_{\{X > e_\alpha^{\mathcal{G}}(X)\}} + (1-\alpha)\mathbb{1}_{\{X \le e_\alpha^{\mathcal{G}}(X)\}}}{\mathbb{E}[\alpha \mathbb{1}_{\{X > e_\alpha^{\mathcal{G}}(X)\}} + (1-\alpha)\mathbb{1}_{\{X \le e_\alpha^{\mathcal{G}}(X)\}} \mid \mathcal{G}]}.$$

Proof. First of all we show that for each $\alpha \in (0,1)$ it holds $\mathbb{E}[\bar{\varphi} X \mid \mathcal{G}] = e_\alpha^{\mathcal{G}}(X)$. To this aim, we compute

$$\begin{aligned}
\mathbb{E}[\bar{\varphi} X \mid \mathcal{G}] &= \frac{\mathbb{E}[\alpha X \mathbb{1}_{\{X > e_\alpha^{\mathcal{G}}(X)\}} + (1-\alpha) X \mathbb{1}_{\{X \le e_\alpha^{\mathcal{G}}(X)\}} \mid \mathcal{G}]}{\mathbb{E}[\alpha \mathbb{1}_{\{X > e_\alpha^{\mathcal{G}}(X)\}} + (1-\alpha)\mathbb{1}_{\{X \le e_\alpha^{\mathcal{G}}(X)\}} \mid \mathcal{G}]} \\
&= e_\alpha^{\mathcal{G}}(X) + \frac{\alpha \mathbb{E}[(X - e_\alpha^{\mathcal{G}}(X))_+ \mid \mathcal{G}] + (1-\alpha)\mathbb{E}[(X - e_\alpha^{\mathcal{G}}(X))_- \mid \mathcal{G}]}{\mathbb{E}[\alpha \mathbb{1}_{\{X > e_\alpha^{\mathcal{G}}(X)\}} + (1-\alpha)\mathbb{1}_{\{X \le e_\alpha^{\mathcal{G}}(X)\}} \mid \mathcal{G}]} \\
&= e_\alpha^{\mathcal{G}}(X),
\end{aligned}$$

where the last equality follows from (7). Let now $\alpha \ge 1/2$. Since $\bar{\varphi} \in \mathcal{M}_\alpha^{\mathcal{G}}$, clearly

$$e_\alpha^{\mathcal{G}}(X) \le \operatorname*{ess\,max}_{\varphi \in \mathcal{M}_\alpha^{\mathcal{G}}} \mathbb{E}[\varphi X \mid \mathcal{G}].$$

To show the opposite inequality, we prove that for each $\varphi \in \mathcal{M}_\alpha^{\mathcal{G}}$

$$\mathbb{E}[\varphi X \mid \mathcal{G}] \le e_\alpha^{\mathcal{G}}(X).$$

Assume w.l.o.g. that $e_\alpha^{\mathcal{G}}(X) = 0$ (otherwise replace X by $X - e_\alpha^{\mathcal{G}}(X)$). Then

$$\begin{aligned}
\mathbb{E}[\varphi X \mid \mathcal{G}] = \mathbb{E}[\varphi(X_+ - X_-) \mid \mathcal{G}] &\le \operatorname{ess\,sup}[\varphi \mid \mathcal{G}]\, \mathbb{E}[X_+ \mid \mathcal{G}] - \operatorname{ess\,inf}[\varphi \mid \mathcal{G}]\, \mathbb{E}[X_- \mid \mathcal{G}] \\
&\le \operatorname{ess\,inf}[\varphi \mid \mathcal{G}]\, (\beta \mathbb{E}[X_+ \mid \mathcal{G}] - \mathbb{E}[X_- \mid \mathcal{G}]) = 0,
\end{aligned}$$

where the last equality follows from (14) and (7). A similar argument applies for $\alpha \le 1/2$. □

Notice that if $\alpha = \frac{1}{2}$ then $\beta = 1$, hence from (14) it follows that $\mathcal{M}_{1/2}^{\mathcal{G}} = \{\varphi = 1, \mathbb{P}\text{-a.s.}\}$, and so we find again that $e_{1/2}^{\mathcal{G}}(X) = \mathbb{E}[X \mid \mathcal{G}]$. If $\mathcal{G} = \{\emptyset, \Omega\}$, then $\mathcal{M}_\alpha^{\mathcal{G}}$ coincides with

$$\mathcal{M}_\alpha = \left\{ \varphi \in L^\infty(\mathcal{F}) \;\middle|\; \varphi > 0, \mathbb{E}[\varphi] = 1, \frac{\operatorname{ess\,sup} \varphi}{\operatorname{ess\,inf} \varphi} \le \beta \right\},$$

that is the dual set of unconditional expectiles, see for instance Bellini et al. (2014).

