Risks **2014**, *4*, 411-424; doi:10.3390/risks2040411

OPEN ACCESS
risks
ISSN 2227-9091
www.mdpi.com/journal/risks

Article

Measuring Risk When Expected Losses Are Unbounded

Alejandro Balbás [1], Iván Blanco [1] and José Garrido [2]

[1] University Carlos III of Madrid. C/ Madrid, 126. 28903 Getafe (Madrid, Spain);
E-Mails: ivan.blanco@uc3m.es (I.B.)

[2] Concordia University. Department of Mathematics and Statistics. 1455 de Maisonneuve Blvd. W., Montreal, QC H3G 1M8, Canada; E-Mail: jose.garrido@concordia.ca

* Author to whom correspondence should be addressed; E-Mail: alejandro.balbas@uc3m.es
Tel.: +34-91-6249636

External Editor: Mogens Steffensen

Received: 26 May 2014; in revised form: 5 September 2014 / Accepted: 10 September 2014 / Published: 30 September 2014

Abstract: This paper proposes a new method to introduce coherent risk measures for risks with infinite expectation, such as those characterized by some Pareto distributions. Extensions of the conditional value at risk, the weighted conditional value at risk and other examples are given. Actuarial applications are analyzed, such as extensions of the expected value premium principle when expected losses are unbounded.

Keywords: heavy tail; risk measures; representation theorem; applications

---

1. Introduction

Risk measurement is becoming increasingly more important in economics, finance and insurance. Although the standard deviation has many interesting properties as a risk measure in a Gaussian world, asymmetries and heavy tails imply inconsistencies between the standard deviation and second order stochastic dominance (or classical utility functions). This also makes it difficult to interpret the standard deviation in terms of potential capital losses. Several recent approaches have attempted to overcome these drawbacks. In particular, a first line of research deals with the axioms that an index of riskiness must satisfy from a Theory of Economics perspective (Aumann and Serrano [1]), whereas a second line deals with the properties allowing us to interpret a risk measure as potential losses and capital requirements (Artzner *et al.* [2]).

This paper focuses on the Artzner *et al.*'s [2] approach with a significant novelty: we allow for risks generating unbounded expected losses. As will be seen, the inclusion of risks with unbounded expectation (for instance, risks with a Cauchy or a Pareto distribution) presents many mathematical problems when extending the notion of coherence (Artzner *et al.* [2]) or expectation boundedness (Rockafellar *et al.* [3]), and previous literature has addressed this caveat by losing some desirable mathematical properties. For instance, if we use value at risk (VaR) as a risk measure, then we lose continuity and sub-additivity. Though there are risk measures for heavy tailed risks that can recover sub-additivity (or convexity at least, Kupper and Svindland [4]), continuity is still lost.

This paper overcomes the mathematical problems above by extending a given coherent or expectation bounded risk measure to a limited new setting. For instance, despite the fact that the conditional value at risk (CVaR) cannot be continuously extended to the whole space of random risks, we will see that, in fact, it can be continuously extended to some "smaller" spaces containing some risks with infinite expected value. In practical situations, most of the involved risks will have a finite expected value, so we should not find infinitely many candidates with infinite expectation. More likely, there will be just a few, or even only one. Then, instead of extending a coherent risk measure to a "too large set of risks", we look for extensions that apply only if the set of heavy tailed risks is finitely generated.

The outline of the paper is as follows. Section 2 introduces the notation, the framework and the main problem to be addressed: the extension of risk measures, so as to conserve continuity and sub-additivity (or convexity, at least) and simultaneously include risks whose fat tails lead to unbounded expected losses. We summarize the mathematical problems affecting this objective.

Theorem 1 is the main result of Section 3. It states the existence of the required extension if the set of fat tailed risks has a finite generator. Furthermore, Remarks 2 and 3 show how to construct the extended risk measure in a recursive manner.

Section 4 provides illustrative examples and applications. In particular, we extend the CVaR and the weighted CVaR (WCVaR) by "integrating in a coherent manner" these risk measures with the VaR of the heavy tailed risks. We have selected CVaR and WCVaR due to their additional properties, since they are consistent with second order stochastic dominance (Ogryczak and Ruszczynski [5]) and may be optimized by linear programming methods (Mansini *et al.* [6], see also Konno *et al.* [7]). We also summarize some actuarial applications, such as some extensions of the expected value premium principle. Empirical applications based on real-world data are not addressed and could be an interesting subject for future research.

