Loss-Based Risk Measures

Rama Cont, Romain Deguest, Xuedong He

Loss-based risk measures

Rama CONT[1,3], Romain DEGUEST [2] and Xue Dong HE[3]

1) Laboratoire de Probabilités et Modèles Aléatoires
CNRS- Université Pierre et Marie Curie, France.
2) EDHEC Risk Institute, Nice (France).
3) IEOR Dept, Columbia University, New York.

2011

Abstract

Starting from the requirement that risk measures of financial portfolios should be based on their losses, not their gains, we define the notion of *loss-based risk measure* and study the properties of this class of risk measures. We characterize loss-based risk measures by a representation theorem and give examples of such risk measures. We then discuss the statistical robustness of estimators of loss-based risk measures: we provide a general criterion for qualitative robustness of risk estimators and compare this criterion with sensitivity analysis of estimators based on influence functions. Finally, we provide examples of statistically robust estimators for loss-based risk measures.

Contents

A main focus of quantitative modeling in finance has been to measure the risk of financial portfolios. In connection with the widespread use of Value-at-Risk and related risk measurement methodologies, a considerable theoretical literature Acerbi (2002, 2007); Artzner et al. (1999); Cont et al. (2010); Föllmer and Schied (2002); Föllmer and Schied (2004); Frittelli and Rosazza Gianin (2002); Heyde et al. (2010); McNeil et al. (2005) has focused on the design of appropriate risk measures for financial portfolios. In this approach, a risk measure is represented as a map $\rho : L^\infty \to \mathbb{R}$ assigning to each (bounded) random variable $X \in L^\infty(\Omega, \mathcal{F}, \mathbb{P})$—representing the gain of a portfolio—a number which measures the risk of this portfolio. A framework often used as a starting point is the axiomatic setting of Artzner et al. (1999), which defines a *coherent risk measure* as a map $\rho : L^\infty \to \mathbb{R}$ that is

1. *Monotone* (*decreasing*): $\rho(X) \leq \rho(Y)$ provided $X \geq Y$.
2. *Cash-additive (additive with respect to cash reserves)*: $\rho(X + c) = \rho(X) - c$ for any $c \in \mathbb{R}$.
3. *Positive homogeneous*: $\rho(\lambda X) = \lambda\rho(X)$ for any $\lambda \geq 0$.
4. *Sub-additive*: $\rho(X + Y) \leq \rho(X) + \rho(Y)$.

Artzner et al. (1999) argue that these axioms correspond to desirable properties of a risk measure, such as the reduction of risk under diversification.

However, some of these axioms have unintended consequences and have been revised in various ways. Replacing positive homogeneity property by the more general convexity property leads to the class of convex risk measures (Föllmer and Schied, 2002; Föllmer and Schied, 2004; Frittelli and Rosazza Gianin, 2002); replacing sub-additivity by co-monotonic sub-additivity (Heyde et al., 2010; Song and Yan, 2009) allows to restore the robustness of risk estimators (Cont et al., 2010).

Even more basic issues arise when considering the above axioms as the sole characterization of a "risk measure". Intuitively, the risk of a portfolio is associated with the magnitude of its losses, not its gains. In the case of a statistical risk measure, this means that the risk measure should be defined in terms of the *left* tail of the portfolio gain (distribution). Indeed, the most popular risk measures—Value at Risk and expected shortfall—are based on the left tail of the loss distribution. However, this natural property is not implied by any of the axioms of coherent risk measures as outlined in Artzner et al. (1999).

If one further distinguishes, as in Heyde et al. (2010), 'external' risk measures used for determining capital requirements from 'internal' risk measures used for a firm's risk management –determining position limits, allocation of capital, etc.—then the seemingly innocent axiom of cash-additivity—that the risk of a portfolio decreases by c if combined with a cash position of size c—also comes under question. This axiom makes sense if $\rho(X)$ is interpreted as a capital requirement for portfolio X, i.e., for external risk measures. But if $\rho(X)$ is used to gauge the risk of the portfolio X, this axiom implies that if one combines a risky position X with a cash position $c = \rho(X)$ then the combined position $X + c$ has a "zero risk": $\rho(X + c) = 0$, which is meaningless in a risk management context. Also, as noted already by Jarrow (2002), the axiom of cash-additivity excludes some natural examples of risk measures, such as the value of the put option on a portfolio's net value (Jarrow, 2002).

El Karoui and Ravanelli (2009) challenge the axiom of cash additivity from a different angle. They show that, when considering risk measures defined on *future* (instead of discounted) value of portfolio gains in order to take into account interest rate risk , cash additivity is inevitably violated and needs to be replaced by cash subadditivity.

Another issue, very important in practice but somewhat neglected in the literature on risk measures, is the issue of statistical estimation of risk, and the design of robust estimators for risk measures: lack of robustness of a risk estimation procedure makes it difficult, if not impossible, to use in practice. Unfortunately, as shown in Cont et al. (2010), another unintended consequence of the axioms of coherent risk measures is that lead to non-robust risk estimators: the subadditivity property requires dependence on extreme tail events, which leads to high sensitivity to outliers Cont et al. (2010).

We consider in this paper an alternative approach to defining risk measures which addresses these concerns. Starting from the requirement that risk measures of financial portfolios should be based on their losses, not their gains, we define the notion of *loss-based risk measure* and study the properties of this class of risk measures.

We claim that our theoretical framework is a more natural one for the purpose of risk measurement and management of financial portfolios, and englobes various natural examples of risk measures used in practice. Examples of such loss-based risk measures have been considered by Jarrow (2002), who studied the put option premium as a risk measure, and Staum (2011), in the context of a finite dimensional probability space. Our analysis extends these results to a more general setting: we use a slightly different set of axioms and obtain a characterization of these risk measures on a general probability space.

Section 1 defines loss-based risk measures and provides a representation theorem in terms of a penalty function (Theorem 1), which adapts the results of Föllmer and Schied (2004) to our setting. This representation result is made more explicit in the case of distribution-based risk measures (Theorem 2) and examples are given in Section 1.5. We then study the statistical robustness, in the sense of Cont et al. (2010), of estimators for loss-based risk measures. In the case of statistical (i.e. distribution-based) risk measures we present a general criterion for qualitative robustness of such risk estimators (Section 2). These results extend results of Cont et al. (2010) to loss-based risk measures. In Section 2.3 we provide statistically robust versions of these risk estimators. The theoretical results on the qualitative robustness of these estimators are confirmed by a quantitative sensitivity analysis based on influence functions (Section 3).

1 Loss-Based Risk Measures

1.1 Notations

Consider an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$ representing market scenarios. For a random variable X, denote by $F_X(\cdot)$ its cumulative distribution function and $G_X(\cdot)$ its left-continuous quantile function. For any $p \in [1, \infty)$, let $\mathbb{L}^p(\Omega, \mathcal{F}, \mathbb{P})$ be the space of random variables X with the norm $\|X\|_p := (\mathbb{E}[|X|^p])^{\frac{1}{p}}$, and let $\mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$

be the space of bounded random variables.

Let

$$\mathcal{P}(\Omega, \mathcal{F}, \mathbb{P}) := \{X \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P}) : X \geq 0, ||X||_1 = 1\},$$

and

$$\mathcal{M}(\Omega, \mathcal{F}, \mathbb{P}) := \{X \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P}) : X \geq 0, ||X||_1 \leq 1\}.$$

Then $\mathcal{P}(\Omega, \mathcal{F}, \mathbb{P})$ can also be regarded as the set of $\mathbb{P}$-absolutely continuous probability measures on $(\Omega, \mathcal{F}, \mathbb{P})$, and $\mathcal{M}(\Omega, \mathcal{F}, \mathbb{P})$ as the set of $\mathbb{P}$-absolutely continuous measures μ such that $\mu(\Omega) \leq 1$.

Denote by $\mathcal{P}((0,1))$ the set of probability measures on the open unit interval $(0,1)$, and $\mathcal{M}((0,1))$ the set of positive measures μ on $(0,1)$ such that $\mu((0,1)) \leq 1$. Let

$$\Psi((0,1)) := \left\{\phi : (0,1) \mapsto \mathbb{R}_+ \mid \phi(\cdot) \text{ is decreasing on (0,1) and } \int_0^1 \phi(z)dz \leq 1\right\}, \tag{1}$$

$$\Phi((0,1)) := \left\{\phi : (0,1) \mapsto \mathbb{R}_+ \mid \phi(\cdot) \text{ is decreasing on (0,1) and } \int_0^1 \phi(z)dz = 1\right\}, \tag{2}$$

both of which can be identified as subsets of $\mathcal{M}((0,1))$. Finally, for any random variables X, let $X \wedge 0 := \min(X, 0)$.

1.2 Definition

Consider a map $\rho : \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}$ which associates to each random variable X, representing the future payoff of a portfolio, a number $\rho(X)$ representing its risk. Artzner et al. (1999) have argued that, in order for ρ to qualify as a monetary risk measure, it has to satisfy at least the two following properties:

(i) *Cash additivity:* adding cash to a portfolio reduces the risk by the same amount.

$$\forall X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}), \forall \alpha \in \mathbb{R}, \rho(\alpha + X) = \rho(X) - \alpha$$

(ii) *Monotonicity:* if a portfolio X has higher payoff than a portfolio Y in all scenarios, then it has lower risk.

$$\forall X, Y \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}), \qquad X \leq Y \Rightarrow \rho(X) \geq \rho(Y).$$

The cash-additivity property (i) is necessary to interpret $\rho(X)$ in terms of capital requirements for X: the capital requirement for a portfolio X may be satisfied by combining it with a position in cash equal to $\rho(X)$ since $\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0$. However if one uses $\rho(X)$ as a measure of risk, one might be tempted to think that the position $X + c$ has "no residual risk", which is counterintuitive and, indeed, nonsensical if taken literally. This observation suggests that the cash-invariance property may not be a reasonable requirement for a good risk measure.

Also, as noted by Jarrow (2002), the cash-additivity requirement excludes a natural risk measure, the premium of a put on the loss:

$$\rho(X) = \mathbb{E}[\max(-X, 0)]$$

Moreover, the initial motivation of risk measures is to quantify the risk associated with portfolio's losses. Thus, a risk measure should focus on the loss of a portfolio, and two portfolios which have the same losses should lead to the same risk measure.

With these concerns in mind, we define a class of risk measures, which we call *loss-based risk measures*, as follows.

Definition 1 (Loss-based risk measures) A mapping $\rho : \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}) \to \mathbb{R}_+$ is called a *loss-based risk measure* if it satisfies

(a) *Cash Loss:* for any $\alpha \in \mathbb{R}_+$, $\rho(-\alpha) = \alpha$;

(b) *Monotonicity:* for any $X, Y \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$,

$$\text{if} \quad X \leq Y, \quad \text{then} \quad \rho(X) \geq \rho(Y).$$

(c) *Convexity:* for any $X, Y \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$ and $0 < \alpha < 1$,

$$\rho(\alpha X + (1-\alpha)Y) \leq \alpha\rho(X) + (1-\alpha)\rho(Y).$$

(d) *Loss-dependency:* the risk of a portfolio only depends on its losses, i.e.,

$$\forall X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}), \rho(X) = \rho(X \wedge 0)$$

.

