Elicitable risk measures

Fabio Bellini* and Valeria Bignozzi†

October 30, 2014

To appear in *Quantitative Finance*

Abstract

A statistical functional is elicitable if it can be defined as the minimizer of a suitable expected scoring function (see Gneiting, 2011; Ziegel, 2013, and the references therein). With financial applications in view, we suggest a slightly more restrictive definition than Gneiting (2011), and we derive several necessary conditions. For monetary risk measures, we show that elicitability leads to a subclass of the shortfall risk measures introduced by Föllmer and Schied (2002). In the coherent case, we show that the only elicitable risk measures are the expectiles. Further, we provide an alternative proof of the result in Ziegel (2013) that the only coherent comonotone elicitable risk measure is the expected loss.

Key-words: Elicitability, expectiles, decision theory, shortfall risk measures, VaR, mixture convexity.
JEL classification D81, C44, C13.

1 Introduction

Informally, a statistical functional T on a set of probability measures $\mathcal{M}$ on the real line is elicitable if it can be defined as the minimizer of a suitable expected scoring function. The simplest examples of elicitable functionals are the mean that minimizes a quadratic score and the quantiles that are the set of minimizers of a piecewise linear score. Despite this notion finds its roots in decision theory, see for example Savage (1971), the term 'elicitable functional' seems to have been introduced in Osband (1985). See also Lambert et al. (2008); Gneiting (2011) and the reference therein for an early history of the notion of elicitability.

We refer to the recent paper of Gneiting (2011) for the formal definition. Let $T : \mathcal{M} \to 2^{\mathbb{R}}$ be a possibly set-valued functional, where $2^{\mathbb{R}}$ is the power set of $\mathbb{R}$. T is elicitable relative to the class $\mathcal{M}$ if there exists a scoring function $S : \mathbb{R}^2 \to [0, +\infty)$ such that

$$T(F) = \arg\min_x \int S(x, y) \mathrm{d}F(y), \qquad \forall F \in \mathcal{M}. \tag{1.1}$$

If (1.1) holds, then we say that the scoring function S is strictly consistent with T.

*Dipartimento di Statistica e Metodi Quantitativi, Università di Milano-Bicocca, Via Bicocca degli Arcimboldi 8, 20126 Milan, Italy. Phone: +39 02 6448 3119. Email: fabio.bellini@unimib.it.

†Corresponding author. RiskLab, Department of Mathematics, ETH Zurich, 101 Ramistrasse, 8092 Zurich, Switzerland. Phone: +41 44 632 8615. Email: valeria.bignozzi@math.ethz.ch.

The main focus of the present contribution is the elicitability property of law-invariant monetary risk measures. For this reason we will define elicitability only for single-valued functionals, that in our opinion have a more transparent financial interpretation. The most common examples of monetary risk measure are Value-at-Risk (VaR) and Conditional Value-at-Risk (CVaR). The former one is defined as

$$\mathrm{VaR}_\alpha(F) = -q_\alpha^+(F), \qquad \alpha \in (0,1) \tag{1.2}$$

where $q_\alpha^+(F)$ is the right α-quantile of the P/L (profit and loss) distribution F. Despite its well known shortcomings (tail insensitivity and non-subadditivity, see Acerbi, 2002, for a review), it is still the most widespread risk measure in the banking and insurance sectors. CVaR_α, defined as

$$\mathrm{CVaR}_\alpha(F) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(F)\mathrm{d}u, \qquad \alpha \in (0,1)$$

was proposed by Acerbi and Tasche (2002) as a valid alternative to VaR, with the advantage that it belongs to the class of coherent risk measures (see Artzner et al., 1999).

The wide majority of risk measures known in the literature, such as VaR_α and CVaR_α, depend on the usually unknown distribution function of the financial position to assess. Hence they need to be estimated from a sample of available data. The robustness of risk measures to changes in the dataset and to different estimation procedures has been investigated, among others by Cont et al. (2010) and Krätschmer et al. (2013). In particular, it emerges that while CVaR_α is generally considered a better risk measure from a mathematical point of view, it requires a higher number of data for an accurate estimation (see e.g. Daníelsson, 2011) and it is less robust than VaR_α (see Cont et al., 2010).

Also the Basel Committee on Banking Supervision (2012), recently asked for a deeper understanding of the consequences of changing the regulatory regime from VaR_α to CVaR_α especially for what concerns robust backtesting. We recall that backtesting is the activity of periodically comparing the forecasted risk measure with the realized value of the variable under interest, in order to try to assess the accuracy of the forecasting methodology.

In this direction, it has been recently suggested by several authors (Gneiting, 2011; Ziegel, 2013; Embrechts and Hofert, 2013; Tasche, 2013) that elicitability is a key property for a risk measure as it provides a natural methodology to perform backtesting. Indeed, suppose that we want to provide an incentive to a forecaster to give an accurate assessment of a statistical functional (in our case a risk measure) T. If the functional T is elicitable and the forecaster is an expected score minimizer, than he/she can be induced to report a correct forecast by means of the expected score. In other words, if T is elicitable, a natural statistics to perform backtesting is given by the average expected score:

$$\widehat{S} = \frac{1}{n}\sum_{i=1}^{n} S(T_i, Y_i), \qquad n \in \mathbb{N}$$

where $T_1, \dots, T_n$ are point forecasts of the functional T and $Y_1, \dots, Y_n$ are outcomes of a random variables Y with distribution $F \in \mathcal{M}$. In particular, it is shown in Gneiting (2011) that while VaR_α (intended as quantile interval) is elicitable, CVaR_α is not. This supports the finding in Cont et al. (2010) that CVaR_α (and more generally the class of spectral risk measures introduced by Acerbi, 2002) are not ideal when estimation procedure is required. Before moving on, we have two remarks concerning the backtesting of VaR_α. First, it is well known (see e.g. Gneiting, 2011) that if

$$S(x,y) = \alpha(y-x)^+ + (1-\alpha)(y-x)^-, \tag{1.3}$$

then the corresponding expected score is minimized by the set of α-quantiles of $F \in \mathcal{M}$. For this reason we say that the 'quantile interval' is an elicitable, set-valued functional. However, the single-valued functional VaR_α as defined in (1.2) is not elicitable in the sense of (1.1). It becomes elicitable only in the case of distributions F with $q_\alpha^+(F) = q_\alpha^-(F)$. Second, the backtesting methodology for VaR_α based on the scoring function (1.3) is different from the standard approach used in practice that is based on a binomial test on the number of violations. We refer to Christoffersen (2010) for a review on the backtesting methodologies for VaR_α and Bellini and Figà-Talamanca (2007) for an early example of consistent backtesting of VaR_α.