Remark 1. *It is possible to check that the dual sets* $\mathcal{M}_\alpha$ *and* $\mathcal{M}_\alpha^{\mathcal{G}}$ *are convex and maximal. Further, in general* $\mathcal{M}_\alpha^{\mathcal{G}} \not\subseteq \mathcal{M}_\alpha$*, since* $\operatorname{ess\,sup}[\varphi \mid \mathcal{G}] \leq \operatorname{ess\,sup} \varphi$ *and* $\operatorname{ess\,inf}[\varphi \mid \mathcal{G}] \geq \operatorname{ess\,inf} \varphi$*, so*

$$\frac{\operatorname{ess\,sup}[\varphi \mid \mathcal{G}]}{\operatorname{ess\,inf}[\varphi \mid \mathcal{G}]} \leq \frac{\operatorname{ess\,sup} \varphi}{\operatorname{ess\,inf} \varphi}.$$

As it was pointed out by an anonymous referee, it is possible to conclude that in general it does not hold that $e_\alpha(X) \geq \mathbb{E}[e_\alpha^{\mathcal{G}}(X)]$*, since* $\mathbb{E}[e_\alpha^{\mathcal{G}}(X)]$ *is a coherent risk measure with dual set* $\mathcal{M}_\alpha^{\mathcal{G}}$*. The same conclusion can be drawn from mixture concavity properties of unconditional expectiles; see Section 3.*

3 Time consistency properties of dynamic expectiles

We consider dynamic risk measures $(\rho_0, \rho^{\mathcal{G}})$ with $\rho_0 \colon L^1(\mathcal{F}) \to \mathbb{R}$ and $\rho^{\mathcal{G}} \colon L^1(\mathcal{F}) \to L^1(\mathcal{G})$. In this context the σ-algebra $\mathcal{G}$ represents the information available at an intermediate time $t \in (0, T)$. Time consistency properties relate future conditional risk assessment at time t given by $\rho^{\mathcal{G}}$ with today's risk assessment given by ρ_0. Recall the following definitions (see for instance Acciaio and Penner, 2011).

Definition 3. *A dynamic risk measure* $(\rho_0, \rho^{\mathcal{G}})$ *is said:*

- *dynamic consistent, if for all* $X, Y \in L^1(\mathcal{F})$

$$\rho^{\mathcal{G}}(X) = \rho^{\mathcal{G}}(Y) \Rightarrow \rho_0(X) = \rho_0(Y),$$

- *sequentially consistent, if for all* $X \in L^1(\mathcal{F})$
  - *a)* $\rho^{\mathcal{G}}(X) \leq 0 \Rightarrow \rho_0(X) \leq 0$ *(acceptance consistent), and*
  - *b)* $\rho^{\mathcal{G}}(X) \geq 0 \Rightarrow \rho_0(X) \geq 0$ *(rejection consistent).*
- *to have the supermartingale property, if for all* $X \in L^1(\mathcal{F})$

$$\rho_0(X) \geq \mathbb{E}[\rho^{\mathcal{G}}(X)]. \tag{15}$$

Dynamic consistency implies sequential consistency, while the supermartingale property implies rejection consistency but does not in general imply acceptance consistency. Kupper and Schachermayer (2009) proved that in the monetary and law invariant case dynamic consistency is satisfied only by the entropic risk measure, defined by the one parameter family

$$\rho_\gamma^{\mathcal{G}}(X) = \begin{cases} \frac{1}{\gamma} \ln \mathbb{E}\left[\exp(\gamma X) \mid \mathcal{G}\right], & \text{if } \gamma \in (-\infty, 0) \cup (0, +\infty) \\ \mathbb{E}\left[X \mid \mathcal{G}\right], & \text{if } \gamma = 0 \\ \operatorname{ess\,sup}[X \mid \mathcal{G}], & \text{if } \gamma = +\infty \end{cases}$$

(Föllmer and Schied, 2011). It follows that dynamic expectiles satisfy dynamic consistency only when they coincide with conditional expectations, that is when $\alpha = 1/2$. Another way to see the same result is to recall that Detlefsen and Scandolo (2005) proved that dynamic consistency is equivalent to the tower property

$$\rho_0(\rho^{\mathcal{G}}(X)) = \rho_0(X),$$

that is not satisfied by conditional expectiles with $\alpha \neq 1/2$, as it can be seen in elementary examples such as the following one.