The last section of the paper summarizes the most important conclusions.

2. Preliminaries and Notations

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ composed of the set of "states of the world" Ω, the σ-algebra $\mathcal{F}$ and the probability measure $\mathbb{P}$. Denote by $\mathbb{E}(y)$ the mathematical expectation of every $\mathbb{R}$-valued random variable y defined on Ω. Let $1 \leq p < \infty$ and denote by L^p the Banach space of random variables y on Ω such that $\mathbb{E}\left(|y|^p\right) < \infty$,[1] is endowed with the norm:

$$\|y\|_p = \left(\mathbb{E}\left(|y|^p\right)\right)^{1/p}$$

According to the Riesz representation theorem, L^q is the dual space of L^p, where $q \in (1, \infty]$ is characterized by $1/p + 1/q = 1$, and L^∞ is the space of essentially bounded random variables endowed with the supremum norm.

Let $[0, T]$ be a time interval. From an intuitive point of view, one can interpret that $y \in L^p$ represents the portfolio pay-off at T for some arbitrary investor (finance) or claims at T for some arbitrary insurer (actuarial science). Throughout this paper, y will represent the random wealth at T, although other interpretations would not modify our main conclusions. If:

$$\rho : L^p \longrightarrow \mathbb{R}$$

is a risk measure, then $\rho(y)$ may be understood as the "risk" associated with the wealth y. Let us assume that ρ satisfies a representation theorem in the line of Artzner *et al.* [1] or Rockafellar *et al.* [3]. More precisely, consider the sub-gradient of ρ:

$$\Delta_\rho = \left\{z \in L^q; -\mathbb{E}(yz) \leq \rho(y), \forall y \in L^p\right\} \subset L^q \tag{1}$$

composed of those linear expressions lower than ρ. We assume that Δ_ρ is convex and $weak^*$-compact[2] and that ρ is its envelope, in the sense that:

$$\rho(y) = Max\ \left\{-\mathbb{E}(yz);\ z \in \Delta_\rho\right\} \tag{2}$$

holds for every $y \in L^p$. Furthermore, we assume the existence of $\tilde{E}_\rho \geq 0$, such that:

$$\Delta_\rho \subset \left\{z \in L^q;\ \mathbb{E}(z) = \tilde{E}_\rho\right\} \tag{3}$$

These assumptions are equivalent to the well-known properties of sub-additivity, homogeneity and translation invariance. To sum up, we have:

Assumption 1. The risk measure ρ satisfies the equivalent Conditions a and b below:

(a) The set Δ_ρ given by (1) is convex and $weak^*$-compact, (2) holds for every $y \in L^p$, and (3) holds.

(b) ρ is continuous, sub-additive ($\rho(y_1 + y_2) \leq \rho(y_1) + \rho(y_2)$), homogeneous ($\rho(\lambda y) = \lambda\rho(y)$ if $\lambda \geq 0$) and $\tilde{E}_\rho$-translation invariant ($\rho(y + k) = \rho(y) - \tilde{E}_\rho k$ if $k \in \mathbb{R}$ is zero-variance).

We will not prove the equivalence between Conditions a and b above, as similar results may be found in several papers (see, for instance, Balbás *et al.* [9]).

Assumption 1 is not at all restrictive, since it is satisfied by every expectation bounded risk measure (Rockafellar *et al.* [3]) with $\tilde{E}_\rho = 1$ and by every deviation measure (Rockafellar *et al.* [3]) with $\tilde{E}_\rho = 0$.

---

1 Therefore, $\mathbb{E}\left(|y|^{p'}\right) < \infty$ if $1 \leq p' \leq p$. Recall that $L^{p'} \supset L^p$ if $1 \leq p' \leq p$.