It is important to emphasize that elicitability alone does not guarantee a correct ranking of different point forecasts. Indeed, while it ensures that the expected score is minimized only by exact reporting of the functional, it does not guarantee that given two point forecasts, the more accurate one (i.e. the closer one to the true value of the functional), receives a lower expected score. This is possible if in addition to elicitability we require that the scoring function is *accuracy rewarding* (in the terminology of Lambert et al., 2008), in the sense that for any $x_1, x_2 \in \mathbb{R}$ and $F \in \mathcal{M}$

$$T(F) < x_1 < x_2 \ \text{ or } \ x_2 < x_1 < T(F) \ \Rightarrow \ \int S(x_1,y)\mathrm{d}F(y) \leqslant \int S(x_2,y)\mathrm{d}F(y).$$

If that is the case, then we can compare different approaches for computing T (say, historical simulation, parametric approaches, Monte Carlo) in a natural and consistent way, by simply comparing the ex post realized expected score. Accuracy rewarding can be guaranteed by imposing weak technical conditions on the functional (as proposed in Lambert, 2011) or by requiring additional properties of the scoring function (as we suggest in this paper), see the discussion in Proposition 3.1.

Given the previous discussion on VaR_α and CVaR_α, one may wonder whether there exist a risk measure with good mathematical properties (e.g. convex or coherent risk measure) that is elicitable. The answer is positive, indeed in Gneiting (2011) it was shown that expectiles, introduced in the literature by Newey and Powell (1987), are elicitable and according to Bellini et al. (2014); Delbaen (2013) they are coherent. We refer the interested reader to Bellini (2012); Bellini et al. (2014); Delbaen (2011, 2013); Rémillard (2013) and the references therein for characterizations and properties of expectiles.

The recent paper of Ziegel (2013) deals with the characterization of coherent elicitable risk measures. Starting with an explicit construction based on the Kusuoka representation, the author proves that among the class of spectral risk measures, the expected loss is the only one to be elicitable. Further, it is shown that expectiles play a central role in the class of coherent risk measures as they provide a lower bound for elicitable risk measures. However, it is left open the question on whether they are the unique coherent risk measure that is elicitable. In the present contribution we continue in this direction. Under weak technical assumptions on the scoring function S and by means of the characterization theorem of shortfall risk measures in Weber (2006), we fully characterize the class of convex and coherent risk measures that are elicitable with an accuracy rewarding scoring function. In particular, we answer the question posed by

Ziegel (2013), showing that expectiles are indeed the only elicitable coherent risk measure.

For further motivations for the use of elicitable risk measures and their connection with backtesting we refer the interested reader to Gneiting (2011); Embrechts and Hofert (2013); Tasche (2013); Ziegel (2013), and the references therein.

Elicitability is also a very interesting mathematical property in itself. Several partial results toward its characterization have been obtained in Lambert et al. (2008) and Lambert (2011), although the problem still seems to be open, as mentioned in Gneiting (2011). An elementary necessary condition, pointed out by several authors (see Osband, 1985; Gneiting, 2011; Ziegel, 2013) is that an elicitable functional should have level sets that are convex with respect to mixtures. For this reason the CVaR_α is not elicitable, as it was shown by Gneiting (2011); an example displaying the same phenomenon can also be found in Weber (2006). We show that under weak hypotheses on the scoring function, another necessary condition to have elicitability and accuracy rewarding is that the functional is mixture continuous. For this we use a general result from set-valued analysis that is known as *Berge maximum theorem* (see for example Aliprantis and Border, 2006).

The paper is structured as follows: in Section 2 we introduce the basic notations and definitions; Section 3 deals with the characterization of elicitability in the general case, while Section 4 deals with monetary, convex and coherent risk measures.

2 Notations and basic definitions

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an atomless probability space and $L^\infty = L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ be the set of bounded random variables. We assume the convention that a random variable $X \in L^\infty$ represents a financial position (e.g. the P/L of a company). A monetary risk measure is a map $\rho: L^\infty \to \mathbb{R}$ that, for any $X, Y \in L^\infty$ and $m \in \mathbb{R}$, satisfies:

- *Monotonicity*: If $X \geqslant Y$, then $\rho(X) \leqslant \rho(Y)$;
- *Translation invariance*: $\rho(Y + m) = \rho(Y) - m$.

A monetary risk measure is convex if it also satisfies:

- *Convexity*: For any $\lambda \in [0,1]$, $\rho(\lambda X + (1-\lambda)Y) \leqslant \lambda\rho(X) + (1-\lambda)\rho(Y)$,

and it is coherent if it is convex and

- *Positive homogeneous*: For any $\alpha \geqslant 0$, $\rho(\alpha Y) = \alpha\rho(Y)$.

We refer to Artzner et al. (1999); Föllmer and Schied (2002); Frittelli and Rosazza Gianin (2002) among many others for a detailed discussion of monetary risk measures. The wide majority of risk measures used in practice and in the academic literature are law-invariant, meaning that the outcome of the risk measure depends uniquely on the distribution of the financial position to assess. Law-invariant risk measures can alternatively be defined as functional on the space of probability measures with compact support $\mathcal{M}_{1,c}(\mathbb{R})$. Hence the rest of the paper (except Section 4), will discuss properties of statistical functionals in general, having risk measures as a special case.

Let $\mathcal{M}_1(\mathbb{R})$ be the set of probability measures on $\mathbb{R}$. Each $\mu \in \mathcal{M}_1(\mathbb{R})$ will be represented by its distribution function $F(x) = \mu(-\infty, x]$. $\mathcal{M}_1(\mathbb{R})$ is endowed with the usual stochastic order $\leqslant_{st}$, also known as first order stochastic dominance, for any $F, G \in \mathcal{M}_1(\mathbb{R})$:

$$F \leqslant_{st} G \quad \text{if and only if} \quad F(x) \geqslant G(x), \qquad \forall x \in \mathbb{R}.$$

A single-valued statistical functional is defined as a map $T : \mathcal{M} \to \mathbb{R}$, where $\mathcal{M} \subseteq \mathcal{M}_1(\mathbb{R})$. We will always assume that $\mathcal{M}$ is *m-convex* (i.e., convex with respect to mixtures), in the sense that for any $\lambda \in [0,1]$

$$F, G \in \mathcal{M} \quad \Rightarrow \quad \lambda F + (1-\lambda)G \in \mathcal{M}.$$