Property (a) is a "normalization" property which states that the risk of a (non-random) cash liability is its face value. Properties (b) and (c) have well-known interpretations in terms of increasing loss diversification, see e.g. Föllmer and Schied (2002). Property (d) entails that *portfolios with the same losses have the same risk*. This implies that in fact the monotonicity property (b) may be replaced by the weaker

$$X \wedge 0 \leq Y \wedge 0 \Rightarrow \quad \rho(X) \geq \rho(Y).$$

Note that property (d) prevents the risk measure from satisfying the cash-additivity property (i): a loss-based risk measure is not "coherent" in the sense of Artzner et al. (1999).

Loss-based risk measures are related to the class of cash subadditive risk measures defined by El Karoui and Ravanelli El Karoui and Ravanelli (2009). Indeed, properties (a) and (d), together with a weak semi-continuity condition, imply the cash subadditivity property defined in El Karoui and Ravanelli (2009, Def 3.1). To see this, let us first assume the following semi-continuity condition: for any $X_n \to X$

in $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$, $\rho(X) \leq \text{liminf}_{n\to\infty}\rho(X_n)$. This condition is weaker than the Fatou property commonly assumed in the risk measure literature. Now, consider any $X \in \mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$ and $\alpha \geq 0$. For any $\epsilon \in (0,1)$, we have

$$\begin{aligned}\rho((1-\epsilon)X-\alpha) &= \rho((1-\epsilon)X+\epsilon(-\frac{\alpha}{\epsilon})) \\ &\leq (1-\epsilon)\rho(X)+\epsilon\rho(-\frac{\alpha}{\epsilon})) \\ &= (1-\epsilon)\rho(X)+\alpha,\end{aligned}$$

where the inequality is due to the convexity property (c) and the last equality is due to the Cash Loss property (a). Letting $\epsilon \downarrow 0$ and using semi-continuity, we have

$$\rho(X-\alpha) \leq \rho(X)+\alpha, \quad \forall X \in \mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P}), \alpha \geq 0,$$

which also implies

$$\rho(X+\alpha) \geq \rho(X)-\alpha, \quad \forall X \in \mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P}), \alpha \geq 0.$$

Thus, loss-based risk measures are cash–subadditive. On the other hand, it is easy to see that cash subadditivity in the sense of El Karoui and Ravanelli (2009) does not imply the Cash Loss property (a).

However, our focus in this study is not only on cash-subadditivity: the Loss-dependency property (d) is at least as important as replacing cash additivity by cash subadditivity in a risk management context. Thus, loss-based risk measures deserve to be highlighted and treated separately although they are special cases of cash subadditive risk measures El Karoui and Ravanelli (2009).

The following useful lemma shows that that Convexity (c) and Monotonicity (b) together imply $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$ continuity for a risk measure.

Lemma 1 *Any map $\rho : \mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P}) \to \mathbb{R}$ which is monotone and convex is continuous on $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$.*

Proof Consider a sequence $(X_n)_{n\geq 1}$ that converges to X in $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$. Then the sequence $\varepsilon_n := ||X_n - X||_\infty$, $n \geq 1$, converges to zero and we can assume without loss of generality that $\varepsilon_n \leq 1, n \geq 1$. Let $\alpha_n := \sqrt{\varepsilon_n}$, $\beta_n := \frac{\varepsilon_n}{\alpha_n} + ||X||_\infty + 1$, then

$$(1-\alpha_n)X_n - \alpha_n\beta_n = X_n - \varepsilon_n - \alpha_n(||X||_\infty + 1 + X_n) \leq X$$

where the inequality is because $X_n - X \leq ||X_n - X||_\infty$ and $||X||_\infty + 1 + X_n \geq 0$. By the monotonicity and convexity of ρ, we derive

$$\begin{aligned}\rho(X) &\leq \rho((1-\alpha_n)X_n - \alpha_n\beta_n) \\ &\leq (1-\alpha_n)\rho(X_n) + \alpha_n\rho(-\beta_n) \\ &\leq (1-\alpha_n)\rho(X_n) + \alpha_n\rho(-(||X||_\infty + 2)),\end{aligned}$$

which leads to

$$\rho(X_n) \geq \frac{\rho(X) - \alpha_n\rho(-(||X||_\infty + 2))}{1-\alpha_n}.$$

Letting $n \to \infty$ and using the fact that α_n converges to zero, we immediately have $\liminf_{n\to\infty}\rho(X_n) \geq \rho(X)$, i.e., ρ is lower semi-continuous in $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$. A similar argument leads to the upper semi-continuity of ρ in $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$. ∎

1.3 Representation Theorem

We give now a characterization of loss-based risk measures that have a semi-continuity property, the Fatou property. A (loss-based) risk measure ρ satisfies the *Fatou property* if, for any sequence $(X_n)_{n\geq 1}$ uniformly bounded in $\mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$ such that $X_n \to X$ almost surely, we have

$$\rho(X) \leq \liminf_{n\to\infty} \rho(X_n). \tag{3}$$

Theorem 1 *The following are equivalent*

1. *ρ is a loss-based risk measure satisfying Fatou property* (3).
2. *There exists a convex function $V : \mathcal{M}(\Omega, \mathcal{F}, \mathbb{P}) \to [0, \infty]$ satisfying*

$$\inf_{\|Y\|_1 \geq 1-\epsilon} V(Y) = 0 \quad \textit{for any} \quad \epsilon \in (0, 1), \tag{4}$$

such that

$$\rho(X) = -\inf_{Y\in\mathcal{M}(\Omega,\mathcal{F},\mathbb{P})} \{\mathbb{E}[(X \wedge 0)Y] + V(Y)\}, \quad \forall X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}). \tag{5}$$

Proof Assume ρ is a loss-based risk measure satisfying the Fatou property. First, following Theorem 2.1 in Delbaen and Schachermayer (1994) or Theorem 3.2 in Delbaen (2002), ρ is lower semi-continuous under weak* topology if and only if it satisfies the Fatou property. By Frittelli and Rosazza Gianin (2002, Theorem 6) there exists $V : \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P}) \to (-\infty, \infty]$ such that

$$\begin{aligned}\rho(X) &= -\inf_{Y\in\mathbb{L}^1(\Omega,\mathcal{F},\mathbb{P})} \{\mathbb{E}[XY] + V(Y)\} \\ &= -\inf_{Y\in\mathbb{L}^1(\Omega,\mathcal{F},\mathbb{P})} \{\mathbb{E}[(X \wedge 0)Y] + V(Y)\}, \quad \forall X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}),\end{aligned}$$

where the second equality is due to the Loss-dependency property (d) of the loss-based risk measure. Furthermore, we have the dual relation

$$V(Y) = \sup_{X\in\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})} \{-\rho(X) - \mathbb{E}[XY]\}, \quad \forall Y \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P}).$$

For any $Y \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, let $A := \{Y < 0\}$. If $\mathbb{P}(A) > 0$,

$$\begin{aligned}V(Y) &\geq \sup_{n\geq 1}\{-\rho(n\mathbf{1}_A) - \mathbb{E}[n\mathbf{1}_A Y]\} \\ &= \sup_{n\geq 1}\{-\mathbb{E}[n\mathbf{1}_A Y]\} \\ &= +\infty,\end{aligned}$$

where the first equality holds because $n\mathbf{1}_A \geq 0$ and ρ is monotone with $\rho(0) = 0$. Now, if $Y \geq 0$ and $\|Y\|_1 > 1$, we have

$$\begin{aligned} V(Y) &\geq \sup_{\alpha \geq 0}\{-\rho(-\alpha) - \mathbb{E}[(-\alpha)Y]\} \\ &= \sup_{\alpha \geq 0}\{(\mathbb{E}[Y] - 1)\alpha\} \\ &= +\infty. \end{aligned}$$

Thus the domain of V lies in $\mathcal{M}(\Omega, \mathcal{F}, \mathbb{P})$. Next, it is easy to see that $V(Y) \geq 0$ for any $Y \in \mathbb{L}^1(\Omega, \mathcal{F}, \mathbb{P})$, because $\rho(0) = 0$. Finally, for any $\epsilon \in (0, 1)$ and any $\alpha > 0$,

$$\begin{aligned} \alpha = \rho(-\alpha) &= \sup_{Y \in \mathcal{M}(\Omega, \mathcal{F}, \mathbb{P})} \{\alpha \mathbb{E}[Y] - V(Y)\} \\ &= \max\left(\sup_{Y \in \mathcal{M}(\Omega, \mathcal{F}, \mathbb{P}), \|Y\|_1 < 1-\epsilon} \{\alpha \mathbb{E}[Y] - V(Y)\}, \sup_{Y \in \mathcal{M}(\Omega, \mathcal{F}, \mathbb{P}), \|Y\|_1 \geq 1-\epsilon} \{\alpha \mathbb{E}[Y] - V(Y)\} \right) \\ &\leq \max\left(\alpha(1-\epsilon), \sup_{Y \in \mathcal{M}(\Omega, \mathcal{F}, \mathbb{P}), \|Y\|_1 \geq 1-\epsilon} \{\alpha \mathbb{E}[Y] - V(Y)\} \right) \\ &\leq \max\left(\alpha(1-\epsilon), \alpha - \inf_{Y \in \mathcal{M}(\Omega, \mathcal{F}, \mathbb{P}), \|Y\|_1 \geq 1-\epsilon} \{V(Y)\} \right). \end{aligned}$$

Thus we conclude that $V(\cdot)$ must satisfy (4).

On the other hand, one can check that ρ represented in (5) is a loss-based risk measure satisfying the lower-semi-continuity under weak* topology and thus the Fatou property. ∎

We can see in the representation theorem that the domain of the penalty function $V(\cdot)$ is a subset of $\mathcal{M}(\Omega, \mathcal{F}, \mathbb{P})$, the set of all measures with total mass less than one. This property is also observed in El Karoui and Ravanelli (2009) (Theorem 4.3-(b)). Compared with the representation theorem in El Karoui and Ravanelli (2009), we have an additional condition (4) due to Cash Loss property. Moreover, the dual representation formula (5) only depends on the negative part of X due to the Loss-dependency property (d).

1.4 Loss-based statistical risk measures

Most of the risk measures used in finance are statistical, or distribution-based risk measures, i.e. they depend on X only through its distribution $F_X(\cdot)$:

$$F_X(\cdot) = F_Y(\cdot) \quad \Rightarrow \quad \rho(X) = \rho(Y).$$

Following ideas from Kusuoka (2001), Frittelli and Rosazza Gianin (2005), and Jouini et al. (2006), we derive a representation theorem for loss-based statistical risk measures.

Theorem 2 *Let ρ be a loss-based statistical risk measure. There exists a convex function $v : \Psi((0,1)) \to [0, \infty]$ satisfying*

$$\inf_{\int_0^1 \phi(z)dz \geq 1-\epsilon} v(\phi) = 0 \text{ for any } \epsilon \in (0,1), \tag{6}$$

such that

$$\rho(X) = -\inf_{\phi\in\Psi}\left\{\int_0^1 (G_X(z)\wedge 0)\phi(z)dz + v(\phi)\right\}, \quad \forall X\in\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P}). \tag{7}$$

Proof First, we remark that a statistical risk measure necessarily satisfies Fatou property. Indeed, as noted in Theorem 2.2, Jouini et al. (2006), distribution-based convex functionals in $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$ satisfy Fatou property if and only if they are lower semi-continuous in $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$ norm. By Lemma 1, any loss-based risk measures is $\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$-continuous, so automatically satisfies the Fatou property if it is distribution-based. We then need to build a connection between (5) and (7) when the risk measure is distribution-based, which can be done following the lines of Jouini et al. (2006, Theorem 2.1). ∎

Motivated by the representation (7), we sometimes abuse the notation by writing

$$\rho(G(\cdot)) = -\inf_{\phi\in\Psi}\left\{\int_0^1 (G(z)\wedge 0)\phi(z)dz + v(\phi)\right\} \tag{8}$$

for any bounded quantile functions, if we are considering a loss-based statistical risk measure.