2 See Rudin [8] for further details about $weak^*$-compact sets in Banach spaces.

Examples of expectation bounded risk measures are, amongst many others, the CVaR and the WCVaR. Recall that the VaR of a random variable y with cumulative distribution function equaling F is given by:

$$VaR_{\mu}\left(y\right) := Inf\ \left\{t \in \mathbb{R};\ F\left(t\right) > 1-\mu\right\}$$

and for $y \in L^1$, the CVaR and the WCVaR are given by:

$$CVaR_{\mu}\left(y\right) := \frac{1}{1-\mu}\int_0^{1-\mu} VaR_{1-t}\left(y\right)dt$$

and:

$$WCVaR_{\nu}\left(y\right) := \int_0^1 CVaR_t\left(y\right)d\nu\left(t\right)$$

$0 < \mu < 1$ denoting the level of confidence of VaR and CVaR and ν being a probability measure on the interval $[0, 1]$. Examples of deviation measures are, amongst others, the classical p-deviation:

$$\sigma_p\left(y\right) := \left[\mathbb{E}\left(\left|\mathbb{E}\left(y\right) - y\right|^p\right)\right]^{1/p}$$

or the upside and downside p-semi-deviations:

$$\sigma_p^+\left(y\right) := \left[\mathbb{E}\left(\left|Max\ \left\{y - \mathbb{E}\left(y\right), 0\right\}\right|^p\right)\right]^{1/p}$$

and:

$$\sigma_p^-\left(y\right) := \left[\mathbb{E}\left(\left|Max\ \left\{\mathbb{E}\left(y\right) - y, 0\right\}\right|^p\right)\right]^{1/p}$$

If $\tilde{E}_\rho = 1$, then it is easy to see that ρ is also coherent in the sense of Artzner *et al.* [1] if and only if:

$$\Delta_\rho \subset L^q_+ = \left\{z \in L^q; \mathbb{P}\left(z \geq 0\right) = 1\right\}$$

Assumption 1 may be relaxed, and the main conclusions of this paper remain true. For instance, sub-additivity and homogeneity may be replaced by convexity (in the line of Balbás *et al.* [10] or Föllmer and Schied [11]). Besides, $\tilde{E}_\rho$-translation invariance may be removed in (1b), in which case, the elements in the sub-gradient of ρ do not necessarily have constant expectation equaling $\tilde{E}_\rho$ (see (3)). Nevertheless, we prefer to impose Assumption 1 because it significantly simplifies the exposition.

We will also deal with the (metric, but not Banach) space L^0. Every random variable belongs to L^0, whose usual metric is given by:

$$d\left(y_1, y_2\right) = \mathbb{E}\left(Min\ \left\{\left|y_1 - y_2\right|, 1\right\}\right)$$

It is known that Metric d above leads to "convergence in probability", which is strictly weaker than the L^p-convergence. As said above, d cannot be given by a norm, and L^0 is not a Banach space. Therefore, the dual space of L^0 may be "too small", and this dual actually reduces to zero if $\mathbb{P}$ is atomless (Rudin [8]). In particular, if a function $\rho : L^0 \longrightarrow \mathbb{R}$ satisfies Condition (1a), then (2) implies that $\rho = 0$. In other words:

Remark 1. *Assumption* 1 *cannot be imposed for functionals* $\rho : L^0 \longrightarrow \mathbb{R}$*, because it would imply* $\rho = 0$ *if* $\mathbb{P}$ *were atomless.*

The latter remark implies that there are no non-null, continuous, sub-additive and homogeneous functionals on L^0 (or even on much smaller proper subspaces of L^0, Delbaen [12]). Yet, in finance, operational risk and insurance, one can find risks whose distribution does not belong to L^1, $i.e.$, it does not have a finite expectation. For instance, the advanced measurement approach (AMA) to Pillar I modeling of operational risk, as defined in Basel II, deals with random risks given by:

$$R = \sum_{i=1}^{k} R_i,$$

where every R_i is related to a specific business line and/or risk type as defined in the Basel II Accord (Nešlehová *et al.* [13]). Several R_i usually follow the Pareto distribution with parameters $\alpha > 0$ and $\beta > 0$ and with density function:

$$f(x) = \frac{\beta^\alpha \alpha}{(x+\beta)^{\alpha+1}},\ x > 0$$

The expectation of R_i is infinite if $\alpha \leq 1$.

Several authors have proposed to use VaR if tails are so heavy that it is impossible to find sub-additive risk measures (Chavez-Demoulin *et al.* [14], Embrechts *et al.* [15], *etc.*). Others have studied non-continuous sub-additive risk measures (Kupper and Svindland [4]). On the other hand, using continuous sub-additive risk measures has many important analytical advantages, since the optimization of such functions is much simpler and many classical financial and actuarial problems (pricing and hedging, portfolio choice, equilibrium, optimal reinsurance, *etc.*) become easier to tackle (Balbás *et al.* [10], among others). For these reasons, it may be worthwhile to look for partial solutions overcoming Remark 1 above, while still preserving some kind of continuity and sub-additivity. This is the main purpose of this paper.