We say that T is *monotone* if for any $F, G \in \mathcal{M}$,

$$F \leqslant_{st} G \quad \Rightarrow \quad T(F) \leqslant T(G),$$

and it is *translation invariant* if for every $h \in \mathbb{R}$ and $F \in \mathcal{M}$

$$T(F_h) = T(F) + h,$$

where F_h denotes the right-translated distribution function $F_h(x) := F(x-h)$.
We say that the functional T is *m-convex* if for each $F, G \in \mathcal{M}$ and $\lambda \in [0,1]$

$$T(\lambda F + (1-\lambda)G) \leqslant \lambda T(F) + (1-\lambda)T(G).$$

It is important to stress the difference between m-convexity, that is defined with respect to mixtures, and the usual notion of convexity of risk measures, that is defined with respect to sums. The two notions of convexity are not related and have very different interpretations: the former one is related to compound lotteries, the latter one to portfolios. We say that T is *m-quasiconvex* if for each $\gamma \in \mathbb{R}$ its lower level sets

$$\{T \leqslant \gamma\} := \{F \in \mathcal{M} \text{ s.t. } T(F) \leqslant \gamma\}$$

are m-convex, i.e. for any $\lambda \in [0,1]$ and $F, G \in \mathcal{M}$

$$F, G \in \{T \leqslant \gamma\} \quad \Rightarrow \quad \lambda F + (1-\lambda)G \in \{T \leqslant \gamma\},$$

while T is *m-quasiconcave* if the functional $-T$ is m-quasiconvex. A functional that is both m-quasiconvex and m-quasiconcave is called *m-quasilinear*. Finally, T has *convex level sets* (CxLS) if for each $\gamma \in \mathbb{R}$ the level sets

$$\{T = \gamma\} := \{F \in \mathcal{M} \text{ s.t. } T(F) = \gamma\}$$

are m-convex, that is if for any $\lambda \in [0,1]$ and $F, G \in \mathcal{M}$

$$T(F) = T(G) = \gamma \quad \Rightarrow \quad T(\lambda F + (1-\lambda)G) = \gamma.$$

Since $\{T = \gamma\} = \{T \leqslant \gamma\} \cap \{T \geqslant \gamma\}$, an m-quasilinear functional T has also CxLS; the converse is not necessarily true, as the following example shows:

example 2.1. Let $\mathcal{M}_1^1(\mathbb{R})$ be the set of probability distributions with finite mean and let $T : \mathcal{M}_1^1(\mathbb{R}) \to \mathbb{R}$ be given by

$$T(F) = \begin{cases} 1 & \text{if } \mathbb{E}[F] = 0 \\ -1 & \text{if } \mathbb{E}[F] > 0 \\ 0 & \text{if } \mathbb{E}[F] < 0. \end{cases}$$

It is clear that the level sets of T are m-convex, however

$$\{T \leqslant 0\} = \{F \in \mathcal{M}_1^1(\mathbb{R}) : \mathbb{E}[F] \neq 0\}$$

is not m-convex.

The *weak topology* on $\mathcal{M}_1(\mathbb{R})$ is the topology generated by the linear functionals of the form

$$F \mapsto \int f(y) \mathrm{d}F(y),$$

with $F \in \mathcal{M}_1(\mathbb{R})$ and $f : \mathbb{R} \to \mathbb{R}$ continuous and bounded. It is well known that the weak topology on $\mathcal{M}_1(\mathbb{R})$ is metrizable and that

$$F_n \to F \text{ weakly} \quad \Leftrightarrow \quad F_n(x) \to F(x) \;\text{ in the continuity points of } F.$$

A statistical functional T is *weakly continuous* if it is continuous with respect to the weak topology. As it has been pointed out by several authors, weak continuity is a very stringent property, not satisfied by the majority of risk measures (see for example Cont et al., 2010; Weber, 2006; Krätschmer et al., 2013). It is then natural to consider the more general notion of ψ-weak continuity. Let $\psi : \mathbb{R} \to [1, +\infty)$ be a continuous function which serves as *gauge function*, and let $C_\psi(\mathbb{R})$ be the linear space of all real continuous functions f for which there exists a constant c such that $|f| \leqslant c\psi$. The *ψ-weak topology* is the topology generated by the family of functionals

$$F \mapsto \int f(y) \mathrm{d}F(y), \qquad \text{with } f \in C_\psi, F \in \mathcal{M}_1(\mathbb{R}).$$

If ψ is bounded, then the ψ-weak topology corresponds to the weak topology; further we have that

$$F_n \to F \;\psi\text{-weakly} \quad \Leftrightarrow \quad F_n \to F \text{ weakly and } \int \psi \mathrm{d}F_n \to \int \psi \mathrm{d}F,$$

or equivalently

$$F_n \to F \;\psi\text{-weakly} \quad \Leftrightarrow \quad \int f \mathrm{d}F_n \to \int f \mathrm{d}F$$

for every continuous function $f \in C_\psi$. A statistical functional T is *ψ-weak continuous* if it is continuous with respect to the ψ-weak topology, that is if

$$F_n \to F \;\psi\text{-weakly} \quad \Rightarrow \quad T(F_n) \to T(F).$$

We refer to Föllmer and Schied (2004) for more details on the ψ-weak continuity. Finally, we say that T is *mixture continuous* if for each $\lambda \in [0, 1]$ and $F, G \in \mathcal{M}$ the function

$$\lambda \mapsto T(\lambda F + (1 - \lambda)G)$$

is continuous in λ. The following implications hold:

$$T \text{ weak continuous} \quad \Rightarrow \quad T\ \psi\text{-weak continuous} \quad \Rightarrow \quad T \text{ mixture continuous.}$$

Under additional assumptions, if T has CxLS then it is also m-quasilinear, as the following simple lemma shows:

Lemma 2.1. *If T has* CxLS *and is mixture continuous or monotone and translation invariant, then the sets $\{T < \gamma\}$, $\{T \leqslant \gamma\}$, $\{T > \gamma\}$ and $\{T \geqslant \gamma\}$ are m-convex.*

Proof. Consider $\lambda \in [0,1]$ and $F, G \in \mathcal{M}$ such that $H_\lambda = \lambda F + (1-\lambda)G \in \mathcal{M}$. Assume first that T is MC and let $T(F) \leqslant \gamma$ and $T(G) \leqslant \gamma$. If $T(F) = T(G) = \gamma$, then from CxLS also $T(H_\lambda) = \gamma$, so we can assume w.l.o.g. that $T(F) < T(G) = \gamma$. Let now by contradiction $T(H_\lambda) > \gamma$. From mixture continuity, since $T(H_\lambda) > T(G) > T(F)$, there exists $\lambda' \in [0,1]$ such that

$$T(\lambda' F + (1-\lambda')H_\lambda) = T(G) = \gamma,$$

that is

$$T((\lambda' + (1-\lambda')\lambda)F + (1-\lambda')(1-\lambda)G) = \gamma.$$

Since $\lambda' + (1-\lambda')\lambda > \lambda$, the distribution H_λ belongs to the segment joining the distributions $(\lambda' + (1-\lambda')\lambda)F + (1-\lambda')(1-\lambda)G$ and G, and hence from CxLS we would have $T(H_\lambda) = \gamma$, a contradiction. The same argument shows the m-convexity of the sets $\{T \geqslant \gamma\}$, $\{T < \gamma\}$ and $\{T > \gamma\}$. Let now T be monotone and translation invariant. Again, assume w.l.o.g. that $T(F) < T(G) = \gamma$. From translation invariance it follows that there exists $h > 0$ such that $T(F_h) = T(G) = \gamma$, and since $F \leqslant_{st} F_h$ when $h > 0$, from monotonicity and CxLS we have that

$$T(\lambda F + (1-\lambda)G) \leqslant T(\lambda F_h + (1-\lambda)G) = \gamma.$$

Again, similar arguments show the m-convexity of $\{T \geqslant \gamma\}$, $\{T < \gamma\}$ and $\{T > \gamma\}$. □

3 Elicitability

As anticipated in the introduction, in this paper we adopt a slightly different definition of elicitability than the one given by Gneiting (2011). First of all, we define the notion of elicitability only for *single-valued functionals* which implies that the minimizer of the expected scoring function is unique. Of course, from a purely mathematical point of view a unified treatment of single-valued and set-valued functionals (as done in Gneiting, 2011) is more natural, but we believe that from a financial point of view the interpretation of a set-valued risk measure is much less straightforward. For example, when we speak about VaR_α, we do not mean the interval of quantiles, but one of its endpoints.