Example 1. Let $\Omega = \{\omega_1, \dots, \omega_4\}$, $\mathbb{P}(\omega_i) = 1/4$, $\mathcal{G} = \{\Omega, \emptyset, A, A^c\}$ with $A = \{\omega_1, \omega_2\}$. Let $X(\omega_i) = i$ and $\alpha = 1/3$. Then

$$e^{\mathcal{G}}_{1/3}(X)(\omega) = \begin{cases} 4/3 \text{ if } \omega \in A \\ 10/3 \text{ if } \omega \in A^c, \end{cases}$$

so $e_{1/3}(e^{\mathcal{G}}_{1/3}(X)) = 2$, while $e_{1/3}(X) = 13/6$.

On the contrary, from Theorem 2.3, f) it follows immediately that dynamic expectiles are sequentially consistent, coherently with Weber (2006) that showed that (under some weak continuity assumptions and on L^∞) the class of sequentially consistent risk measures coincides with the class of dynamic shortfall risk measures. Thus the only question that remains open is whether dynamic expectiles have the supermartingale property.

Since conditional expectiles are conditionally law invariant, a necessary condition for the supermartingale property of dynamic expectiles is mixture concavity of unconditional expectiles (see e.g. Pflug and Römisch, 2007). We report the argument for completeness. Let $\mathcal{M}_1 := \mathcal{M}_1(\mathbb{R})$ be the space of distributions on $\mathbb{R}$ with finite first moment and recall that a law invariant risk measure on L^1 can be alternatively defined as a functional on $\mathcal{M}_1$, where, with slight abuse of notation, we denote $\rho(F) = \rho(X)$ for any random variable $X \sim F \in \mathcal{M}_1$. Since $(\Omega, \mathcal{F}, \mathbb{P})$ is nonatomic, for each $\lambda \in (0, 1)$ and for each $F, G \in \mathcal{M}_1$ there exist $\mathcal{A} \in \mathcal{F}$ with $\mathbb{P}(A) = \lambda$ and $X, Y \in L^1(\mathcal{F})$ with $X \sim F$ and $Y \sim G$ such that $\mathbb{1}_A, X$ and Y are pairwise independent. Letting $Z = \mathbb{1}_A X + \mathbb{1}_{A^C} Y$ and $\mathcal{G} = \{A, A^C, \emptyset, \Omega\}$, clearly $Z \sim \lambda F + (1 - \lambda)G$, and from conditional law invariance we get

$$e^{\mathcal{G}}_\alpha(Z)(\omega) = \begin{cases} e_\alpha(F) \text{ if } \omega \in A \\ e_\alpha(G) \text{ if } \omega \in A^c \end{cases}.$$

Hence the supermartingale property (15) implies that for each $F, G \in \mathcal{M}_1$ and for each $\lambda \in (0, 1)$,

$$e_\alpha(\lambda F + (1 - \lambda)G) \geq \lambda e_\alpha(F) + (1 - \lambda)e_\alpha(G), \tag{16}$$

that corresponds to the concavity of the function

$$z(\lambda) := e_\alpha(\lambda F + (1 - \lambda)G). \tag{17}$$

It turns out that (17) is in general not satisfied, as it can be seen in the following example.

Example 2. Let $F = \frac{1}{2}\,\delta_{-1} + \frac{1}{2}\,\delta_1$ and $G = \frac{2}{3}\,\delta_0 + \frac{1}{3}\,\delta_5$. Then $e_{1/3}(F) = -1/3$, $e_{1/3}(G) = 1$,

$$z(\lambda) = \begin{cases} \frac{10-13\lambda}{10-\lambda} \text{ if } \lambda \in [0, 10/13) \\ \frac{10-13\lambda}{6+3\lambda} \text{ if } \lambda \in [10/13, 1] \end{cases}$$

$$z''(\lambda) = \begin{cases} \frac{240}{(\lambda-10)^3} < 0 \text{ if } \lambda \in (0, 10/13) \\ \frac{24}{(\lambda+2)^3} > 0 \text{ if } \lambda \in (10/13, 1), \end{cases}$$

so $z(\lambda)$ changes its concavity in $\lambda = 10/13$.