Remark 1 If, in addition, the loss-based risk measure ρ satisfies

(e) *Cash-loss Invariance:* for any $X\in\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$, $X\leq 0$, and $\alpha\in\mathbb{R}_+$, $\rho(X-\alpha) = \rho(X)+\alpha$,

then one can select V in the representation (5) and v in the representation (7) such that the domain of V lies in $\mathcal{P}(\Omega,\mathcal{F},\mathbb{P})$ and the domain of v lies in $\Phi((0,1))$. In other words, $\rho(X) = \widetilde{\rho}(X\wedge 0)$ for some convex monetary risk measure $\widetilde{\rho}$. Indeed, for any $X\in\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$, one can define

$$\widetilde{\rho}(X) = \rho(X-\alpha_X) - \alpha_X,$$

where α_X is any upper-bound of X. By the Cash-loss Invariance property of ρ, $\widetilde{\rho}$ is well-defined. Furthermore, it is easy to check that $\widetilde{\rho}$ is a cash-invariant risk measure and $\rho(X) = \rho(X\wedge 0) = \widetilde{\rho}(X\wedge 0)$.

However, even though ρ satisfies the Cash-loss Invariance property, it is still not cash invariant. Indeed, in general, we only have

$$\rho(X+\alpha) = \widetilde{\rho}((X+\alpha)\wedge 0) \geq \widetilde{\rho}(X\wedge 0+\alpha) = \widetilde{\rho}(X\wedge 0) - \alpha = \rho(X) - \alpha,$$

for any $X\in\mathbb{L}^\infty(\Omega,\mathcal{F},\mathbb{P})$ and $\alpha\in\mathbb{R}$.

1.5 Examples

Example 1 (Put option premium) Jarrow (2002) argues that a natural measure of a firm's insolvency risk is the premium of a put option on the firms equity, which is given by the positive part of its net value X (assets minus liabilities), i.e.

$\mathbb{E}^{\mathbb{Q}}[-\min(X,0)]$ where $\mathbb{Q}$ is an appropriately chosen pricing model. One can generalize this to any portfolio whose net value is represented by a random variable X: the downside risk of the portfolio can be measured by

$$\rho(X) := \mathbb{E}[-\min(X,0)] \tag{9}$$

This example satisfies all the properties of Definition 1: it is a loss-based risk measure. In particular, as noted by Jarrow (2002), it is not cash-additive.

Example 2 (Scenario-based margin requirements) When determining margin requirements for derivative transactions, the objective of a central clearing facility is to compute the margin requirement of each participant in the clearinghouse in order to cover losses incurred by the clearing participants portfolio over a liquidation horizon T (typically, a few days). A popular method for computing such margin requirements –used for example by various futures and options exchanges– is to select a certain number of "stress scenarios" for risk factors affecting the portfolio and compute the margin requirement as the maximum loss over these scenarios. If one denotes the portfolio P&L over the horizon T by X, then the margin requirement $\rho(X)$ is given by

$$\rho(X) = \max\{-\min(X(\omega_1),0), ..., -\min(X(\omega_n),0)\} \tag{10}$$

where $\omega_1, ..., \omega_n$ are the stress scenarios. Naturally, the clearinghouse only consider the losses of the portfolio when computing the margin: as a result, such margin requirements may be viewed as a loss-based risk measure. This example satisfies all the properties of Definition 1: it is a loss-based risk measure. This is the main idea behind the SPAN method used by the Chicago Merchantile Exchange (CME). Interestingly, the SPAN method was considered as an initial motivation for the definition of coherent risk measures in Artzner et al. (1999). Yet it is easy to check that (10) is loss-dependent and therefore not cash-additive, so is not a coherent risk measure.

Example 3 (Expected tail-loss) This popular risk measure is defined as

$$\rho(X) := -\frac{1}{\beta}\int_0^\beta (G_X(z) \wedge 0)dz, \quad X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}) \tag{11}$$

Note that, by construction, this risk measure focuses on the left tail of the P&L distribution, since it only involves the quantile function on $(0, \beta)$. Nonetheless, the classical definition of expected shortfall $-\frac{1}{\beta}\int_0^\beta G_X(z)dz$ does not satisfy Loss-dependency property (d) of Definition 1 since $G_X(\beta)$ might be greater that 0 for some X. Therefore, we insert $G_X(z) \wedge 0$ in its definition to turn it into a loss-based risk measure. We notice that the put option premium is an expected tail-loss by taking $\beta = 1$.

Example 4 (Spectral tail measures) A large class of loss-based statistical risk measures is obtained by taking weighted averages of quantiles with various weight functions $\phi \in \Phi$:

$$\rho(X) := -\int_0^1 (G_X(z) \wedge 0)\phi(z)dz, \quad X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}) \tag{12}$$

We call such a risk measures a *spectral tail measure* : it is a loss-based version of the spectral risk measures defined by Acerbi (2002). This example satisfies the Cash-loss Invariance property (e) and is also *positively homogeneous*: for any $\lambda > 0$, $\rho(\lambda X) = \lambda\rho(X)$. Notice that the expected tail loss is a spectral tail measure, with $\phi(z) = \frac{1}{\beta}1_{(0,\beta)}(z)$.

Example 5 (Loss certainty equivalent) Consider $u(\cdot) \in C^4(\mathbb{R}_+)$ which is strictly increasing and strictly convex. Assume u'/u'' is concave. Consider the following map:

$$\rho(X) := u^{-1}\left(\mathbb{E}u(|X \wedge 0|)\right), \quad X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}). \tag{13}$$

It is clear that ρ satisfies Cash Loss, Monotonicity, and Loss-dependency in Definition 1. By Hardy et al. (1959, Theorem 106), ρ is also convex. Thus, ρ is a loss-based convex risk measure. We call ρ a *loss certainty equivalent* . By definition, ρ is distributional-based and

$$\rho(X) = u^{-1}\left(\int_0^1 u(|G_X(z) \wedge 0|)dz\right), \quad X \in \mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P}). \tag{14}$$

If $u(x) = x^p, x \geq 0$ for certain $p \geq 1$, we will speak of the L^p *loss certainty equivalent* . Here, when $p = 1$, $u(\cdot)$ is not strictly convex. However, in this case, (13) is still well-defined and the risk measure is actually the put option premium, which is also a loss-based risk measure. Thus, we include the case $p = 1$ here and identify the put option premium as a special case of L^p loss certainty equivalent in the following. One can show that the L^p loss certainty equivalent is the only loss certainty equivalent satisfying *positive homogeneity*, i.e., for any $\lambda > 0$, $\rho(\lambda X) = \lambda\rho(X)$. The L^p loss certainty equivalent has the following dual representation

$$\rho(X) = -\inf_{Y \in \mathcal{M}^q(\Omega,\mathcal{F},\mathbb{P})} \mathbb{E}\left[(X \wedge 0)Y\right] \tag{15}$$

where $1 < q \leq \infty$ is the conjugate of p, i.e., $\frac{1}{p} + \frac{1}{q} = 1$, and $\mathcal{M}^q(\Omega, \mathcal{F}, \mathbb{P})$ is the set of all nonnegative random variables with L^q norms less than or equal to one. Moreover, it has the distribution-based representation

$$\rho(X) = -\inf_{\phi \in \Psi^q((0,1))} \int_0^1 (G_X(z) \wedge 0)\phi(z)dz. \tag{16}$$

where $\Psi^q((0,1))$ is the set of $\phi \in \Psi((0,1))$ such that $\int_0^1 \phi(z)^q dz \leq 1$.

Let $u(x) = e^{\beta x}, x \geq 0$ for certain $\beta > 0$, then the loss certainty equivalent becomes the *entropic loss certainty equivalent* , a loss-based version of the entropic risk measure studied in Frittelli and Rosazza Gianin (2002); Föllmer and Schied (2002); Föllmer and Schied (2004). One can show that the entropic loss certainty equivalent and L^1 loss certainty equivalent are the loss certainty equivalent satisfying Cash-loss Invariance property. It is well known in the literature that the entropic loss certainty equivalent has the dual representation

$$\rho(X) = -\inf_{Y \in \mathcal{P}(\Omega,\mathcal{F},\mathbb{P})} \{\mathbb{E}[(X \wedge 0)Y] + V(Y)\}, \tag{17}$$

where

$$V(Y) = \mathbb{E}\left[Y \ln Y\right] - \inf_{Y \in \mathcal{P}(\Omega,\mathcal{F},\mathbb{P})} \mathbb{E}\left[Y \ln Y\right], \tag{18}$$

and the distributional-based representation

$$\rho(X) = - \inf_{\phi \in \Phi((0,1))} \int_0^1 (G_X(z) \wedge 0)\phi(z)dz + v(\phi), \tag{19}$$

where

$$v(\phi) = \int_0^1 \phi(z) \ln \phi(z) dz - \inf_{\phi \in \Phi((0,1))} \int_0^1 \phi(z) \ln \phi(z) dz. \tag{20}$$

See e.g. Frittelli and Rosazza Gianin (2002); Föllmer and Schied (2002); Föllmer and Schied (2004); Frittelli and Rosazza Gianin (2005).

2 Robustness of risk estimators

In practice, for measuring the risk of a portfolio, aside from the theoretical choice of a risk measure, a key issue is the estimation of the risk measure, which requires the choice of a *risk estimator* (Cont et al. (2010)). In this section we study the robustness property of empirical risk estimators built from some distribution-based risk measure. We follow the ideas in Cont et al. (2010) but we will view risk measures as functionals on the set of quantile functions, rather than the set of distribution functions. As we will see, this makes the study of continuity properties of loss-based risk measures easier.

Moreover, from Theorem 2, a loss-based statistical risk measure can be represented by (7), which suggests that it is more natural to work directly on quantile functions.

Denote by $\mathcal{Q}$ the set of all quantile functions and by $\mathcal{D}$ the set of all distribution functions. The Lévy-Prokhorov metric between two distribution functions $F_1 \in \mathcal{D}$ and $F_2 \in \mathcal{D}$ is defined as

$$d_P(F_1, F_2) \triangleq \inf\{\epsilon > 0 \,:\, F_1(x-\epsilon) - \epsilon \leq F_2(x) \leq F_1(x+\epsilon) + \epsilon,\ \forall x \in \mathbb{R}\}.$$

This metric appears to be the most tractable one on $\mathcal{D}$ and it induces the same topology as the usual weak topology on $\mathcal{D}$.

The quantile set $\mathcal{Q}$ and the distribution set $\mathcal{D}$ are connected by the following one-on-one correspondence

$$\begin{array}{ccc} \mathcal{D} & \to & \mathcal{Q} \\ F(\cdot) & \mapsto & F^{-1}(\cdot) \end{array}$$

where

$$F^{-1}(t) = \inf\left\{x \in \mathbb{R} \mid F(x) \geq t\right\}, \quad t \in (0,1)$$

is the left continuous inverse of $F(\cdot)$. Such a correspondence, together with the Lévy-Prokhorov metric on $\mathcal{D}$ induces a metric on $\mathcal{Q}$ which we denote by d. The convergence

under this metric can be characterized by the following: for any $G_n, G \in \mathcal{Q}$, $G_n \to G$ if and only if $G_n(z) \to G(z)$ at any continuity points of G.