Secondly we adopt the following properties for the scoring function.

Definition 3.1. A scoring function $S : \mathbb{R}^2 \to [0, \infty)$ satisfies: for any $x, y \in \mathbb{R}$

(a) $S(x,y) \geqslant 0$ and $S(x,y) = 0$ if and only if $x = y$;

(b) $S(x,y)$ is increasing in x for $x > y$ and decreasing for $x < y$;

(c) $S(x,y)$ is continuous in x.

Remark 3.1. While the definition of elicitability given in Gneiting (2011) does not in principle require any additional property on the scoring function S, most of the scoring functions used in practice (e.g. the scoring functions that elicit the mean, the quantile set, the median and the expectiles), verify (a), (c) and admit partial derivative $\frac{\partial S(x,y)}{\partial x}$ that is continuous in x whenever $x \neq y$. For this reason our assumptions are not really restrictive. Further, we will show in Propositions 3.1 and 3.2 that the extra condition (b) is necessary to ensure that S is also accuracy rewarding.

Definition 3.2. A statistical functional T is elicitable on $\mathcal{M}_T \subseteq \mathcal{M}$ if there exists a scoring function S as in Definition 3.1 such that for each $F \in \mathcal{M}_T$

a)
$$g_F(x) := \int S(x,y)\mathrm{d}F(y) < +\infty \qquad \forall x \in \mathbb{R}; \tag{3.1}$$

b)
$$T(F) = \arg\min_{x\in\mathbb{R}} \int S(x,y)\mathrm{d}F(y). \tag{3.2}$$

In this case we say that the scoring function S is strictly consistent with T.

Differently from Gneiting (2011) and Ziegel (2013), we allow the functional T to be elicitable on a subset of its domain. We call the maximal subset $\mathcal{M}_T$ of $\mathcal{M}$ on which (3.1) and (3.2) hold the *domain of elicitability* of T. Despite its simplicity, Definition 3.2 is not fully satisfactory since typically the domain of elicitability of T is smaller than its natural domain. For example, the mean is elicited by the scoring function $S(x,y) = (x-y)^2$, but its natural domain is $\mathcal{M}_1^1(\mathbb{R})$, not only on $\mathcal{M}_1^2(\mathbb{R})$. To overcome this problem, it is possible to generalize Definition 3.2 as follows:

Definition 3.3. A statistical functional T is elicitable on $\mathcal{M}_T \subseteq \mathcal{M}$ if there exists a scoring function S as in Definition 3.1 such that for each $F \in \mathcal{M}_T$

a)
$$h_F(x_1,x_2) := \int [S(x_2,y) - S(x_1,y)]\mathrm{d}F(y) \quad \text{is well defined for each } x_1,x_2 \in \mathbb{R};$$

b)
$$h_F(T(F),x) > 0 \qquad \forall x \in \mathbb{R},\ x \neq T(F).$$

A similar approach has been used in the literature on generalized quantiles in Jaworski (2006) and in robust statistics in Huber (1981). However, to avoid unnecessary technical complications we will always adopt the more simple and transparent Definition 3.2 in what follows. Consistently with the literature on risk measures, the rest of the paper focuses on functionals defined on $\mathcal{M}_{1,c}(\mathbb{R})$. When the domain of elicitability is not specified, it is intended to be the whole $\mathcal{M}_{1,c}(\mathbb{R})$.

It is well known (see Osband, 1985; Gneiting, 2011; Ziegel, 2013) that an elicitable functional T has CxLS. Our conditions on the scoring function imply additional properties for the elicited functional T:

Proposition 3.1. *If T is elicitable on $\mathcal{M}_{1,c}(\mathbb{R})$, then:*

a) $T(F) \in [\operatorname{ess\,inf}(F), \operatorname{ess\,sup}(F)]$. *In particular* $T(\delta_x) = x$, *for any* $x \in \mathbb{R}$*;*

b) T *has* CxLS*;*

c) T *is mixture continuous;*

d) T *is m-quasilinear;*

e) any scoring function S *that elicits* T *is accuracy rewarding.*

Proof. a) From Definition 3.1, we have that $g_F(x) := \int S(x, y)\mathrm{d}F(y)$ is decreasing for $x \leqslant \operatorname{ess\,inf}(F)$ and increasing for $x \geqslant \operatorname{ess\,sup}(F)$; it follows that $T(F) \in [\operatorname{ess\,inf}(F), \operatorname{ess\,sup}(F)]$. The second part is a direct consequence of the definition of δ_x that is the Dirac measure at the point $x \in \mathbb{R}$.
b) (See Osband, 1985; Gneiting, 2011; Ziegel, 2013); for completeness, we repeat the argument: let $F, G \in \mathcal{M}_{1,c}$ and $\lambda \in [0, 1]$. For $T(F) = T(G) = \gamma$, we have

$$T(\lambda F + (1-\lambda)G) = \arg\min_x \left\{\lambda \int S(x, y)\mathrm{d}F(y) + (1-\lambda) \int S(x, y)\mathrm{d}G(y)\right\} = \gamma.$$

c) Consider F, G, λ as before and

$$h_{F,G}(x, \lambda) := \lambda \int S(x, y)\mathrm{d}F(y) + (1-\lambda) \int S(x, y)\mathrm{d}G(y) < +\infty.$$

We claim that $h_{F,G}(x, \lambda)$ is jointly continuous in x and λ. Indeed, for $(x, \lambda) \in [\overline{x} - \varepsilon, \overline{x} + \varepsilon] \times [0, 1]$, we have that

$$\lambda S(x, y) \leqslant S(x, y) \leqslant \max\{S(\overline{x} - \varepsilon, y), S(\overline{x} + \varepsilon, y)\},$$

so from condition (3.1) and the dominated convergence theorem we have that

$$\lambda_n \int S(x_n, y)\mathrm{d}F(y) \to \lambda \int S(x, y)\mathrm{d}F(y) \text{ whenever } (x_n, \lambda_n) \to (x, \lambda).$$