Remark 2. *Alternatively, the lack of mixture concavity of expectiles may be derived from Corollary 6 in Acciaio and Svindland (2013), since the set of generalised scenarios $\mathcal{M}_\alpha$ in their dual representation is not convex with respect to mixtures (although it is convex with respect to sums, as we noted in Remark 1).*

However, for fixed F and G, we can prove an equivalent characterisation and two simple sufficient conditions for (16) (or for the opposite inequality).

Theorem 3.1. *Let $F, G \in \mathcal{M}_1$ with $e_\alpha(F) < e_\alpha(G)$ and let $\bar{z} := \lambda e_\alpha(F) + (1-\lambda)e_\alpha(G)$, with $\lambda \in (0,1)$. If $\alpha < 1/2$, then*

$$e_\alpha(\lambda F + (1-\lambda)G) \leq \lambda e_\alpha(F) + (1-\lambda)e_\alpha(G) \tag{18}$$

if and only if

$$\lambda \int_{e_\alpha(F)}^{\bar{z}} F(t)\mathrm{d}t - (1-\lambda) \int_{\bar{z}}^{e_\alpha(G)} G(t)\mathrm{d}t \geq 0, \tag{19}$$

while

$$e_\alpha(\lambda F + (1-\lambda)G) \geq \lambda e_\alpha(F) + (1-\lambda)e_\alpha(G) \tag{20}$$

if and only if

$$\lambda \int_{e_\alpha(F)}^{\bar{z}} F(t)\mathrm{d}t - (1-\lambda) \int_{\bar{z}}^{e_\alpha(G)} G(t)\mathrm{d}t \leq 0. \tag{21}$$

If $\alpha > 1/2$, then (18) $\iff$ (21) and (20) $\iff$ (19).

Proof. Let $x := e_\alpha(F)$ and $y := e_\alpha(G)$, so that $\bar{z} = \lambda x + (1-\lambda)y$. Let $H_\lambda := \lambda F + (1-\lambda)G$ and $\tilde{z} := e_\alpha(H_\lambda)$. Define also

$$I(z) := \alpha \int_{-\infty}^{+\infty} (t-z)_+ \mathrm{d}H_\lambda(t) - (1-\alpha) \int_{-\infty}^{+\infty} (t-z)_- \mathrm{d}H_\lambda(t).$$

Since $I(\tilde{z}) = 0$ and $I(z)$ is strictly decreasing, it follows that

$$\tilde{z} \leq \bar{z} \iff I(\bar{z}) \leq 0.$$

Integrating by parts, we compute

$$I(z) = \alpha \int_z^{+\infty} \overline{H}_\lambda(t)\mathrm{d}t - (1-\alpha) \int_{-\infty}^z H_\lambda(t)\mathrm{d}t = \alpha\lambda \int_z^{+\infty} \overline{F}(t)\mathrm{d}t +$$
$$+ \alpha(1-\lambda) \int_z^{+\infty} \overline{G}(t)\mathrm{d}t - (1-\alpha)\lambda \int_{-\infty}^z F(t)\mathrm{d}t - (1-\alpha)(1-\lambda) \int_{-\infty}^z G(t)\mathrm{d}t,$$

where $\overline{H}_\lambda = 1 - H_\lambda$, $\overline{F} = 1 - F$, $\overline{G} = 1 - G$. For $z \in (x, y)$, we break the integrals getting

$$I(z) = \alpha\lambda \int_z^y \overline{F}(t)\mathrm{d}t + \alpha\lambda \int_y^{+\infty} \overline{F}(t)\mathrm{d}t + \alpha(1-\lambda) \int_z^y \overline{G}(t)\mathrm{d}t + \alpha(1-\lambda) \int_y^{+\infty} \overline{G}(t)\mathrm{d}t +$$
$$- (1-\alpha)\lambda \int_{-\infty}^x F(t)\mathrm{d}t - (1-\alpha)\lambda \int_x^z F(t)\mathrm{d}t - (1-\alpha)(1-\lambda) \int_{-\infty}^x G(t)\mathrm{d}t +$$
$$- (1-\alpha)(1-\lambda) \int_x^z G(t)\mathrm{d}t. \tag{22}$$