Consider a finite collection of linearly independent final wealth:

$$\{w_1, w_2, ..., w_m\} \subset L^0$$

and suppose that their tails are very heavy and $w_i \notin L^1$, $i = 1, 2, ..., m$ ($i.e.$, $\mathbb{E}(|w_i|) = \infty$, $i = 1, 2, ..., m$). Consider the linear manifold L generated by $\{w_1, w_2, ..., w_m\}$ and suppose that it does not contain non-null elements of L^p. Since L has finite dimension, it only has a unique separated vector topology (Rudin [8]), and this is the one induced by the topology of L^0. In other words, the sequence $\left(\sum_{i=1}^m x_{i,n} w_i\right)_{n=1}^\infty$ converges in probability to $\sum_{i=1}^m x_i w_i$ if and only if $(x_{i,n})_{n=1}^\infty$ converges to x_i, $i = 1, 2, ..., m$. Thus, manifold L recovers the structure of a Banach space, and we can define non-trivial risk measures on L, that we will denote ρ_L.

Assumption 2. A risk measure $\rho_L : L \longrightarrow \mathbb{R}$ satisfies the equivalent Conditions a and b below:
(a) The set

$$\begin{aligned}\Delta_L &= \left\{ (\xi_i)_{i=1}^m \in \mathbb{R}^m;\ -\sum_{i=1}^m x_i \xi_i \leq \rho_L \left(\sum_{i=1}^m x_i w_i \right) \forall (x_i)_{i=1}^m \in \mathbb{R}^m \right\} \\ &\subset \mathbb{R}^m\end{aligned}$$

is convex and compact, and:

$$\rho_L(w) = Max \left\{ -\sum_{i=1}^m x_i \xi_i;\ (\xi_1, \xi_2, ..., \xi_m) \in \Delta_L \right\}$$

holds for every $w = \sum_{i=1}^{m} x_i w_i \in L$.

(b) ρ_L is continuous, sub-additive and homogeneous.

3. Extending the Risk Measure

As said above, L does not present the drawbacks of L^0, and the risk measure ρ_L does satisfy the required properties. According to Remark 1, ρ_L cannot be extended to the whole space L^0 unless we lose its good properties. Thus, let us propose a partial extension that allows us "to integrate" those risks included in L and those included in L^p. In practical applications, we do not expect to find infinitely many risks involving infinite expectations. More likely, we will just find a few (or even only one). Then, the proposed solution may be sufficient, since we will be able to have a "global risk measure" containing both ρ and ρ_L.

In order to jointly manage the risk given by ρ and ρ_L, we need to deal with the space:

$$L^p + L = \left\{ y + \sum_{i=1}^{m} x_i w_i \in L^0;\ y \in L^p,\ \ (x_1, x_2, ..., x_m) \in \mathbb{R}^m \right\}$$

which contains those risks included in L^p, those ones included in L and their linear combinations.

Theorem 1. *There exists an extension $\tilde{\rho}_L : L^p + L \longrightarrow \mathbb{R}$, such that:*

(a) $\tilde{\rho}_L$ is continuous, sub-additive, homogeneous and $\tilde{E}_\rho$-translation invariant (i.e., $\tilde{\rho}_L(y + w + k) = \tilde{\rho}_L(y + w) - \tilde{E}_\rho k$ if $y \in L^p$, $w \in L$ and $k \in \mathbb{R}$).

(b) $\tilde{\rho}_L(y) = \rho(y)$ if $y \in L^p$ and $\tilde{\rho}_L(w) = \rho_L(w)$ if $w \in L$.

(c) $\tilde{\rho}_L$ is minimal among the functionals $\Gamma : L^p + L \longrightarrow \mathbb{R}$ satisfying a and b.

(d) The set (sub-gradient of $\tilde{\rho}_L$):

$$\begin{aligned} \tilde{\Delta}_L &= \left\{ (z, \xi) ; - \mathbb{E}(yz) - \sum_{i=1}^{m} x_i \xi_i \le \tilde{\rho}_L \left(y + \sum_{i=1}^{m} x_i w_i \right) \ \forall (y, x) \in L^p \times \mathbb{R}^m \right\} \\ &\subset L^q \times \mathbb{R}^m \end{aligned} \tag{4}$$

is convex and $weak^*$-compact.