A similar argument applies to the second term $(1-\lambda)\int S(x, y)\mathrm{d}G(y)$, so that joint continuity of $h_{F,G}(x, \lambda)$ is established. It follows that the minimization problem

$$\min_x \ h_{F,G}(x, \lambda)$$

has a jointly continuous objective function and from a) can be equivalently defined on a compact domain. We can then apply Berge maximum theorem (see for example Aliprantis and Border, 2006) to conclude that

$$\arg\min_x h_{F,G}(x, \lambda) \quad \text{is continuous in } \lambda,$$

that corresponds to the mixture continuity of T.
d) Since T has CxLS and is mixture continuous, it follows from Lemma 2.1 that T is m-quasilinear.
e) The result follows from mixture continuity and Proposition 2 in Lambert (2011). □

Proposition 3.2. *On* $\mathcal{M}_{1,c}(\mathbb{R})$ *the condition (b) in Definition 3.1 is necessary to have the accuracy rewarding of the scoring function S.*

Proof. Let us assume that S is accuracy rewarding and there exists $x_1, x_2, y \in \mathbb{R}$ such that $y \leqslant x_1 < x_2$ and $S(x_1, y) > S(x_2, y)$. Then we can choose a random variable Y with distribution F, where $F = \delta_y \in \mathcal{M}_{1,c}$ and

$$\mathbb{E}[S(x_1, Y)] = S(x_1, y) > S(x_2, y) = \mathbb{E}[S(x_2, Y)],$$

which contradicts the accuracy rewarding property. Similar contradictions can be obtained for $x_2 < x_1 \leqslant y$. □

If $S(x, y)$ is convex in x for each y, we have the following further characterization of the elicited functional T that generalizes the first order condition for generalized quantiles (see Bellini et al., 2014, and the references therein):

Proposition 3.3. *If T is elicitable with a scoring function $S(x, y)$ that is convex in x for each $y \in \mathbb{R}$, then*

$$T(F) = \sup\{m \in \mathbb{R} : \int \frac{\partial S}{\partial x^-}(m, y) dF(y) \leqslant 0\} = \inf\{m \in \mathbb{R} : \int \frac{\partial S}{\partial x^+}(m, y) dF(y) \geqslant 0\}. \quad (3.3)$$

Proof. From a) in Definition 3.2 and from the convexity of $S(x, y)$ in x we have that

$$g_F(x) := \int S(x, y) \mathrm{d}F(y)$$

is finite, convex, and hence continuous. From the dominated convergence theorem the left and the right derivatives of g_F are given by

$$\begin{aligned} g'_{F-}(x) &= \int \frac{\partial S}{\partial x^-}(x, y) \mathrm{d}F(y), \\ g'_{F+}(x) &= \int \frac{\partial S}{\partial x^+}(x, y) \mathrm{d}F(y), \end{aligned}$$

and hence the first order condition for the problem (3.2) can be written as

$$\int \frac{\partial S}{\partial x^-}(T(F), y) \mathrm{d}F(y) \leqslant 0 \leqslant \int \frac{\partial S}{\partial x^+}(T(F), y) \mathrm{d}F(y). \quad (3.4)$$

By the definition of elicitability there is exactly one real number $T(F)$ that satisfies this system of inequalities, and since the one sided partial derivatives $\frac{\partial S}{\partial x^-}(x, y)$ and $\frac{\partial S}{\partial x^+}(x, y)$ are increasing in x, representation (3.3) follows. □

In the recent literature on financial risk measures the notion of *robustness* has become increasingly important. In Krätschmer et al. (2013) the authors suggested that a proper notion of robustness for a risk measure is continuity with respect to the ψ-weak convergence. A very interesting point of this proposal is that it defines a *continuum* of degrees of robustness, in contrast with the more traditional binary notion of the statistical literature. For elicitable functionals, we can prove the following general robustness result:

Proposition 3.4. *Let $T : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ be elicited by the scoring function S, with*
a) $S(x, y)$ is convex in x for each $y \in \mathbb{R}$
b) $\frac{\partial S}{\partial x^+}(x, y)$ is continuous in y for each $x \in \mathbb{R}$
c) $\frac{\partial S}{\partial x^+}(x, y) \leqslant A(x) + B(x)\psi(y)$, for each $x, y \in \mathbb{R}$,
for some gauge function ψ. Then T is ψ-weak continuous.

Proof. From the characterization of $T(F)$ in Proposition 3.3, we know that

$$T(F) = \inf\{m \in \mathbb{R} : \int \frac{\partial S}{\partial x^+}(m, y)\mathrm{d}F(y) \geqslant 0\} = g^{-1}(0), \tag{3.5}$$

where g^{-1} is the left continuous inverse of

$$g(x) := \int \frac{\partial S}{\partial x^+}(x, y)\mathrm{d}F(y).$$

Let $F, F_1, F_2, \ldots$ be a sequence of measures in $\mathcal{M}_{1,c}(\mathbb{R})$ such that $F_n \to F$ ψ-weakly. Then the assumptions on $S(x, y)$ imply that

$$g_n(x) := \int \frac{\partial S}{\partial x^+}(x, y)\mathrm{d}F_n(y) \to g(x),$$

that in turn implies

$$T(F_n) = g_n^{-1}(0) \to g^{-1}(0) = T(F),$$

since under the elicitability hypothesis 0 is a continuity point of g^{-1}. $\square$

Remark 3.2. The mixture continuity and ψ-weak continuity results can also be generalized to the case in which the argmin of the expected scoring function does not reduce to a single point, that is to the case of *set-valued functionals*; the notion of continuity has then to be replaced with upper semicontinuity (see for example Aliprantis and Border, 2006).

We now present several examples on the elicitability property of well known statistical functionals:

example 3.1. (Left quantile) Let $T(F)$ be the left α-quantile of F, that is

$$T(F) = q_\alpha^-(F) = \inf\{m \in \mathbb{R} : F(m) \geqslant \alpha\}.$$

Then T is monotone and translation invariant, has CxLS, but is not mixture continuous, as it can be seen by considering for example $\alpha = \frac{1}{2}$, $F = \delta_0$, $G = \delta_1$. Indeed the function

$$\lambda \mapsto T(\lambda F + (1 - \lambda)G)$$

has a discontinuity in $\lambda = \frac{1}{2}$. So the left quantile is not elicitable on $\mathcal{M}_{1,c}(\mathbb{R})$ since it is not mixture continuous. However, it can be elicited on the set

$$\tilde{\mathcal{M}} = \{F \in \mathcal{M}_{1,c}(\mathbb{R}),\, F \text{ strictly increasing}\} = \{F \in \mathcal{M}_{1,c}(\mathbb{R}) : q_\alpha^+(F) = q_\alpha^-(F)\},$$

by means of the scoring function

$$S(x, y) = \alpha(y - x)^+ + (1 - \alpha)(y - x)^-.$$

example 3.2. Let $T(F) = \operatorname{ess\,sup} F$, this functional corresponds to the worst-case risk measure. Also in this case T is monotone, translation invariant and has CxLS, but since

$$T(\lambda F + (1 - \lambda)G) = \max(T(F), T(G)),$$

T is not mixture continuous and hence not elicitable.