Since $x = e_\alpha(F)$ and $y = e_\alpha(G)$, it holds that

$$\alpha \int_y^{+\infty} \overline{G}(t)\mathrm{d}t = (1-\alpha) \int_{-\infty}^y G(t)\mathrm{d}t \quad \text{and} \quad (1-\alpha) \int_{-\infty}^x F(t)\mathrm{d}t = \alpha \int_x^{+\infty} \overline{F}(t)\mathrm{d}t,$$

so substituting the fourth and fifth integral in (22) we get

$$I(z) = \alpha\lambda \int_z^y \overline{F}(t)\mathrm{d}t + \alpha\lambda \int_y^{+\infty} \overline{F}(t)\mathrm{d}t + \alpha(1-\lambda) \int_z^y \overline{G}(t)\mathrm{d}t + (1-\alpha)(1-\lambda) \int_{-\infty}^y G(t)\mathrm{d}t +$$
$$- \alpha\lambda \int_x^{+\infty} \overline{F}(t)\mathrm{d}t - (1-\alpha)\lambda \int_x^z F(t)\mathrm{d}t - (1-\alpha)(1-\lambda) \int_{-\infty}^x G(t)\mathrm{d}t +$$
$$- (1-\alpha)(1-\lambda) \int_x^z G(t)\mathrm{d}t = -\alpha\lambda \int_x^z \overline{F}(t)\mathrm{d}t + \alpha(1-\lambda) \int_z^y \overline{G}(t)\mathrm{d}t +$$
$$+ (1-\alpha)(1-\lambda) \int_z^y G(t)\mathrm{d}t - (1-\alpha)\lambda \int_x^z F(t)\mathrm{d}t.$$

Recalling that $\bar{z} = \lambda x + (1-\lambda)y$, we compute

$$I(\bar{z}) = -\alpha\lambda \int_x^{\bar{z}} \overline{F}(t)\mathrm{d}t + \alpha(1-\lambda) \int_{\bar{z}}^y \overline{G}(t)\mathrm{d}t + (1-\alpha)(1-\lambda) \int_{\bar{z}}^y G(t)\mathrm{d}t +$$
$$- (1-\alpha)\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t = -\alpha\lambda \int_x^{\bar{z}} (1-F(t))\mathrm{d}t + \alpha(1-\lambda) \int_{\bar{z}}^y (1-G(t))\mathrm{d}t +$$
$$+ (1-\alpha)(1-\lambda) \int_{\bar{z}}^y G(t)\mathrm{d}t - (1-\alpha)\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t = -\alpha\lambda(\bar{z}-x) +$$
$$+ \alpha\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t + \alpha(1-\lambda)(y-\bar{z}) - \alpha(1-\lambda) \int_{\bar{z}}^y G(t)\mathrm{d}t + (1-\alpha)(1-\lambda) \int_{\bar{z}}^y G(t)\mathrm{d}t +$$
$$- (1-\alpha)\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t = (2\alpha-1) \left(\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t - (1-\lambda) \int_{\bar{z}}^y G(t)\mathrm{d}t \right),$$

since $-\alpha\lambda(\bar{z}-x) + \alpha(1-\lambda)(y-\bar{z}) = 0$. It follows that, in the case $\alpha < 1/2$,

$$I(\bar{z}) \leq 0 \iff \lambda \int_x^{\bar{z}} F(t)\mathrm{d}t - (1-\lambda) \int_{\bar{z}}^y G(t)\mathrm{d}t \geq 0,$$

which gives the thesis. The case $\alpha > 1/2$ follows immediately. □

Inequalities (19) and (21) are not easy to check directly. We provide two simple sufficient conditions in the following theorem.