Most of the time, we work with quantile functions that are continuous on $(0,1)$ in order to avoid irregularities due to the presence of atoms. In practice, it is not restrictive to focus on continuous quantile functions. Indeed, people do assume the continuity of quantile functions in many applications, e.g., when computing the Value at Risk. The study of discontinuous quantile functions is more technical and of little interest, so we choose not to pursue in this direction. In the following, we denote by $\mathcal{Q}_c$ the set of all continuous quantile functions. We also denote by $\mathcal{Q}^\infty$ the set of all bounded quantile functions, and $\mathcal{Q}_c^\infty$ the set of all bounded continuous quantile functions.

2.1 A Family of Distribution-based Risk Measures

Motivated by the representation (7), we consider the following *Fenchel-Legendre transform* $\rho : \mathcal{Q} \mapsto [0, \infty]$:

$$\rho(G(\cdot)) := - \inf_{m \in \mathrm{dom}(v)} \left\{ \int_{(0,1)} (G(z) \wedge 0) m(dz) + v(m) \right\}, \tag{21}$$

where $v : \mathcal{M}((0,1)) \to [0, \infty]$ is the *penalty function* satisfying

$$\inf_{m((0,1)) \geq 1-\epsilon} v(m) = 0 \text{ for any } \epsilon \in (0,1), \tag{22}$$

and $\mathrm{dom}(v)$ is the domain of v, i.e.,

$$\mathrm{dom}(v) := \{m \in \mathcal{M}((0,1)) \mid v(m) < \infty\}. \tag{23}$$

It is easy to see that $\rho(G(\cdot))$ is well-defined for any $G(\cdot) \in \mathcal{Q}$, and $\mathrm{dom}(\rho)$, the domain of ρ, contains $\mathcal{Q}^\infty$. It is clear that (21) is a generalization of (7). Such a generalization in the context of risk measures has been considered by Cont et al. (2010) in the case of spectral risk measures, by Heyde et al. (2010) for risk measures on finite-dimensional probability spaces, and by Song and Yan (2009) in a general setting.

The generalization (21) sacrifices the convexity of ρ viewed as a map on $\mathbb{L}^\infty(\Omega, \mathcal{F}, \mathbb{P})$. However, if ρ is a map defined on $\mathcal{Q}$, it satisfies the following properties

1. *Cash Loss:* for any $\alpha \in \mathbb{R}_+$, $\rho(-\alpha) = \alpha$;
2. *Monotonicity:* for any $G_1(\cdot), G_2(\cdot) \in \mathcal{Q}$ such that $G_1(z) \leq G_2(z), 0 < z < 1$, $\rho(G_1(\cdot)) \geq \rho(G_2(\cdot))$;
3. *Quantile Convexity:* for any $G_1(\cdot), G_2(\cdot) \in \mathcal{Q}$ and $0 < \alpha < 1$, $\rho(\alpha G_1(\cdot) + (1-\alpha) G_2(\cdot)) \leq \alpha \rho(G_1(\cdot)) + (1-\alpha)\rho(G_2(\cdot))$.
4. *Loss-dependency:* for any $G(\cdot) \in \mathcal{Q}$, $\rho(G(.)) = \rho(G(.) \wedge 0)$.

2.2 Qualitative Robustness

In practice, in order to compute the risk measure $\rho(G(\cdot))$ of portfolio whose P&L has quantile $G(\cdot)$, one has first to estimate the distribution or quantile of the P&L distribution of the portfolio from data, and then apply the risk measure ρ to the estimated distribution. One of the popular ways is to apply the risk measure to the empirical distribution.

A *sequence of samples* of $G(\cdot) \in \mathcal{Q}$ is a sequence of random variables $X_1, X_2, \cdots$ which are i.i.d and follow the distribution G^{-1}. We denote by $\mathbf{X}$ this sequence of samples and $\mathbf{X}^n$ its first n samples. For each sample size n, the *empirical distribution* is defined as

$$F^{\text{emp}}_{\mathbf{X}^n}(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{X_i \le x}, \quad x \in \mathbb{R}, \tag{24}$$

and the *empirical quantile* is defined as

$$G^{\text{emp}}_{\mathbf{X}^n}(z) := (F^{\text{emp}}_{\mathbf{X}^n})^{-1}(z) = X_{(\lfloor nz \rfloor + 1)}, \quad z \in (0,1), \tag{25}$$

where $\lfloor a \rfloor$ denotes the integer part of a and $X_{(1)} \le \cdots \le X_{(n)}$. In practice, the quantity $\widehat{\rho}(\mathbf{X}^n) := \rho(G^{\text{emp}}_{\mathbf{X}^n}(\cdot))$ is computed as the estimated risk measure $\rho(G(\cdot))$ given the n samples $\mathbf{X}^n$. The risk estimator $\widehat{\rho}$ is defined on $\cup_{n\ge 1}\mathbb{R}^n$, the set of all possible samples $(\mathbf{X}^n)_{n\ge 1}$, and has values in $\mathbb{R}^+$. Since the samples can be regarded as random variables, so is the risk estimator $\widehat{\rho}(\mathbf{X}^n)$. We denote by $\mathcal{L}_n(\widehat{\rho}, G)$ the distribution function of $\widehat{\rho}(\mathbf{X}^n)$.

ρ is said to be *consistent* at $G = F^{-1} \in \text{dom}(\rho)$ if

$$\lim_{n\to\infty} \widehat{\rho}(\mathbf{X}^n) = \rho(F^{\text{emp}}_{\mathbf{X}^n})^{-1}) \qquad \text{almost surely.} \tag{26}$$

Because the true risk measure $\rho(G(\cdot))$ is estimated by $\widehat{\rho}(\mathbf{X}^n)$, the consistency is the minimal requirement for a meaningful risk measure. In the following we denote by $\mathcal{Q}^\rho$ the set of quantiles G at which ρ is well defined and consistent, and $\mathcal{Q}^\rho_c$ the continuous quantiles in $\mathcal{Q}^\rho$.

The following definition of *robust risk estimator* is considered in Cont et al. (2010, Definition 4):

Definition 2 (Cont et al. (2010)) Let ρ be defined by (21) and $\mathcal{C} \subset \mathcal{Q}^\rho$ be a a set of plausible loss quantiles. $\widehat{\rho}$ is *$\mathcal{C}$-robust* at $G \in \mathcal{C}$ if for any $\varepsilon > 0$ there exist $\delta > 0$ and $n_0 \ge 1$ such that, for all $\widetilde{G} \in \mathcal{C}$,

$$d(\widetilde{G}, G) \le \delta \Longrightarrow d_P(\mathcal{L}_n(\widehat{\rho}, \widetilde{G}), \mathcal{L}_n(\widehat{\rho}, G)) < \varepsilon, \qquad \forall n \ge n_0.$$

A risk estimator $\widehat{\rho}$ is called *$\mathcal{C}$-robust* if it is $\mathcal{C}$-robust at any $G \in \mathcal{C}$.

The choice of the Lévy-Prokhorov distance is natural in robust statistics (see Huber (1981)). It is worth mentioning, as pointed out by Huber (1981), that the use of a different metric, even if it also metrizes the weak topology, may lead to a different class of robust estimators.

The following proposition which is cited from Cont et al. (2010) shows that the robustness of the risk estimator $\widehat{\rho}$ is equivalent to the continuity of the risk measure ρ.

Proposition 1 (Cont et al. (2010)) *Let ρ be a risk measure and $G \in \mathcal{C} \subset \mathcal{Q}_\rho$. The following are equivalent:*

1. *ρ, when restricted to $\mathcal{C}$, is continuous at G;*
2. *$\widehat{\rho}$ is $\mathcal{C}$-robust at G.*

In the following, we are going to investigate the continuity of ρ, which finally clarifies whether $\widehat{\rho}$ is robust or not. The following lemma is useful.

Lemma 2 *Let $-\infty < a < b < +\infty$, and G_n, G be increasing functions on $[a,b]$. Suppose $G(z)$ is continuous on $[a,b]$ and $G_n(z) \to G(z)$ for each $z \in [a,b]$, then $G_n(z) \to G(z)$ uniformly on $[a,b]$.*

Proof Because $[a,b]$ is a closed interval and $G(z)$ is continuous on $[a,b]$, for each $\epsilon > 0$, there exists $a = z_1 < \cdots < z_m = b$ such that $\sup_{1\le i\le m-1} |G(z_{i+1}) - G(z_i)| < \frac{\epsilon}{2}$. On the other hand, there exists N such that when $n \ge N$, $|G_n(z_i) - G(z_i)| < \frac{\epsilon}{2}$ for every $z_i, i = 1, \dots m$. Now for any $z \in [a,b]$ there exist z_i, z_{i+1} such that $z_i \le z \le z_{i+1}$. Thus

$$\begin{aligned} G_n(z) - G(z) &\le G_n(z_{i+1}) - G(z_i) \\ &= G_n(z_{i+1}) - G(z_{i+1}) + G(z_{i+1}) - G(z_i) \\ &< \epsilon \end{aligned}$$

when $n \ge N$. Similarly, we have $G_n(z) - G(z) > -\epsilon$ when $n \ge N$. Therefore, $G_n(z) \to G(z)$ on $[a,b]$ uniformly. ∎

The following lemma shows that any risk measure ρ in (21) is consistent on $\mathcal{Q}_c^\infty$, i.e., $\mathcal{Q}_c^\infty \subset \mathcal{Q}_c^\rho$.

Lemma 3 *Let ρ be given in* (21)*. Then ρ is consistent at any $G \in \mathcal{Q}_c$ which is bounded from below. In particular, $\mathcal{Q}_c^\infty \subset \mathcal{Q}_c^\rho$.*

Proof Let $G \in \mathcal{Q}_c$ which is bounded from below, and $X_1, X_2, \dots$ be its samples. By Glivenko-Cantelli theorem, $G^{\text{emp}}_{\mathbf{X}^n}(z) \to G(z), 0 < z < 1$ almost surely. Furthermore, $\inf_{i=1,\dots,n} X_i \to \text{essinf} X_1$ almost surely, which shows that $G^{\text{emp}}_{\mathbf{X}^n}(0+) \to G(0+)$. Thus if we extend $G^{\text{emp}}_{\mathbf{X}^n}$ and G from $(0,1)$ to $[0,1)$ by setting $G^{\text{emp}}_{\mathbf{X}^n}(0) := G^{\text{emp}}_{\mathbf{X}^n}(0+)$ and $G(0) := G(0+)$, then $G(\cdot)$ is continuous on $[0,1)$ and $G^{\text{emp}}_{\mathbf{X}^n}(z) \to G(z), 0 \le z < 1$ almost surely.