example 3.3. Let us now consider the case of CVaR_α, that corresponds to the functional

$$T(F) = \frac{1}{1-\alpha}\int_\alpha^1 q_\beta^-(F)\,\mathrm{d}\beta.$$

It has been shown by Weber (2006) and Gneiting (2011) by means of simple examples that CVaR_α does not have CxLS and hence is not elicitable. Ziegel (2013) proved that actually the only spectral risk measure that is elicitable is the expected loss, we obtain the same result as a direct consequence of Theorem 4.4.

example 3.4. The *expectiles*, introduced in the statistical literature by Newey and Powell (1987) and further studied in Bellini et al. (2014) and Delbaen (2013), are defined as

$$T(F) = \arg\min_x \int [\alpha((y-x)^+)^2 + (1-\alpha)((y-x)^-)^2]\mathrm{d}F(y).$$

They are elicitable by definition and hence they satisfy properties a)-e). They are also monotone and translation invariant. We will see in the next section that they are the only example of an *elicitable coherent risk measure.*

example 3.5. An interesting and less well known example is the so called ΛVaR (Lambda-VaR), introduced in Frittelli et al. (2013), of which we consider a particular case. Let $\Lambda : \mathbb{R} \to (0,1)$ be continuous and strictly decreasing, with $\Lambda(x) \to 1^-$ for $x \to -\infty$ and $\Lambda(x) \to 0^+$ for $x \to +\infty$. Define

$$T(F) = \inf\{m \in \mathbb{R} : F(m) \geqslant \Lambda(m)\}.$$

In Frittelli et al. (2013) the authors have shown that T is monotone and m-quasiconvex, along with several other interesting properties. It is not difficult to see that this functional is elicitable, with a scoring function

$$S(x,y) = (x-y)^+ - \varphi(x,y),$$

where

$$\varphi(x,y) = \int_y^x \Lambda(t)\mathrm{d}t.$$

4 Monetary elicitable risk measures

In this section, we deal with financial risk measures on $\mathcal{M}_{1,c}(\mathbb{R})$. For $F \in \mathcal{M}_{1,c}(\mathbb{R})$, we define $\tilde{F}(t) := 1 - F_-(-t)$, where F_- is the left continuous version of F. $\tilde{F}$ is the distribution function of the losses corresponding to F, in other words for any random variable $X \in L^\infty$ with distribution $F \in \mathcal{M}_{1,c}(\mathbb{R})$, $-X$ has distribution $\tilde{F}$. This remark is important to have consistency between the academic literature on risk measures where X represents a financial position and the statistical literature where X is generally a loss. We give the following:

Definition 4.1. A risk measure $\rho : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ is *monetary* if $\rho(F) = T(\tilde{F})$, with T monotone and translation invariant; it is *elicitable* if $\rho(F) = T(\tilde{F})$, with T elicitable.

We say that a monetary risk measure is elicitable if it is an elicitable statistical functional of the loss.

As it is well known, a monetary risk measure induces an *acceptance set at the level of distributions* defined as

$$\mathcal{N} := \{F \in \mathcal{M}_{1,c}(\mathbb{R}) : \rho(F) \leqslant 0\},$$

and a corresponding rejection set

$$\mathcal{N}^c := \{F \in \mathcal{M}_{1,c}(\mathbb{R}) : \rho(F) > 0\}.$$

Under the hypotheses that $\mathcal{N}$ is non-empty, that $\inf\{m : \delta_m \in \mathcal{N}\} > -\infty$ and that

$$\text{for any } F \in \mathcal{N} \text{ s.t } F \leqslant_{st} G \quad \Rightarrow \quad G \in \mathcal{N}.$$

The risk measure ρ can be recovered from $\mathcal{N}$ as

$$\rho(F) = \inf\{x \in \mathbb{R} : F(\cdot - x) \in \mathcal{N}\},$$

(see for example Weber, 2006; Föllmer and Schied, 2004). In this section, we will see that the class of monetary elicitable risk measures coincides with a subclass of the *shortfall risk measures* introduced by Föllmer and Schied (2002). We recall the basic definition:

Definition 4.2. Let $\ell : \mathbb{R} \to \mathbb{R}$ be increasing and not constant, with x_0 in the interior of the convex hull of the range of ℓ. We define the shortfall risk measure ρ_ℓ as

$$\rho_\ell(F) := \inf\{x \in \mathbb{R} : \int \ell(-y - x)\mathrm{d}F(y) \leqslant x_0\}. \tag{4.1}$$

We recall from Föllmer and Schied (2002) that if ℓ is continuous, then from the dominated convergence theorem $\rho_\ell(F)$ is a solution of the equation

$$\int \ell(-y - \rho_\ell(F))\mathrm{d}F(y) = x_0, \tag{4.2}$$

while if ℓ is strictly increasing then the solution of (4.2) is unique.

Remark 4.1. It is easy to see that without loss of generality it is always possible to assume that $x_0 = 0$ and that $\ell(-\infty, 0) \subseteq (-\infty, 0)$ and $\ell(0, +\infty) \subseteq (0, +\infty)$. The last condition guarantees that the shortfall risk measure ρ_ℓ has the *constancy* property $\rho_\ell(\delta_y) = -y$ that must be satisfied by any elicitable monetary risk measure from Proposition 3.1 item a).

In order to show that an elicitable monetary risk measure is a shortfall, we will use the characterization of shortfall risk measures that has been given in Weber (2006). For completeness we report here its main result:

Theorem 4.1. *(Weber) Let ρ be a monetary risk measure on $\mathcal{M}_{1,c}(\mathbb{R})$ with acceptance set $\mathcal{N}$. If there exists $x \in \mathbb{R}$ with $\delta_x \in \mathcal{N}$ such that, for each $y \in \mathbb{R}$ with $\delta_y \in \mathcal{N}^c$, it holds*

$$(1 - \alpha)\delta_x + \alpha\delta_y \in \mathcal{N} \qquad \textit{for sufficiently small } \alpha > 0, \tag{4.3}$$

then the following assumptions are equivalent:
a) the acceptance and the rejection sets $\mathcal{N}$ and $\mathcal{N}^c$ are m-convex and $\mathcal{N}$ is ψ-weakly closed for some gauge function ψ;
b) ρ is a shortfall risk measure with a left continuous ℓ.

We can now prove that under a very weak hypothesis on S, an elicitable and monetary risk measure is a shortfall.