Theorem 3.2. *Let $F, G \in \mathcal{M}_1$ with $e_\alpha(F) < e_\alpha(G)$. Let $\alpha < 1/2$.*

a) If

$$F(e_\alpha(F)) \geq G(e_\alpha(G)), \tag{23}$$

then for each $\lambda \in (0,1)$ it holds that

$$e_\alpha(\lambda F + (1-\lambda)G) \leq \lambda e_\alpha(F) + (1-\lambda)e_\alpha(G). \tag{24}$$

b) If

$$F(z) \leq G(z), \tag{25}$$

for each $z \in (e_\alpha(F), e_\alpha(G))$, then for each $\lambda \in (0,1)$ it holds that

$$e_\alpha(\lambda F + (1-\lambda)G) \geq \lambda e_\alpha(F) + (1-\lambda)e_\alpha(G). \tag{26}$$

If $\alpha > 1/2$ then (23) $\Rightarrow$ (26) and (25) $\Rightarrow$ (24).

Proof. Since F and G are non decreasing, we have the straightforward inequality

$$\begin{aligned}&\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t - (1-\lambda)\int_{\bar{z}}^y G(t)\mathrm{d}t \geq \lambda F(x)(\bar{z}-x) - (1-\lambda)G(y)(y-\bar{z})\\ &= \lambda(1-\lambda)(y-x)[F(x) - G(y)].\end{aligned}$$

If (23) holds, then (19) holds, from which a) follows again from Theorem 3.1. Similarly, from the specular inequality

$$\begin{aligned}&\lambda \int_x^{\bar{z}} F(t)\mathrm{d}t - (1-\lambda)\int_{\bar{z}}^y G(t)\mathrm{d}t \leq \lambda F(\bar{z})(\bar{z}-x) - (1-\lambda)G(\bar{z})(y-\bar{z})\\ &= \lambda(1-\lambda)(y-x)[F(\bar{z}) - G(\bar{z})]\end{aligned}$$

we get that (25) $\Rightarrow$ (21), from which b) follows, again from Theorem 3.1. □

We illustrate the previous results in the following examples.

Example 3. Let $F = \delta_x$ and $G = \delta_y$, with $x < y$. Since $e_\alpha(F) = x$ and $e_\alpha(G) = y$, condition (23) is satisfied. By a direct computation,

$$\begin{aligned}&e_\alpha(\lambda F + (1-\lambda)G) - \lambda e_\alpha(F) - (1-\lambda)e_\alpha(G) = \frac{\lambda(1-\alpha)x + \alpha(1-\lambda)y}{\lambda(1-\alpha) + \alpha(1-\lambda)} - \lambda x - (1-\lambda)y\\ &= (2\alpha - 1)\frac{\lambda(1-\lambda)(y-x)}{\lambda(1-\alpha) + \alpha(1-\lambda)},\end{aligned}$$

that has the same sign of $2\alpha - 1$, in accordance with Theorem 3.2.

Example 4. Let $F = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ and $G = \frac{2}{3}\delta_{1/3} + \frac{1}{3}\delta_1$. Since $e_{1/3}(F) = 1/3$, $e_{1/3}(G) = 7/15$, and for $z \in (1/3, 7/15)$ it holds that $F(z) \leq G(z)$, condition (25) is satisfied. hence the inequality

$$e_\alpha(\lambda F + (1-\lambda)G) = \frac{14 - 5\lambda}{30 - 3\lambda} \geq \frac{\lambda}{3} + (1-\lambda)\frac{7}{15} = \lambda e_\alpha(F) + (1-\lambda)e_\alpha(G),$$

is satisfied for any $\lambda \in (0,1)$, as can be easily verified.

In the following Corollary we summarise the sufficient conditions for the supermartingale property that follows from Theorem 3.2. The straightforward proof is omitted.

Corollary 3.1. *Let $\mathcal{A} \in \mathcal{F}$ and let $X, Y \in L^1(\mathcal{F})$ such that $\mathbb{1}_A, X$ and Y are pairwise independent. Let $Z = \mathbb{1}_A X + \mathbb{1}_{A^C} Y$ and let $\mathcal{G} = \{A, A^C, \emptyset, \Omega\}$. Let F and G be respectively the distribution functions of X and Y. The following conditions are sufficient for the supermartingale property $e_\alpha(Z) \geq \mathbb{E}[e_\alpha^{\mathcal{G}}(Z)]$.*

i) $\alpha > 1/2$ *and* $F(e_\alpha(X)) \geq G(e_\alpha(Y))$,

ii) $\alpha < 1/2$ *and* $F(z) \leq G(z)$, *for each* $z \in (e_\alpha(X), e_\alpha(Y))$.