In the following, for each fixed ω, let $G_n := G^{\text{emp}}_{\mathbf{X}^n}$. For simplicity, we work with $U(\cdot) := -\rho(\cdot)$. We want to show that $U(G_n(\cdot)) \to U(G(\cdot))$. On the one hand,

$$\begin{aligned} \limsup_{n\to\infty} U(G_n(\cdot)) &\le \inf_{m\in \text{dom}(v)} \left[\limsup_{n\to\infty} \int_{(0,1)} (G_n \wedge 0)(z) m(dz) + v(m) \right] \\ &\le \inf_{m\in \text{dom}(v)} \left[\int_{(0,1)} (G(z) \wedge 0) m(dz) + v(m) \right] \\ &= U(G(\cdot)), \end{aligned}$$

where the second inequality is due to Fatou's lemma. On the other hand, for each $\eta < 1$, by Lemma 2, $G_n(z) \to G(z)$ uniformly for $z \in (0, \eta]$. Thus we have

$$\begin{aligned}
\liminf_{n\to\infty} U(G_n(\cdot)) &= \liminf_{n\to\infty} \inf_{m\in \mathrm{dom}(v)} \left[\int_{(0,1)} (G_n(z) \wedge 0) m(dz) + v(m)\right] \\
&\geq \liminf_{n\to\infty} \inf_{m\in \mathrm{dom}(v)} \left[\int_{(0,1)} (G_n(z \wedge \eta) \wedge 0) m(dz) + v(m)\right] \\
&= \inf_{m\in \mathrm{dom}(v)} \left[\int_{(0,1)} (G(z \wedge \eta) \wedge 0) m(dz) + v(m)\right] \\
&= U(G(\cdot \wedge \eta)),
\end{aligned}$$

where the second equality holds because $G_n(z \wedge \eta) \to G(z \wedge \eta)$ for $z \in (0, 1)$ uniformly. Finally,

$$\begin{aligned}
U(G(\cdot)) \geq U(G(\cdot \wedge \eta)) &= \inf_{m\in \mathrm{dom}(v)} \left[\int_{(0,1)} (G(z \wedge \eta) \wedge 0) m(dz) + v(m)\right] \\
&= \inf_{m\in \mathrm{dom}(v)} \Big[\int_{(0,1)} (G(z) \wedge 0) m(dz) + v(m) \\
&\qquad - \int_{(0,1)} (G(z) \wedge 0 - (G(z \wedge \eta) \wedge 0)) m(dz)\Big] \\
&\geq U(G(\cdot)) - \left[\lim_{z\uparrow 1} G(z) \wedge 0 - G(\eta) \wedge 0\right] \\
&\to U(G(\cdot))
\end{aligned}$$

as $\eta \uparrow 1$. Therefore we conclude $\liminf_{n\to\infty} U(G_n(\cdot)) \geq U(G(\cdot))$. ∎

The following result characterizes risk measures $\rho : \mathcal{Q} \mapsto \mathbb{R}_+$ defined by (21) which are continuous on $\mathcal{Q}$.

Theorem 3 *Let ρ be defined by (21) and $\mathcal{C}$ be any subset of $\mathcal{Q}_c$ such that $\mathcal{C} \supseteq \mathcal{Q}_c^\infty$. The following are equivalent:*

(i) $\rho(\cdot)$, when restricted to $\mathcal{C}$, is continuous at any $G(\cdot) \in \mathcal{C}$.

(ii) There exists $0 < \delta < 1$ such that

$$\sup_{m\in dom(v)} m((0, \delta)) = 0. \tag{27}$$

Furthermore, if (ii) holds, $\mathcal{Q}_c^\rho = \mathcal{Q}_c$.

Proof

$(ii) \Rightarrow (i)$ The proof is analogous to the proof of Lemma 3. Let (ii) hold for some $0 < \delta < 1$. For simplicity, we work with $U(.) \triangleq -\rho(.)$ and first show that U is lower semi-

continuous. For each $\eta < 1$, we have

$$\begin{aligned}
\liminf_{n\to\infty} U(G_n(\cdot)) &= \liminf_{n\to\infty} \inf_{m\in\mathrm{dom}(v)} \left[\int_{[\delta,1)} (G_n(z)\wedge 0)m(dz) + v(m)\right] \\
&\geq \liminf_{n\to\infty} \inf_{m\in\mathrm{dom}(v)} \left[\int_{[\delta,1)} (G_n(z\wedge\eta)\wedge 0)m(dz) + v(m)\right] \\
&= \inf_{m\in\mathrm{dom}(v)} \left[\int_{[\delta,1)} (G(z\wedge\eta)\wedge 0)m(dz) + v(m)\right] \\
&= U(G(\cdot\wedge\eta)),
\end{aligned}$$

where the second equality holds because $G_n(z\wedge\eta)$ converges to $G(z\wedge\eta)$ uniformly for $z\in[\delta,1)$ where Lemma 2 applies. Then, by monotonicity, we have $0 \leq U(G(\cdot)) - U(G(\cdot\wedge\eta))$. Finally,

$$\begin{aligned}
U(G(\cdot)) \geq U(G(\cdot\wedge\eta)) &= \inf_{m\in\mathrm{dom}(v)} \left[\int_{[\delta,1)} (G(z\wedge\eta)\wedge 0)m(dz) + v(m)\right] \\
&= \inf_{m\in\mathrm{dom}(v)} \Big[\int_{[\delta,1)} (G(z)\wedge 0)m(dz) + v(m) \\
&\qquad - \int_{(0,1)} (G(z)\wedge 0 - (G(z\wedge\eta)\wedge 0))m(dz)\Big] \\
&\leq U(G(\cdot)) - \left[\lim_{z\uparrow 1} G(z)\wedge 0 - G(\eta)\wedge 0\right] \\
&\to U(G(\cdot))
\end{aligned}$$

as $\eta\uparrow 1$. Thus we have $\liminf_{n\to\infty} U(G_n(\cdot)) \geq U(G(\cdot))$, and can conclude that U is lower semi-continuous.

Next, we show that U is also upper-semi-continuous.

$$\begin{aligned}
\limsup_{n\to\infty} U(G_n(\cdot)) &\leq \inf_{m\in\mathrm{dom}(v)} \left[\limsup_{n\to\infty} \int_{(0,1)} (G_n\wedge 0)(z)m(dz) + v(m)\right] \\
&\leq \inf_{m\in\mathrm{dom}(v)} \left[\int_{(0,1)} (G(z)\wedge 0)m(dz) + v(m)\right] \\
&= U(G(\cdot)),
\end{aligned}$$

where the second inequality is due to Fatou's lemma. Together with the lower-semi-continuity, $U(\cdot)$ is continuous at G, and so is ρ.

$(i)\Rightarrow(ii)$ We prove it by contradiction. If (ii) is not true, there exists $\delta_n > 0$, and $m_n \in \mathrm{dom}(v)$, such that $\delta_n \to 0$ and $m_n((0,\delta_n)) > 0$. Define

$$G_n(z) := \begin{cases} -\beta_n & 0 < z \leq \delta_n, \\ \frac{\beta_n}{\delta_n}(z - 2\delta_n) & \delta_n < z \leq 2\delta_n \\ 0 & 2\delta_n < z < 1, \end{cases} \quad n \geq 1,$$

where

$$\beta_n := \frac{v(m_n)+1}{m_n((0,\delta_n))}, \quad n \geq 1.$$

It is obvious that $G_n \in \mathcal{Q}_c^\infty \subset \mathcal{C}$, $n \geq 1$. Because $\delta_n \downarrow 0$, $G_n(\cdot) \to 0$ in $\mathcal{C}$. On the other hand,

$$U(G_n(\cdot)) \leq \int_{(0,1)} G_n(z) m_n(dz) + v(m_n) \leq -1,$$

and $U(0) = 0$. Thus, $U(\cdot)$ is not continuous at 0 which is a contradiction.

Finally, if (ii) holds, for any $G(\cdot) \in \mathcal{Q}_c$, we have $\rho(G(\cdot)) = \rho(G(\cdot \vee \delta))$. Because $G(\cdot \vee \delta)$ is bounded from below, by Lemma 3, ρ is consistent at $G(\cdot \vee \delta)$, and therefore consistent at $G(\cdot)$. In other words, $\mathcal{Q}_c^\rho = \mathcal{Q}_c$. ∎

Remark 2 Jouini et al. (2006) consider the *Lebesgue property* of a convex monetary risk measure and give an equivalent dual characterization. Actually, the Lebesgue property is a weaker continuity property of ρ compared with the continuity we consider here. To compare these two continuity properties, we mimic the result in Jouini et al. (2006) to give a dual characterization in Appendix A.

Combining Proposition 1 and Theorem 3, we have

Corollary 1 *Let ρ be defined by (21) and $\mathcal{C}$ be any subset of $\mathcal{Q}_c^\rho$ such that $\mathcal{C} \supseteq \mathcal{Q}_c^\infty$. Then the following are equivalent*

1. *$\widehat{\rho}$ is $\mathcal{C}$-robust*
2. *There exists $0 < \delta < 1$ such that*

$$\sup_{m \in dom(v)} m((0, \delta)) = 0. \tag{28}$$

Furthermore, if $\widehat{\rho}$ is $\mathcal{C}$-robust, then $\mathcal{Q}_c^\rho = \mathcal{Q}_c$.

As the following corollary shows, loss-based statistical risk measures do not lead to robust risk estimators.

Corollary 2 *Let ρ be a loss-based statistical risk measure and $\mathcal{C}$ be any subset of $\mathcal{Q}_c^\rho$ such that $\mathcal{C} \supseteq \mathcal{Q}_c^\infty$. Then $\widehat{\rho}$ is not $\mathcal{C}$-robust.*

Proof By Theorem 2, ρ can be represented as (7) where the penalty function v satisfies (6). As a result, there exists a $\phi \in \Psi((0,1)) \cap \text{dom}(v)$ such that $\int_0^1 \phi(z)dz \geq \frac{1}{2}$. Because $\phi(\cdot)$ is decreasing on (0,1), we must have $\int_0^\delta \phi(z)dz > 0$ for any $\delta > 0$. By Corollary 1, $\widehat{\rho}$ is not $\mathcal{C}$-robust. ∎

2.3 Statistical robustness of risk estimators

Corollary 1 gives a sufficient and necessary condition for a loss-based risk measure ρ represented by equation (21) to be continuous and therefore lead to a robust empirical risk estimator. In particular, loss-based statistical risk measures lead to non robust risk estimators. In the following, we provide one way to robustify risk

estimator computed from loss-based statistical risk measures. Fixed some $\delta \in (0,1)$, for any loss-based statistical risk measure ρ, consider its δ-*truncation*

$$\rho_\delta(G(\cdot)) := \rho(G(\cdot \vee \delta)), \quad G \in \mathcal{Q}. \tag{29}$$

Note that we could also consider risk measures given by representation (21) and propose a modified version of this larger class of risk measures leading to robust risk estimators. The restriction to loss-based statistical risk measures is only made to illustrate further the conflict between robustness and convexity.

From the representation (7), we can find the representation of ρ_δ, which in turns shows that $\widehat{\rho}_\delta$ is robust. Define the map $\pi : \Psi \mapsto \mathcal{M}((0,1))$ which associates to any density function ϕ in Ψ a measure m defined by

$$m(dz) := \begin{cases} \phi(z)dz, & \delta < z < 1, \\ \left(\int_{(0,\delta]} \phi(t)dt\right)\delta_z, & z = \delta, \\ 0, & 0 < z < \delta, \end{cases}$$

where δ_z is the Dirac measure at z. Noticing that π is not a bijective map leads to an additional definition since the penalty function $v(m)$ for $m \in \pi(\Psi((0,1))) := \{m \mid m = \pi(\phi) \text{ for some } \phi \in \Psi((0,1))\}$ cannot be derived uniquely from $v(\phi)$ where $m = \pi(\phi)$. Therefore by denoting $\pi^{-1}(m) = \{\phi \in \Psi((0,1)) \mid \pi(\phi) = m\}$, we define for $m \in \pi(\Psi)$,

$$v_\delta(m) := \inf_{\phi \in \Psi \cap \pi^{-1}(m)} v(\phi).$$

From the representation (7), we have

$$\begin{aligned} \rho_\delta(G) &= -\inf_{\phi\in\Psi}\left\{(G(\delta)\wedge 0)\int_{(0,\delta]}\phi(z)dz + \int_{(\delta,1)}(G(z)\wedge 0)\phi(z)dz + v(\phi)\right\} \\ &= -\inf_{m\in\pi(\Psi)}\left\{\int_{(0,1)}(G(z)\wedge 0)m(dz) + v_\delta(m)\right\}. \end{aligned} \tag{30}$$

Finally, from (6), it is easy to see that v_δ satisfies (22).