Theorem 4.2. *Let $\rho : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ be a monetary and elicitable risk measure with a scoring function $S(x, y)$ that is continuous in y and, for any $x \in [-\epsilon, \epsilon]$ with $\epsilon > 0$, satisfies $S(x, y) \leqslant \psi(y)$ for some gauge function ψ. Then it is a shortfall risk measure.*

Proof. If ρ is elicitable, then from Proposition 3.1 it has CxLS and it is mixture continuous. From Lemma 2.1 it follows that the set $\mathbb{N} = \{F : \rho(F) \leqslant 0\}$ and $\mathbb{N}^c = \{F : \rho(F) > 0\}$ are m-convex. Let $\delta_x \in \mathbb{N}$ and $\delta_y \in \mathbb{N}^c$; from mixture continuity it is easy to see that (4.3) in Theorem 4.1 holds. To apply Theorem 4.1 and conclude that ρ is a shortfall, it remains to show that $\mathbb{N}$ is ψ-weakly closed for some gauge function ψ. Recall that

$$\mathcal{N} = \{F \in \mathcal{M}_{1,c}(\mathbb{R}) : \rho(F) \leqslant 0\},$$

where

$$\rho(F) = T(\tilde{F}) = \arg\min_x \int S(x, y) d\tilde{F}(y) = \arg\min_x \int S(x, -y)\mathrm{d}F(y).$$

Since under our hypotheses the scoring function $S(x, y)$ is accuracy rewarding, we have that

$$\rho(F) \leqslant 0 \iff \int S(0, -y)\mathrm{d}F(y) \leqslant \int S(\epsilon, -y)\mathrm{d}F(y),$$

for each $\epsilon > 0$. Let now $F_1, F_2, \ldots$ be a sequence of measures in $\mathcal{N}$, with $F_n \to F$ ψ-weakly. We have that

$$\int S(0, -y)\mathrm{d}F_n(y) \leqslant \int S(\epsilon, -y)\mathrm{d}F_n(y).$$

Since $S(x, y)$ is bounded by $\psi(y)$ for $x \in (-\epsilon, \epsilon)$ and continuous in the second argument, it follows

$$\int S(0, -y)\mathrm{d}F_n(y) \to \int S(0, -y)\mathrm{d}F(y), \quad \int S(\epsilon, -y)\mathrm{d}F_n(y) \to \int S(\epsilon, -y)\mathrm{d}F(y) \quad \text{and} \quad \rho(F) \leqslant 0.$$

Thus

$$\int S(0, -y)\mathrm{d}F(y) \leqslant \int S(\epsilon, -y)\mathrm{d}F(y),$$

so that $\mathcal{N}$ is ψ-weakly closed and we can apply Theorem 4.1 to conclude that

$$\rho(F) = \inf\{x \in \mathbb{R} : \int \ell(-y - x)\mathrm{d}F(y) \leqslant 0\},$$

with a left continuous ℓ. □

Unfortunately, not all shortfall risk measures are elicitable; for example $\mathrm{VaR}_\alpha(F)$ is a shortfall risk measure with loss

$$\ell(z) = 1_{\{z \geqslant 0\}} - \alpha, \tag{4.4}$$

but as we saw in Example 3.1 it is not mixture continuous and hence not elicitable on the whole $\mathcal{M}_{1,c}(\mathbb{R})$. A similar problem can also arise with a continuous loss function, as the following example shows:

example 4.1. Let

$$\ell(z) = \begin{cases} -2 & \text{if } z < -1 \\ 2z & \text{if } z \in [-1, 1] \\ 2 & \text{if } z > 1. \end{cases}$$

The corresponding shortfall ρ_ℓ is given by

$$\rho_\ell(F) = \inf\{m \in \mathbb{R} : \int \ell(-y - m)\mathrm{d}F(y) \leqslant 0\},$$

and is not mixture continuous, as it can be seen by considering $F = \delta_{-2}$, $G = \delta_2$ and $H_\lambda = \lambda F + (1 - \lambda)G$. Indeed, it is not difficult to check that

$$\rho(H_\lambda) = \begin{cases} -1 & \text{if } \lambda = 1/2 \\ \frac{2-3\lambda}{1-\lambda} & \text{if } \lambda > 1/2, \end{cases}$$

which has a discontinuity in $\lambda = 1/2$.

These examples show that a converse of Theorem 4.2 does not hold in general; there can be shortfall risk measures that are not elicitable. From an intuitive point of view, the problem is that while it is always possible to interpret a shortfall risk measure as a first order condition for a corresponding expected scoring function, the uniqueness of the minimizer is not always guaranteed. However, under additional assumptions on the strict monotonicity of the loss function ℓ, the following converse of Theorem 4.2 holds:

Theorem 4.3. *Let $\rho_\ell(F)$ be a shortfall risk measure as in (4.1). If ℓ is left continuous and strictly increasing on $(-\infty, \epsilon)$ or on $(-\epsilon, +\infty)$ for some $\epsilon > 0$, then ρ can be elicited by the scoring function*

$$S(x, y) = \varphi(y - x), \text{ with } \varphi(z) = \int_0^z \ell(t)\, dt. \tag{4.5}$$

Proof. Clearly $S(x, x) = 0$, and since $\ell(-\infty, 0) \subseteq (-\infty, 0)$, $\ell(0, +\infty) \subseteq (0, +\infty)$ and ℓ is strictly increasing on $(-\epsilon, \epsilon)$, it follows that $S(x, y) = 0 \Rightarrow x = y$. Further, $S(x, y)$ is convex in x since φ is convex, so $S(x, y)$ is strictly increasing for $x \geqslant y$ and strictly decreasing for $x \leqslant y$ and thus it is a scoring function according to Definition 3.1. In order to prove the elicitability, we have to show that the shortfall risk measure

$$\rho(F) = \inf\{m : \int \ell(-y - m)\mathrm{d}F(y) \leqslant 0\}$$

is the only minimizer of the expected scoring function

$$g_F(x) := \int S(x, -y)\mathrm{d}F(y) = \int \varphi(-y - x)\mathrm{d}F(y).$$

From the left continuity of ℓ and (4.5), we have that

$$\varphi'_-(z) = \ell(z),$$

and hence from the dominated convergence theorem it follows that

$$g'_{F-}(x) = -\int \ell(-y - x)\mathrm{d}F(y).$$

Hence the shortfall risk measure $\rho(F)$ satisfies

$$\rho(F) = \inf\{m : g'_{F-}(m) \geqslant 0\},$$

that shows that

$$\rho(F) \in \arg\min g_F,$$

since g_F is convex. To prove the uniqueness of the minimizer, it is enough to note that under our assumption on ℓ it follows that $g'_{F-}(x)$ is strictly increasing for $x \in (\operatorname{ess\,inf}(F), \operatorname{ess\,sup}(F))$, so that g_F is strictly convex in this interval. □

Remark 4.2. A shortfall risk measure with a convex ℓ is always elicitable, since from the assumption that 0 belongs to the interior of the convex hull of the range of ℓ it follows that ℓ is continuous on $\mathbb{R}$ and strictly increasing on $(-\epsilon, +\infty)$ for some $\epsilon > 0$, so that the hypotheses of Theorem 4.3 are satisfied.

example 4.2. Let us consider the strictly increasing loss function

$$\ell(z) = \begin{cases} z+1 & \text{if } z \geqslant 0 \\ z-1 & \text{if } z < 0. \end{cases}$$