(e)

$$\tilde{\rho}_L(y + w) = Max \left\{ - \mathbb{E}(yz) - \sum_{i=1}^{m} x_i \xi_i;\ \ (z, \xi_1, \xi_2, ..., \xi_m) \in \tilde{\Delta}_L \right\} \tag{5}$$

holds for every $y \in L^p$ and every $w = \sum_{i=1}^{m} x_i w_i \in L$.[3]

See the Appendix. □

From an intuitive viewpoint, Theorem 1 has a simple interpretation. One can extend both ρ and ρ_L in such a manner that they become "integrated" in a global measure $\tilde{\rho}_L$, which preserves the required

[3] A similar statement may be also proven if $\{w_1, w_2, ..., w_m\}$ is not linearly independent or if $L^p \cap L$ contains non-null risks and $\rho(y) = \rho_L(y)$ for every $y \in L^p \cap L$. Nevertheless, we prefer to prove Theorem 1, in order to simplify the exposition.

properties, despite the fact that $L^p + L$ has infinitely many dimensions and the convergence in this global space still involves convergence in probability. Therefore, one can simultaneously deal with those standard risks y with finite expectations and those much heavier tailed risks w whose expectations are not finite.

Theorem 1 is an existence result, but it does not indicate how to construct $\tilde{\rho}_L$ in practice. Let us address this point.

Remark 2. *Building $\tilde{\rho}_L$ in practice for a single heavy tailed risk: Suppose firstly that $m = 1$, i.e., L is a linear manifold generated by only one heavy tailed risk w with no finite expected value.*

Step 1: Construct the sub-gradient $\tilde{\Delta}_L$ of $\tilde{\rho}_L$ in such a way that the natural projections $\Pi_q : L^q \times \mathbb{R}^m \longrightarrow L^q$ and $\Pi_m : L^q \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$ satisfy $\Pi_q\left(\tilde{\Delta}_L\right) = \Delta_\rho$ and $\Pi_m\left(\tilde{\Delta}_L\right) = \Delta_L$. This may be easily done as follows.

Step 2: Fix $\phi \in L^p$,

$$-\rho(\phi) = Min\left\{\mathbb{E}(\phi z); z \in \Delta_\rho\right\} \tag{6}$$

and:

$$\rho(-\phi) = Max\left\{\mathbb{E}(\phi z); z \in \Delta_\rho\right\} \tag{7}$$

(see (2)). Choose ϕ in such a manner that $-\rho(\phi) < \rho(-\phi)$.[4] *Notice that (6) and (7) obviously imply that:*

$$-\rho(\phi) \leq \mathbb{E}(\phi z) \leq \rho(-\phi) \tag{8}$$

holds for every $z \in \Delta_\rho$.

Step 3: Transform the interval:

$$[-\rho(\phi), \rho(-\phi)]$$

into the interval

$$[-\rho_L(w), \rho_L(-w)]$$

by means of the (one to one, unless $-\rho_L(w) = \rho_L(-w)$) increasing affine function[5]*:*

$$\begin{aligned}&[-\rho(\phi), \rho(-\phi)] \ni t \longrightarrow F(t) = \\ &-\rho_L(w) + \frac{\rho_L(-w) + \rho_L(w)}{\rho(-\phi) + \rho(\phi)}(t + \rho(\phi)) \in [-\rho_L(w), \rho_L(-w)]\end{aligned} \tag{9}$$

Step 4: $\tilde{\Delta}_L$ will be chosen according to the affine function above. More precisely,

$$\begin{aligned}\tilde{\Delta}_L &= \left\{(z, \xi); z \in \Delta_\rho \text{ and } \xi = F(\mathbb{E}(\phi z))\right\} \\ &\subset L^q \times \mathbb{R}\end{aligned} \tag{10}$$

[4] Notice that $0 = \rho(0) \leq \rho(\phi) + \rho(-\phi)$ leads to $-\rho(\phi) \leq \rho(-\phi)$. Moreover, the existence of ϕ satisfying the strict inequality holds if ρ is not linear (Δ_ρ is not a singleton), which is the case for all of the usual risk measures.

[5] Once again, $0 = \rho_L(0) \leq \rho_L(w) + \rho_L(-w)$ leads to $-\rho_L(w) \leq \rho_L(-w)$.