From the representation (30), it is immediate to see that the δ-truncation ρ_δ is no longer convex since measures $m \in \pi(\Psi((0,1)))$ have a point mass at δ and therefore do not admit a density on $(0,1)$. On the other hand, it is also straightforward to see that $\widehat{\rho}_\delta$ is $\mathcal{Q}_c$-robust because each $m \in \pi(\Psi((0,1)))$ satisfies $m((0,\delta)) = 0$.

Example 6 The δ-truncation of the spectral tail measure (12) is given by

$$\rho_\delta(G) = \int_\delta^1 (G(z)\wedge 0)\phi(z)dz + G(\delta)\int_0^\delta \phi(z)dz, \quad G \in \mathcal{Q}. \tag{31}$$

Example 7 The δ-truncation of the loss certainty equivalent is given by

$$\rho_\delta(G(\cdot)) = u^{-1}\left(\int_0^1 u(|G(t\vee\delta)\wedge 0|)dt\right), \quad G \in \mathcal{Q}. \tag{32}$$

In particular, the δ-truncation of the L^p risk measure is given by

$$\rho_\delta(G) = \left[\int_0^1 |G(z \vee \delta) \wedge 0|^p dz\right]^{\frac{1}{p}}, \quad G \in \mathcal{Q}, \tag{33}$$

and its representation is expressed as

$$\rho_\delta(G) = -\inf_{m \in \pi(\Psi^q((0,1)))} \int_0^1 (G(z) \wedge 0) m(dz), \quad G \in \mathcal{Q}. \tag{34}$$

The δ-truncation of the entropic risk measure is given by

$$\rho_\delta(G) = \frac{1}{\beta} \log\left(\int_{(0,1)} e^{-\beta\, G(z \vee \delta) \wedge 0}\, dz\right), \quad G \in \mathcal{Q}, \tag{35}$$

and its representation is expressed as

$$\rho_\delta(G) = -\inf_{m \in \pi(\Phi((0,1)))} \left\{\int_{(0,1)} (G(z) \wedge 0) m(dz) + v_\delta(m)\right\}, \quad G \in \mathcal{Q} \tag{36}$$

where for each $m \in \pi(\Phi((0,1)))$

$$\begin{aligned} v_\delta(m) &= \inf_{\phi \in \Phi \cap \pi^{-1}(m)} v(\phi) \\ &= \frac{1}{\beta} \int_{(\delta,1)} m'(z) \log\left(m'(z)\right)\, dz + \frac{1}{\beta} \lim_{z \downarrow \delta} \left(m'(z) \log\left(m'(z)\right)\, z\right) - \inf_{\phi \in \Phi((0,1))} \int_0^1 \phi(z) \ln \phi(z) dz. \end{aligned}$$

Here $m'(z), \delta < z < 1$ denotes the density of m on $(\delta, 1)$ w.r.t to Lebesgue measure. This density is well-defined because $m \in \pi(\Phi((0,1)))$.

Remark 3 In the case of the Expected Shortfall, Cont et al. (2010) have proposed another truncation for ES_α

$$\mathrm{ES}_{\delta,\alpha}(G) = \frac{1}{\alpha - \delta} \int_{(\delta,\alpha)} G(z)\, dz.$$

Hence, applying this idea to the larger class of loss-based statistical risk measures would lead to defining the following $\delta-$truncation

$$\widetilde{\rho}_\delta(G) = -\inf_{\widetilde{\phi} \in \widetilde{\pi}(\Psi)} \left\{\int_{(0,1)} (G(z) \wedge 0) \widetilde{\phi}(z) dz + \widetilde{v}_\delta(\widetilde{\phi})\right\},$$

where the map $\widetilde{\pi} : \Psi \to \mathcal{M}((0,1))$ associates to any function $\phi \in \Psi$ another function $\widetilde{\phi} \in \mathcal{M}((0,1))$ defined by

$$\widetilde{\phi}(z) := \frac{\phi(z) \mathbf{1}_{(\delta,1)}(z)}{\int_{(\delta,1)} \phi(z) dz} \notin \Psi,$$

and where the penalty function is given by

$$\widetilde{v}_\delta(\widetilde{\phi}) := \inf_{\phi \in \Psi \cap \widetilde{\pi}^{-1}(\widetilde{\phi})} v(\phi).$$

In this context, such truncations $\widetilde{\rho}_\delta$ are less tractable than ρ_δ since they cannot be computed from the initial risk measure ρ as in (29).

3 Sensitivity analysis of risk estimators

In the previous section, we have studied the robustness of risk estimators in a qualitative sense. One may argue that the above results rely on the choice of a topology (weak topology) together with a distance (Lévy-Prokhorov distance) on the space of probability distributions. As illustrated in Appendix A, the choice of a weaker topology could lead to different robustness properties. Nonetheless, as noted in Huber (1981), the choice of the weak continuity to study robustness is natural in statistics. To further illustrate this statement in this section, we study sensitivity properties of risk estimators without relying on any topology and show that it leads to the same conclusions as before, singling out the same class of risk estimators as non-robust.

This may be done by quantifying the sensitivity of risk estimators using *influence functions* (Cont et al., 2010; Hampel, 1974). Fix a risk estimator $\widehat{\rho}$ for which the estimation is based on applying the risk measure ρ to empirical quantile functions. Then, its sensitivity function at the quantile function G of a distribution F, in the direction of the Dirac mass at z is equal to

$$S(z;G) \triangleq \lim_{\epsilon \to 0^+} \frac{\rho(\epsilon\delta_z + (1-\epsilon)F) - \rho(F)}{\epsilon},$$

for any $z \in \mathbb{R}$ such that the limit exists. Note that $S(z;G)$ is nothing but the directional derivative of the risk measure ρ at F in the direction $\delta_z \in \mathcal{D}$. $S(z,G)$ measures the sensitivity of the risk estimator based on a large sample to the addition of a new data point. It has already been used in the field of risk measure by Heyde et al. (2010); Cont et al. (2010). In the latter, the authors consider different estimation methods, using both empirical and parametric distributions. In that case, the definition of the sensitivity function should be considered with more attention since the risk measure ρ would have to be replaced with an *effective risk measure* incorporating both the choice of the risk measure and the estimation method as explained in Cont et al. (2010).

3.1 Unbounded sensitivity functions

In this section, we compute the sensitivity function of the loss certainty equivalent. We find that this risk measure has unbounded sensitivity function which is consistent with our findings of Section 2. Note that, unlike the setting of Section 2, this result makes no reference to any topology on the set of loss distributions.

Proposition 2 *The sensitivity function of the loss certainty equivalent* (13) *is given by*

$$S(z;G) = \frac{u(|z \wedge 0|) - u(\rho(G))}{u'(\rho(G))}. \tag{37}$$

Proof By denoting $G_\epsilon(\cdot)$ the quantile function corresponding to the distribution function $F_\epsilon(\cdot) = (1-\epsilon)F(\cdot) + \epsilon\delta_z(\cdot)$, we have

$$\begin{aligned}\rho(G_\epsilon) &= u^{-1}\left(\int_{\mathbb{R}} u(|x \wedge 0|)dF_\epsilon(x)\right) \\ &= u^{-1}\left((1-\epsilon)\int_{\mathbb{R}} u(|x \wedge 0|)dF(x) + \epsilon u(|z \wedge 0|)\right).\end{aligned}$$

Now, simple calculus leads to (37). ∎

From Proposition 2, if $\lim_{x\to\infty} u(x) = \infty$, $\lim_{z\downarrow-\infty} S(z;G) = +\infty$, showing that the sensitivity function is unbounded. In particular, for the L^p and entropic risk measures, the sensitivity functions are unbounded.

3.2 Boundedness of sensitivity functions for robust risk estimators

In this section, we compute the sensitivity functions of the $\delta-$truncated versions of the loss certainty equivalents. These truncated versions were introduced in Section 2.3 in order to obtain robust risk estimators. The conclusion of the following proposition is that by truncating these risk measures, their sensitivity functions become bounded. Therefore, the robustness properties of risk estimators derived in Section 2 are consistent with the sensitivity functions computed in this section.

Proposition 3 *Consider the δ-truncation of the certainty equivalent* (32) *with $0 < \delta < 1$. Assume $F(z) < 1$ and $G(\cdot)$ is differentiable at δ. Then the sensitivity $S(z;G)$ can be computed as follows:*

(i) *When $G(\delta) > 0$,*

$$S(z;G) = 0. \tag{38}$$

(ii) *When $G(\delta) = 0$,*

$$S(z;G) = \frac{1}{u'(\rho_\delta(G))}\begin{cases}0, & z \geq G(\delta), \\ \delta\,(1-\delta)\,G'(\delta), & z < G(\delta).\end{cases} \tag{39}$$

(iii) *When $G(\delta) < 0$,*

$$S(z;G) = \frac{1}{u'(\rho_\delta(G))}\left[-u(\rho_\delta(G)) + \begin{cases}u(|z \wedge 0|) - \delta^2 u'(|G(\delta)|)G'(\delta), & z > G(\delta), \\ u(|z \wedge 0|), & z = G(\delta), \\ u(|G(\delta)|) + \delta(1-\delta)u'(|G(\delta)|)G'(\delta), & z < G(\delta).\end{cases}\right] \tag{40}$$

Proof First let us recall some properties of left-continuous inverse functions. For any distribution function $\widetilde{F}$, denote by $\widetilde{F}^{-1}$ its left-continuous inverse. Then for any $t \in [0,1], x \in \mathbb{R}$,

$$\widetilde{F}(x) \geq t \Longleftrightarrow x \geq \widetilde{F}^{-1}(t), \quad \widetilde{F}(x) < t \Longleftrightarrow x < \widetilde{F}^{-1}(t), \tag{41}$$

and thus

$$\widetilde{F}(x-) < t \leq \widetilde{F}(x) \Longleftrightarrow x = \widetilde{F}^{-1}(t). \tag{42}$$

In the following, we denote by $F(\cdot)$ the distribution function associated with $G(\cdot)$. By assumption, $G(\cdot)$ is differentiable and thus continuous at δ. We claim that

$$F(z-) \leq \delta \leq F(z) \Longleftrightarrow z = G(\delta), \quad F(z) > \delta \Longleftrightarrow z > G(\delta), \quad F(z) < \delta \Longleftrightarrow z < G(\delta). \tag{43}$$

Indeed, from (41), we deduce that $F(z) < \delta$ if only if $z < G(\delta)$. From (42), we deduce that if $z = G(\delta)$, then $F(z-) \leq \delta \leq F(z)$. Thus, to prove (43), we only need to show that if $F(z-) \leq \delta \leq F(z)$, $z = G(\delta)$. Suppose $F(z-) \leq \delta \leq F(z)$. On the one hand, from (41), $z \geq G(\delta)$. On the other hand, for any $\varepsilon > 0$ small enough, $F(z-\varepsilon) < \delta + \varepsilon$, which again by (41) leads to $z - \epsilon < G(\delta + \varepsilon)$. Letting $\varepsilon \downarrow 0$, by the continuity of $G(\cdot)$ at δ, we conclude that $z \leq G(\delta)$.