The corresponding shortfall risk measure can be written explicitly as

$$\rho(F) = \inf\{m : E[Y] - m + P(Y > m) - P(Y \leqslant m) \leqslant 0\},$$

and is elicitable, as a consequence of Theorem 4.3. Indeed, it is not difficult to see that it is elicited by the scoring function

$$S(x, y) = \varphi(y - x), \text{ with } \varphi(z) = \frac{z^2}{2} + |z|\,,$$

that is the sum of the scores corresponding to the mean and to the median. S is strictly convex in x for each y (as in the case of the mean) but it is not differentiable in $x = y$ (as in the case of the median).

example 4.3. Consider the strictly increasing loss function

$$\ell(z) = e^{\beta z} - 1, \text{ with } \beta > 0,$$

that corresponds to the entropic risk measure given by

$$\rho(F) = \frac{1}{\beta} \log\Big(\int e^{-\beta y} \mathrm{d}F(y)\Big).$$

Since ℓ is strictly increasing we can apply Theorem 4.3 and obtain

$$\varphi(z) = \frac{e^{\beta z} - 1}{\beta} - z,$$

so that $T(F)$ is elicited by the scoring function

$$S(x, y) = \frac{e^{\beta(y-x)} - 1}{\beta} - (y - x).$$

Combining Theorem 4.2 with the characterization of convexity and coherency properties of shortfall risk measures that have been obtained in Weber (2006), it is not difficult to characterize elicitable convex risk measures, elicitable coherent risk measures and elicitable coherent comonotone risk measures. We recall that a risk measure is comonotone if for any $X, Y \in L^\infty$ such that

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geqslant 0 \quad \text{for all } (\omega, \omega') \in \Omega \times \Omega$$

then $\rho(X+Y) = \rho(X) + \rho(Y)$. The class of law-invariant comonotone coherent risk measures coincides with the class of spectral risk measures introduced by Acerbi (2002). We have the following:

Theorem 4.4. *Let $\rho : \mathcal{M}_{1,c}(\mathbb{R}) \to \mathbb{R}$ be an elicitable monetary risk measure satisfying the hypotheses of Theorem 4.2. Then:*

a) If ρ is convex then it is a shortfall risk measure with a convex loss ℓ;

b) if ρ is coherent then it is an expectile with $\alpha \geqslant \frac{1}{2}$;

c) if ρ is coherent and comonotone then it is the expected loss.

Proof. From Theorem 4.2 a monetary elicitable functional is a shortfall risk measures, so a) follows from the well known characterization of convex shortfall risk measures that can be found in Föllmer and Schied (2002). From Weber (2006) we know that any coherent shortfall risk measure can be represented as the unique solution of

$$\mathbb{E}[a(Y - T(F))^+ - b(Y - T(F))^-] = 0 \tag{4.6}$$

where $a \geqslant b > 0$. By defining $\alpha = \frac{a}{a+b}$ we can rewrite the loss function as $\ell(y) = \alpha y^+ - (1-\alpha) y^-$ that is the loss function of an expectile with $\alpha \geqslant \frac{1}{2}$. In general, expectiles are not comonotone, as it has been shown by many authors (see for example Bellini et al., 2014; Delbaen, 2013); the only exception is the mean, that gives (c). □

These results are in full accordance with Ziegel (2013), that starting from the Kusuoka representation derived several characterizations of elicitable coherent risk measures and showed that the expected loss is the only coherent and comonotone elicitable risk measure (in other words, no other coherent distorted risk measure is elicitable). Both papers highlight the central role played by expectiles and thus provide a further motivation for their study.

Acknowledgments

The authors would like to thank Paul Embrechts, Johanna Ziegel, Ruodu Wang, Elena Di Bernardino and two anonymous referees for valuable comments and suggestions.

References

Acerbi, C. (2002). Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking and Finance 26*(1), 1505–1518.

Acerbi, C. and D. Tasche (2002). On the coherence of Expected Shortfall. *Journal of Banking and Finance 26*(7), 1487–1503.

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

Artzner, P., F. Delbaen, J. Eber, and D. Heath (1999). Coherent measures of risk. *Mathematical Finance 9*(3), 203–228.

Basel Committee on Banking Supervision (May 2012). Fundamental review of the trading book. Available from www.bis.org/publ/bcbs219.htm.

Bellini, F. (2012). Isotonicity results for generalized quantiles. *Statistics and Probability Letters 82*(11), 2017–2024.

Bellini, F. and G. Figà-Talamanca (2007). Conditional tail behaviour and Value at Risk. *Quantitative Finance 7*(6), 599–607.

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

Christoffersen, P. (2010). *Backtesting. Encyclopedia of Quantitative Finance.* Wiley.

Cont, R., R. Deguest, and G. Scandolo (2010). Robustness and sensitivity analysis of risk measurement procedures. *Quantitative Finance 10*(6), 593–606.

Daníelsson, J. (2011). *Financial Risk Forecasting.* Chichester: Wiley.

Delbaen, F. (2011). Monetary utility functions. In *Osaka University CSFI Lecture Notes Series 3*.

Delbaen, F. (2013). A remark on the structure of expectiles. Working paper, ETH Zurich.

Embrechts, P. and M. Hofert (2013). Statistics and quantitative risk management for banking and insurance. To appear in Annual Review of Statistics and its Applications.

Föllmer, H. and A. Schied (2002). Convex measures of risk and trading constraints. *Finance and Stochastics 6*(4), 429–47.

Föllmer, H. and A. Schied (2004). *Stochastic Finance: An Introduction in Discrete Time* (Second ed.). de Gruyter.

Frittelli, M., M. Maggis, and I. Peri (2013). Risk measures on $\mathcal{P}(\mathbb{R})$ and Value at Risk with Probability/Loss function. To appear in Mathematical Finance.

Frittelli, M. and E. Rosazza Gianin (2002). Putting order in risk measures. *Journal of Banking and Finance 26*(7), 1473–1486.

Gneiting, T. (2011). Making and evaluating point forecast. *Journal of the American Statistical Association 106*(494), 746–762.

Huber, P. (1981). *Robust Statistics.* Wiley.

Jaworski, P. (2006). On a subjective approach to risk measurement. *Quantitative Finance 6*(6), 495–511.

Krätschmer, V., A. Schied, and H. Zähle (2013). Comparative and qualitative robustness for law-invariant risk measures. Working paper, University of Mannheim.

Lambert, N., D. M. Pennock, and Y. Shoham (2008). Eliciting properties of probability distributions. *Proceedings of the 9th ACN Conference on Electronic commerce*, 129–138.

Lambert, N. S. (2011). Elicitation and evaluation of statistical forecasts. Working paper, Stanford University.

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

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

Rémillard, B. (2013). *Statistical Methods For Financial Engineering.* Chapman & Hall/CRC.

Savage, L. J. (1971). Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association 66*(336), 783–801.

Tasche, D. (2013). Risk measures: Yet another search for the holy grail. Available at: http://www.newton.ac.uk/programmes/TGM/seminars/2013032809301.pdf.

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

Ziegel, J. (2013). Coherency and elicitability. Working paper, University of Bern.