According to (8), the construction of $\tilde{\Delta}_L$ is correct and its $weak^*$-compactness follows from the $weak^*$-compactness of Δ_ρ and the continuity of F and $\mathbb{E}$. It may be also proven that $\Pi_q\left(\tilde{\Delta}_L\right) = \Delta_\rho$ and $\Pi_m\left(\tilde{\Delta}_L\right) = \Delta_L$ hold, though this proof is trivial and therefore omitted.

Step 5: Once $\tilde{\Delta}_L$ is known, $\tilde{\rho}_L$ is given by (5), which becomes now:

$$\tilde{\rho}_L\left(y+xw\right) = Max\left\{-\,\mathbb{E}\left(yz\right) - xF\left(\mathbb{E}\left(\phi z\right)\right);\ z\in\Delta_\rho\right\} \tag{11}$$

for every $y \in L^p$ and every $x \in \mathbb{R}$. Manipulating,

$$\begin{aligned}
&\mathbb{E}\left(yz\right) + xF\left(\mathbb{E}\left(\phi z\right)\right) = \\
&\mathbb{E}\left(yz\right) + x\left(-\rho_L\left(w\right) + \frac{\rho_L\left(-w\right)+\rho_L\left(w\right)}{\rho\left(-\phi\right)+\rho\left(\phi\right)}\left(\mathbb{E}\left(\phi z\right)+\rho\left(\phi\right)\right)\right) = \\
&\mathbb{E}\left(z\left(y + x\left(\frac{\rho_L\left(-w\right)+\rho_L\left(w\right)}{\rho\left(-\phi\right)+\rho\left(\phi\right)}\phi\right)\right)\right) \\
&+x\left(-\rho_L\left(w\right) + \frac{\rho_L\left(-w\right)+\rho_L\left(w\right)}{\rho\left(-\phi\right)+\rho\left(\phi\right)}\rho\left(\phi\right)\right)
\end{aligned}$$

and (2) and (11) imply that:

$$\begin{aligned}
&\tilde{\rho}_L\left(y+xw\right) = \\
&\rho\left(y + x\left(\frac{\rho_L\left(-w\right)+\rho_L\left(w\right)}{\rho\left(-\phi\right)+\rho\left(\phi\right)}\phi\right)\right) \\
&+\left(\rho_L\left(w\right) - \frac{\rho_L\left(-w\right)+\rho_L\left(w\right)}{\rho\left(-\phi\right)+\rho\left(\phi\right)}\rho\left(\phi\right)\right)x
\end{aligned} \tag{12}$$

Notice that (12) leads to $\tilde{\rho}_L\left(y\right) = \rho\left(y\right)$ if $x = 0$, and $\tilde{\rho}_L\left(w\right) = \rho_L\left(w\right)$ if $y = 0$ and $x = 1$. Furthermore, $\tilde{\rho}_L\left(xw\right) = \rho_L\left(xw\right)$ if $y = 0$ and $x \in \mathbb{R}$. In other words, $\tilde{\rho}_L$ really extends ρ and ρ_L.

The selection of ϕ (Step 2) is only constrained by the inequality $-\rho\left(\phi\right) < \rho\left(-\phi\right)$, which generates many degrees of freedom. In other words, we really have a choice when computing in practice the extension $\tilde{\rho}_L$ of Theorem 1, since it is not unique. The next section gives some rules to select ϕ, mainly related to the specific risk $\tilde{\rho}_L\left(y+xw\right)$ that one would like to associate with some reachable strategies $y + xw$ of $L^p + L$.

The selection of F in (9) is not unique, since the decreasing affine surjective function:

$$\begin{aligned}
&\left[-\rho\left(\phi\right),\rho\left(-\phi\right)\right] \ni t \longrightarrow G\left(t\right) = \\
&\rho_L\left(-w\right) - \frac{\rho_L\left(-w\right)+\rho_L\left(w\right)}{\rho\left(-\phi\right)+\rho\left(\phi\right)}\left(t+\rho\left(\phi\right)\right) \in \left[-\rho_L\left(w\right),\rho_L\left(-w\right)\right]
\end{aligned}$$

may play the role of F. Thus, if G replaces F, one will find a second extension of ρ and ρ_L still satisfying Theorem 1. Nevertheless, bearing in mind Condition c in Theorem 1, F and G are the unique valid choices. Other functions will make Condition c fail. We will not prove this result, because the proof is complex and beyond the scope of this paper.