Denote by $G_\epsilon(\cdot)$ the quantile function corresponding to the distribution function $F_\epsilon(\cdot) = (1-\epsilon)F(\cdot) + \epsilon\delta_z(\cdot)$. We only need to compute

$$A := \lim_{\epsilon \downarrow 0} \frac{\int_0^1 u(|G_\epsilon(t \vee \delta) \wedge 0|)dt - \int_0^1 u(|G(t \vee \delta) \wedge 0|)dt}{\epsilon},$$

and $S(z; G)$ follows from the chain rule.

Straightforward computation shows that

$$G_\epsilon(t) = \begin{cases} G\left(\frac{t-\epsilon}{1-\epsilon}\right), & t > \epsilon + (1-\epsilon)F(z), \\ G\left(\frac{t}{1-\epsilon}\right), & t \leq (1-\epsilon)F(z-), \\ z, & (1-\epsilon)F(z-) < t \leq \epsilon + (1-\epsilon)F(z). \end{cases}$$

Thus

$$\begin{aligned}
&\int_\delta^1 u(|G_\epsilon(t) \wedge 0|)dt \\
=&\int_\delta^1 u(|G\left(\frac{t-\epsilon}{1-\epsilon}\right) \wedge 0|)\mathbf{1}_{\{t>\epsilon+(1-\epsilon)F(z)\}}dt + \int_\delta^1 u(|G\left(\frac{t}{1-\epsilon}\right) \wedge 0|)\mathbf{1}_{\{t\leq(1-\epsilon)F(z-)\}}dt \\
&\quad + \int_\delta^1 u(|z \wedge 0|)\mathbf{1}_{\{(1-\epsilon)F(z-)<t\leq\epsilon+(1-\epsilon)F(z)\}}dt \\
=&(1-\epsilon)\int_{\frac{\delta-\epsilon}{1-\epsilon}}^1 u(|G\left(t\right) \wedge 0|)\mathbf{1}_{\{t>F(z)\}}dt + (1-\epsilon)\int_{\frac{\delta}{1-\epsilon}}^1 u(|G\left(t\right) \wedge 0|)\mathbf{1}_{\{t\leq F(z-)\}}dt \\
&\quad + \int_\delta^1 u(|z \wedge 0|)\mathbf{1}_{\{(1-\epsilon)F(z-)<t\leq\epsilon+(1-\epsilon)F(z)\}}dt \\
=&(1-\epsilon)\int_\delta^1 u(|G\left(t\right) \wedge 0|)dt + (1-\epsilon)\int_{\frac{\delta-\epsilon}{1-\epsilon}}^\delta u(|G\left(t\right) \wedge 0|)\mathbf{1}_{\{t>F(z)\}}dt \\
&\quad - (1-\epsilon)\int_\delta^{\frac{\delta}{1-\epsilon}} u(|G\left(t\right) \wedge 0|)\mathbf{1}_{\{t\leq F(z-)\}}dt - (1-\epsilon)\int_\delta^1 u(|G\left(t\right) \wedge 0|)\mathbf{1}_{\{F(z-)<t\leq F(z)\}}dt \\
&\quad + \int_\delta^1 u(|z \wedge 0|)\mathbf{1}_{\{(1-\epsilon)F(z-)<t\leq\epsilon+(1-\epsilon)F(z)\}}dt.
\end{aligned}$$

It is easy to compute that

$$\int_{\frac{\delta-\epsilon}{1-\epsilon}}^{\delta} u(|G(t)\wedge 0|)\mathbf{1}_{\{t>F(z)\}}dt=\begin{cases}0, & F(z)\geq\delta,\\ (1-\delta)u(|G(\delta)\wedge 0|)\epsilon+o(\epsilon), & F(z)<\delta,\end{cases}$$
$$\int_{\delta}^{\frac{\delta}{1-\epsilon}} u(|G(t)\wedge 0|)\mathbf{1}_{\{t\leq F(z-)\}}dt=\begin{cases}\delta u(|G(\delta)\wedge 0|)\epsilon+o(\epsilon), & F(z-)>\delta,\\ 0, & F(z-)\leq\delta.\end{cases}$$

Noticing (42), we can compute that

$$\int_{\delta}^{1} u(|G(t)\wedge 0|)\mathbf{1}_{\{F(z-)<t\leq F(z)\}}dt=\begin{cases}0, & F(z)<\delta,\\ (F(z)-\delta)u(|z\wedge 0|), & F(z-)\leq\delta\leq F(z),\\ (F(z)-F(z-))u(|z\wedge 0|), & \delta<F(z-).\end{cases}$$

Similarly, we can compute that

$$\int_{\delta}^{1} u(|z\wedge 0|)\mathbf{1}_{\{(1-\epsilon)F(z-)<t\leq\epsilon+(1-\epsilon)F(z)\}}dt=\begin{cases}o(\epsilon), & F(z)<\delta,\\ \left[(1-\epsilon)(F(z)-\delta)+\epsilon(1-\delta)\right]u(|z\wedge 0|), & F(z-)\leq\delta\leq F(z),\\ \left[(1-\epsilon)(F(z)-F(z-))+\epsilon\right]u(|z\wedge 0|), & \delta<F(z-).\end{cases}$$

Therefore

$$\int_{\delta}^{1} u(|G_\epsilon(t)\wedge 0|)dt=(1-\epsilon)\int_{\delta}^{1} u(|G(t)\wedge 0|)dt+\begin{cases}\left[u(|z\wedge 0|)-\delta u(|G(\delta)\wedge 0|)\right]\epsilon+o(\epsilon), & \delta<F(z-),\\ (1-\delta)u(|z\wedge 0|)\epsilon+o(\epsilon), & F(z-)\leq\delta\leq F(z),\\ (1-\delta)u(|G(\delta)\wedge 0|)\epsilon+o(\epsilon), & \delta>F(z).\end{cases}$$

On the other hand

$$G_\epsilon(\delta)\wedge 0-G(\delta)\wedge 0=\begin{cases}G\left(\frac{\delta-\epsilon}{1-\epsilon}\right)\wedge 0-G(\delta)\wedge 0, & \delta>\epsilon+(1-\epsilon)F(z),\\ G\left(\frac{\delta}{1-\epsilon}\right)\wedge 0-G(\delta)\wedge 0, & \delta\leq(1-\epsilon)F(z-),\\ z\wedge 0-G(\delta)\wedge 0, & (1-\epsilon)F(z-)<\delta\leq\epsilon+(1-\epsilon)F(z).\end{cases}$$

We discuss case by case.

1. $G(\delta)>0$. Because $G(\cdot)$ is continuous at δ, it is easy to show that $G_\epsilon(\delta)\wedge 0-G(\delta)\wedge 0=0$ when ϵ is sufficiently small.

2. $G(\delta)=0$. In this case, we have

$$G_\epsilon(\delta)\wedge 0-G(\delta)\wedge 0=\begin{cases}G\left(\frac{\delta-\epsilon}{1-\epsilon}\right), & \delta>\epsilon+(1-\epsilon)F(z),\\ 0, & \delta\leq(1-\epsilon)F(z-),\\ z\wedge 0, & (1-\epsilon)F(z-)<\delta\leq\epsilon+(1-\epsilon)F(z)\end{cases}$$

when ϵ is sufficiently small. Recalling (43), we conclude that

$$G_\epsilon(\delta)\wedge 0-G(\delta)\wedge 0=\begin{cases}0, & \delta<F(z-),\\ 0, & F(z-)\leq\delta\leq F(z),\\ -(1-\delta)G'(\delta)\epsilon+o(\epsilon), & \delta>F(z).\end{cases}$$

3. $G(\delta) < 0$. In this case, we have

$$G_\epsilon(\delta) \wedge 0 - G(\delta) \wedge 0 = \begin{cases} G\left(\frac{\delta-\epsilon}{1-\epsilon}\right) - G(\delta), & \delta > \epsilon + (1-\epsilon)F(z), \\ G\left(\frac{\delta}{1-\epsilon}\right) - G(\delta), & \delta \le (1-\epsilon)F(z-), \\ z - G(\delta), & (1-\epsilon)F(z-) < \delta \le \epsilon + (1-\epsilon)F(z). \end{cases}$$

when ϵ is small enough, leading to

$$G_\epsilon(\delta) \wedge 0 - G(\delta) \wedge 0 = \begin{cases} \delta G'(\delta)\epsilon + o(\epsilon), & \delta < F(z-), \\ 0, & F(z-) \le \delta \le F(z), \\ -(1-\delta)G'(\delta)\epsilon + o(\epsilon), & \delta > F(z). \end{cases}$$

Notice that

$$\int_0^1 u(|G_\epsilon(t \vee \delta) \wedge 0|)dt - \int_0^1 u(|G(t \vee \delta) \wedge 0|)dt = \int_\delta^1 u(|G_\epsilon(t) \wedge 0|)dt - \int_\delta^1 u(|G(t) \wedge 0|)dt + \delta \left(u(|G_\epsilon(\delta) \wedge 0|) - u(|G(\delta) \wedge 0|)\right).$$

Then, when $G(\delta) > 0$,

$$\begin{aligned} &\lim_{\epsilon \downarrow 0} \frac{\int_0^1 u(|G_\epsilon(t \vee \delta) \wedge 0|)dt - \int_0^1 u(|G(t \vee \delta) \wedge 0|)dt}{\epsilon} \\ =&\lim_{\epsilon \downarrow 0} \frac{\int_\delta^1 u(|G_\epsilon(t) \wedge 0|)dt - \int_\delta^1 u(|G(t) \wedge 0|)dt + \delta \left(u(|G_\epsilon(\delta) \wedge 0|) - u(|G(\delta) \wedge 0|)\right)}{\epsilon} \\ =&\lim_{\epsilon \downarrow 0} \frac{\int_\delta^1 u(|G_\epsilon(t) \wedge 0|)dt - \int_\delta^1 u(|G(t) \wedge 0|)dt - \delta u'(|G(\delta) \wedge 0|) \left(G_\epsilon(\delta) \wedge 0 - G(\delta) \wedge 0\right)}{\epsilon} \\ =&0. \end{aligned}$$

Similarly, when $G(\delta) = 0$,

$$\lim_{\epsilon \downarrow 0} \frac{\int_0^1 u(|G_\epsilon(t \vee \delta) \wedge 0|)dt - \int_0^1 u(|G(t \vee \delta) \wedge 0|)dt}{\epsilon} = \begin{cases} 0, & \delta < F(z-), \\ 0, & F(z-) \le \delta \le F(z), \\ u'(0)\delta(1-\delta)G'(\delta), & \delta > F(z). \end{cases}$$

When $G(\delta) < 0$,

$$\begin{aligned} &\lim_{\epsilon \downarrow 0} \frac{\int_0^1 u(|G_\epsilon(t \vee \delta) \wedge 0|)dt - \int_0^1 u(|G(t \vee \delta) \wedge 0|)dt}{\epsilon} \\ =&-\int_\delta^1 u(|G(t) \wedge 0|)dt + \begin{cases} u(|z \wedge 0|) - \delta u(|G(\delta)|) - \delta^2 u'(|G(\delta)|)G'(\delta), & \delta < F(z-), \\ (1-\delta)u(|z \wedge 0|), & F(z-) \le \delta \le F(z), \\ (1-\delta)u(|G(\delta)|) + \delta(1-\delta)u'(|G(\delta)|)G'(\delta), & \delta > F(z). \end{cases} \\ =&-\int_0^1 u(|G(t \vee \delta) \wedge 0|)dt + \begin{cases} u(|z \wedge 0|) - \delta^2 u'(|G(\delta)|)G'(\delta), & \delta < F(z-), \\ u(|z \wedge 0|), & F(z-) \le \delta \le F(z), \\ u(|G(\delta)|) + \delta(1-\delta)u'(|G(\delta)|)G'(\delta), & \delta > F(z). \end{cases} \end{aligned}$$

Finally, applying chain rule and (43), we immediately have (38)-(40). ∎

From Proposition 3, it is easy to compute that

$$\sup_{z\in\mathbb{R}} S(z;G) \leq \frac{1}{u'(\rho_\delta(G))} \times \begin{cases} 0, & G(\delta) > 0, \\ \delta(1-\delta)G'(\delta), & G(\delta) = 0, \\ -u(\rho_\delta(G)) + u(|G(\delta)|) + \delta(1-\delta)u'(|G(\delta)|)G'(\delta), & G(\delta) < 0. \end{cases}$$

Thus the truncated loss certainty equivalent has bounded sensitivity functions, which is consistent with the robustness properties of risk estimators derived in Section 2.