Acknowledgements The authors are grateful to an Associate Editor and to two anonymous referees for their valuable reports which highly improved the paper. They would also like to thank Ruodu Wang for useful discussions and comments on the topic.

References

Acciaio, B. and I. Penner (2011). Dynamic risk measures. In *Advanced Mathematical Methods for Finance* (G. Di Nunno and B. Øksendal Eds.), pp. 1–34. Springer, Heidelberg.

Acciaio, B. and G. Svindland (2013). Are law-invariant risk functions concave on distributions? *Depend. Model. 1*, 54–64.

Acerbi, C. and B. Szekely (2014). Back-testing expected shortfall. *Risk Magazine, December*, 76–81.

Aliprantis, C. D. and K. C. Border (2006). *Infinite Dimensional Analysis: A Hitchhiker's Guide* (Third ed.). Springer, Berlin.

Ash, R. B. (1972). *Real Analysis and Probability.* Academic Press, New York-London.

Barron, E. N., P. Cardaliaguet, and R. Jensen (2003). Conditional essential suprema with applications. *Appl. Math. Optim. 48*(3), 229–253.

Bellini, F. and V. Bignozzi (2015). On elicitable risk measures. *Quant. Finance 15*(5), 725–733.

Bellini, F. and E. Di Bernardino (2017). Risk management with expectiles. *Europ. J. Finance 23*(6), 487–506.

Bellini, F., B. Klar, A. Müller, and E. Rosazza Gianin (2014). Generalized quantiles as risk measures. *Insurance Math. Econom. 54*, 41–48.

Bellini, F., R. J. A. Laeven, and E. Rosazza Gianin (2018). Robust return risk measures. *Math. Financ. Econ. 12*(1), 5–32.

Delbaen, F., F. Bellini, V. Bignozzi, and J. F. Ziegel (2016). Risk measures with the CxLS property. *Finance Stoch. 20*(2), 433–453.

Detlefsen, K. and G. Scandolo (2005). Conditional and dynamic convex risk measures. *Finance Stoch. 9*(4), 539–561.

Dudley, R. M. (2002). *Real Analysis and Probability.* Cambridge University Press, Cambridge.

Fissler, T., J. F. Ziegel, and T. Gneiting (2016). Expected shortfall is jointly elicitable with value at risk-implications for backtesting. *Risk Magazine, January*, 58–61.

Föllmer, H. and A. Schied (2011). *Stochastic Finance: An Introduction in Discrete Time* (extended ed.). Walter de Gruyter & Co., Berlin.

Gneiting, T. (2011). Making and evaluating point forecasts. *J. Amer. Statist. Assoc. 106*(494), 746–762.

Kaina, M. and L. Rüschendorf (2009). On convex risk measures on L^p-spaces. *Math. Methods Oper. Res. 69*(3), 475–495.

Koenker, R (2005). *Quantile Regression.* Cambridge University Press, Cambridge.

Kupper, M. and W. Schachermayer (2009). Representation results for law invariant time consistent functions. *Math. Financ. Econ. 2*(3), 189–210.

Kurdila, A. and M. Zabarankin (2005). *Convex Functional Analysis.* Birkhäuser Verlag, Basel.

Laeven, R. J. A. and M. Stadje (2013). Entropy coherent and entropy convex measures of risk. *Math. Oper. Res. 38*(2), 265–293.

Newey, W. K. and J. L. Powell (1987). Asymmetric least squares estimation and testing. *Econometrica 55*(4), 819–847.

Osband, K. (1985). *Providing incentives for better cost forecasting.* Ph. D. thesis, University of California, Berkeley.

Pflug, G. C. and W. Römisch (2007). *Modeling, Measuring and Managing Risk.* World Scientific, Hackensack, NJ.

Roorda, B. and J. M. Schumacher (2007). Time consistency conditions for acceptability measures, with an application to Tail Value at Risk. *Insurance Math. Econom. 40*(2), 209–230.

Weber, S. (2006). Distribution-invariant risk measures, information, and dynamic consistency. *Math. Finance 16*(2), 419–441.

Ziegel, J. F. (2016). Coherence and elicitability. *Math. Finance 26*(4), 901–918.