Remark 4 If F has a continuous positive density f in the neighborhood of $G(\delta)$, then $G(\cdot)$ is differentiable at δ and

$$G'(\delta) = \frac{1}{f(G(\delta))}.$$

Moreover, denoting by $\mathrm{SVaR}_\delta(z;G)$ the sensitivity function of VaR_δ at G, which, from the proof of Proposition 3, is

$$\mathrm{SVaR}_\delta(z;G) = \begin{cases} \frac{\delta}{f(G(\delta))}, & z > G(\delta), \\ 0, & z = G(\delta), \\ -\frac{1-\delta}{f(G(\delta))}, & z < G(\delta), \end{cases}$$

we can rewrite, for $G(\delta) < 0$, the sensitivity of the $\delta-$truncation of the loss certainty equivalent as:

$$S(z;G) = \frac{1}{u'(\rho_\delta(G))}\left[-u(\rho_\delta(G)) + u(|(z \vee G(\delta)) \wedge 0|) - \delta u'(|G(\delta)|)\mathrm{SVaR}_\delta(z;G)\right].$$

A Lebesgue Continuity

Following Jouini et al. (2006), we consider a weaker continuity property for ρ:

Definition 3 (Lebesgue continuity) Let ρ be defined in (21). ρ is *Lebesgue continuous* at $G \in \mathcal{Q}$ if for any $G_n \in \mathcal{Q}$ such that G_n is uniformly bounded and $G_n \to G$, $\rho(G_n) \to \rho(G)$.

The following result gives the dual characterization of the Lebesgue continuity and shows how Theorem 3 is modified when weak continuity is replaced by Lebesgue continuity.

Theorem 4 *Let ρ be defined as in* (21)*. Then, the following are equivalent*

(i) $\rho(\cdot)$ is Lebesgue continuous at any $G \in \mathcal{Q}_c^\infty$.

(ii) For each $c > 0$,

$$\lim_{\delta\downarrow 0} \sup_{v(m)\leq c} m((0,\delta)) = 0.$$

Proof We work with $U(\cdot) := -\rho(\cdot)$.

$(i) \Rightarrow (ii)$ If (ii) does not hold, there exists $c > 0$, $\epsilon_0 > 0$, $\delta_n \downarrow 0$, and $m_n \in \mathcal{P}((0,1))$ with $v(m_n) \le c$, such that $m_n((0,\delta_n)) \ge \epsilon_0$. Define

$$G_n(z) := \begin{cases} -\beta, & 0 < z \le \delta_n, \\ \frac{\beta}{\delta_n}(z - 2\delta_n), & \delta_n < z \le 2\delta_n, \\ 0, & 2\delta_n < z < 1, \end{cases} \quad n \ge 1,$$

where $\beta := \frac{2c}{\epsilon_0}$. Then we have

$$\begin{aligned} U(G_n(\cdot)) &\le \int_{(0,1)} (G_n(z) \wedge 0) m_n(dz) + v(m_n) \\ &\le -2c + c = -c < 0. \end{aligned}$$

On the other hand, $G_n(\cdot)$ is uniformly bounded, and because $\delta_n \downarrow 0$, $G_n(\cdot) \to G(\cdot) \equiv 0$. Because $U(0) = 0$, $U(\cdot)$ is not Lebesgue continuous at 0, which is a contradiction.

$(ii) \Rightarrow (i)$ Suppose $G, G_n \in \mathcal{Q}_c^\infty$ are uniformly bounded and $G_n \to G$. On the one hand

$$\begin{aligned} \limsup_{n\to\infty} U(G_n(\cdot)) &\le \inf_{m\in\mathcal{P}((0,1))} \left\{ \limsup_{n\to\infty} \int_{(0,1)} (G_n(z) \wedge 0) m(dz) + v(m) \right\} \\ &= \inf_{m\in\mathcal{P}((0,1))} \left\{ \int_{(0,1)} (G(z) \wedge 0) m(dz) + v(m) \right\} \\ &= U(G(\cdot)), \end{aligned}$$

where the first equality is due to bounded convergence theorem.

On the other hand, let $M > 0$ be a uniform bound of G_n, i.e., $\sup_{n\ge1, z\in(0,1)} |G_n(z)| \le M$. For each fixed $0 < \epsilon < 1$, and $n \ge 1$, we can find $m_n \in \mathrm{dom}(v)$ such that

$$U(G_n(\cdot)) \ge \int_{(0,1)} (G_n(z) \wedge 0) m_n(dz) + v(m_n) - \epsilon. \tag{44}$$

Let $c_n := v(m_n) < \infty$, then

$$\begin{aligned} c_n &\le U(G_n(\cdot)) - \int_{(0,1)} (G_n(z) \wedge 0) m_n(dz) + \epsilon \\ &\le 2M + 1 =: c. \end{aligned}$$

Thus we have

$$U(G_n(\cdot)) \ge \inf_{v(m)\le c} \left\{ \int_{(0,1)} (G_n(z) \wedge 0) m(dz) + v(m) \right\} - \epsilon, \quad n \ge 1.$$

By monotonicity of U, we can find $0 < \eta < 1$ such that $U(G(\cdot \wedge \eta)) \ge U(G(\cdot)) - \epsilon$. Because $\lim_{\delta\downarrow0} \sup_{v(m)\le c} m((0,\delta)) = 0$, we can find $0 < \delta < \eta$ such that $\sup_{v(m)\le c} m((0,\delta)) < \epsilon/M$. By Lemma 2, $G_n(z) \to G(z)$ uniformly on $[\delta, \eta]$.

Thus, there exists N such that for any $n \geq N$, $\sup_{z\in[\delta,1)} |G_n(z\wedge\eta) - G(z\wedge\eta)| < \epsilon$. Now, for each $n \geq N$, we have

$$\begin{aligned}
U(G_n(\cdot)) &\geq \inf_{v(m)\leq c} \left\{ \int_{(0,1)} (G_n(z\wedge\eta)\wedge 0) m(dz) + v(m) \right\} - \epsilon \\
&= \inf_{v(m)\leq c} \Big\{ \int_{(0,1)} (G(z\wedge\eta)\wedge 0) m(dz) + \int_{(0,\delta)} [(G_n(z\wedge\eta)\wedge 0) - (G(z\wedge\eta)\wedge 0)] m(dz) \\
&\quad + \int_{[\delta,1)} [(G_n(z\wedge\eta)\wedge 0) - (G(z\wedge\eta)\wedge 0)] m(dz) + v(m) \Big\} - \epsilon \\
&\geq \inf_{v(m)\leq c} \Big\{ \int_{(0,1)} (G(z\wedge\eta)\wedge 0) m(dz) - 2M \cdot \frac{\epsilon}{M} - \epsilon + v(m) \Big\} - \epsilon \\
&= \inf_{v(m)\leq c} \Big\{ \int_{(0,1)} (G(z\wedge\eta)\wedge 0) m(dz) + v(m) \Big\} - 4\epsilon \\
&\geq U(G(\cdot\wedge\eta)) - 4\epsilon \\
&\geq U(G(\cdot)) - 5\epsilon.
\end{aligned}$$

Thus we have $\liminf_{n\to\infty} U(G_n(\cdot)) \geq U(G(\cdot))$. In summary, U is Lebesgue continuous at G and so is ρ.

■

References

Carlo Acerbi. Spectral measures of risk: a coherent representation of subjective risk aversion. *Journal of Banking and Finance*, 26(7):1505–1518, 2002.

Carlo Acerbi. Coherent measures of risk in everyday market practice. *Quantitative Finance*, 7(4):359–364, 2007.

Philippe Artzner, Freddy Delbaen, Jean-Marc Eber, and David Heath. Coherent Measures of Risk. *Mathematical Finance*, 9(3):203–228, 1999.

Rama Cont, Romain Deguest, and Giacomo Scandolo. Robustness and sensitivity analysis of risk meausurement procedures. *Quantitative Finance*, 10(6):593–606, 2010.

Freddy Delbaen. Coherent risk measures on general probability spaces. *Advances in finance and stochastics. Essays in Honour of Dieter Sondermann*, pages 1–37, 2002.

Freddy Delbaen and Walter Schachermayer. A general version of the fundamental theorem of asset pricing. *Mathematische Annalen*, 300(1):1432–1807, 1994.

Nicole El Karoui and Claudia Ravanelli. Cash subadditive risk measures and interest rate ambiguity. *Mathematical Finance*, 19(4):561–590, 2009.

H. Föllmer and A. Schied. Convex measures of risk and trading constraints. *Finance and Stochastics*, 6(4):429–447, 2002.

H. Föllmer and A. Schied. *Stochastic Finance: An Introduction in Discrete Time*. Walter de Gruyter, 2004.

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

Marco Frittelli and Emanuela Rosazza Gianin. Law invariant convex risk measures. *Advances in Mathematical Economics*, 7:33–46, June 2005.

F.R. Hampel. The influence curve and its role in robust estimation. *Journal of the American Statistical Association*, 69(346):383–393, 1974.

G. H. Hardy, J. E. Littlewood, and G. Pólya. *Inequalities*. Cambridge University Press, Cambridge, 1959.

C.C. Heyde, S.G. Kou, and X.H. Peng. External Risk Measures and Basel Accords. *Working Paper*, 2010.

P.J. Huber. *Robust Statistics*. Wiley, NJ, 1981.

Robert Jarrow. Put option premiums and Coherent Risk Measures. *Mathematical Finance*, 12(2):135–142, 2002.

E. Jouini, W. Schachermayer, and N. Touzi. Law invariant risk measures have the Fatou property. *Advances in mathematical economics*, pages 49–71, 2006.

S. Kusuoka. On law invariant coherent risk measures. *Advances in Mathematical Economics*, 3:83–95, 2001.

A.J. McNeil, R. Frey, P. Embrechts, et al. *Quantitative risk management: concepts, techniques, and tools*. Princeton University Press, Princeton, NJ, 2005.

Yongsheng Song and Jia-An Yan. Risk measures with comonotonic subadditivity or convexity and respecting stochastic orders. *Insurance: Mathematics and Economics*, 45(3):459–465, Dec. 2009.

Jeremy Staum. Excess invariance and shortfall risk measures. *Working Paper*, 2011.