Notice that (12) leads to:

$$\tilde{\rho}_L\left(y + xw\right) = \rho\left(y + k_1 x\phi\right) + k_2 x \tag{13}$$

where the parameters k_1 *and* k_2 *must satisfy* $k_1 \geq 0$ *and other additional constraints. If* ϕ *may be selected in such a way that* $\rho\left(\phi\right) = \rho_L\left(w\right)$ *and* $\rho\left(-\phi\right) = \rho_L\left(-w\right)$*, then (12) implies hat* $k_1 = 1$ *and* $k_2 = 0$*. Thus, (13) becomes:*

$$\tilde{\rho}_L\left(y + xw\right) = \rho\left(y + x\phi\right) \tag{14}$$

for every $x \in \mathbb{R}$*. An "intuitive interpretation" of (14) could be like this: pick an integrable random variable* $\phi \in L^p$ *that corresponds to the same risks as* $\rho_L\left(w\right)$ *and* $\rho_L\left(-w\right)$*, and then combinations* $y + xw$ *may be "identified" with combinations* $y + x\phi$ *in terms of risk.*

Remark 3. *Building* $\tilde{\rho}_L$ *in practice (the general case): Let us apply the induction method on the number* m *of heavy tailed risks. Suppose that we have an extension* $\tilde{\rho}_L$ *of* ρ *and* ρ_L *on* $L^p + L_{m-1}$*,* L_{m-1} *denoting the linear manifold generated by* $\{w_1, w_2, ..., w_{m-1}\}$*. In such a case, we have to extend* $\tilde{\rho}_L$ *to one more dimension, and it is obvious that the methodology described in Step 1–Step 5 above applies again. Thus, bearing in mind (12), we can select* $\phi_m \in L^p$ *with* $-\rho\left(\phi_m\right) < \rho\left(-\phi_m\right)$*, and the global risk measure* $\tilde{\rho}_L$ *will be given by:*

$$\begin{aligned}&\tilde{\rho}_L\left(y + \textstyle\sum_{i=1}^m x_i w_i\right) = \\ &\tilde{\rho}_L\left(y + x_m\left(\frac{\rho_L\left(-w_m\right) + \rho_L\left(w_m\right)}{\rho\left(-\phi_m\right) + \rho\left(\phi_m\right)}\phi_m\right) + \textstyle\sum_{i=1}^{m-1} x_i w_i\right) \\ &+ \left(\rho_L\left(w_m\right) - \frac{\rho_L\left(-w_m\right) + \rho_L\left(w_m\right)}{\rho\left(-\phi_m\right) + \rho\left(\phi_m\right)}\rho\left(\phi_m\right)\right)x_m\end{aligned} \tag{15}$$

4. Examples

Expressions (12) and (15) provide us with the continuous, sub-additive, homogeneous and $\tilde{E}_\rho$-translation invariant extensions for which we were looking. Moreover, they may be easily computed in practice, because we have closed formulas, and they are easily optimized (minimized), because we have a convex and $weak^*$-compact sub-gradient (see (4), (5), (10) and (11)). Let us show some examples whose sole objective is to illustrate how $\tilde{\rho}_L$ is in practice.

Example 1. *Suppose that we would like to use the conditional value at risk* $\rho = CVaR_\mu$ *with the level of confidence* $0 < \mu < 1$*, but one of the risks* w *has a Pareto distribution with unbounded expectation. Obviously,* $\rho\left(w\right) = CVaR_\mu\left(w\right) = \infty$*, which means that* $CVaR_\mu$ *is not continuous any more, and therefore, the computation and optimization of risks, such as* $CVaR_\mu\left(y + xw\right)$*, with* $y \in L^1$ *and* $x \in \mathbb{R}$*, is quite difficult to address from a mathematical perspective (Balbás et al. [10]). Then, following Embrechts et al. [15], one can deal with the value at risk in order to measure the risk of the heavy tailed* *distribution* w *and then integrate this risk measure with the conditional value at risk. In such a case, (12) leads to the risk measure:*

$$\begin{gathered}\tilde{C}VaR_{\mu}\left(y+xw\right)= \\ CVaR_{\mu}\left(y+x\left(\frac{VaR_{\mu}\left(-w\right)+VaR_{\mu}\left(w\right)}{CVaR_{\mu}\left(-\phi\right)+CVaR_{\mu}\left(\phi\right)}\phi\right)\right) \\ +\left(VaR_{\mu}\left(w\right)-\frac{VaR_{\mu}\left(-w\right)+VaR_{\mu}\left(w\right)}{CVaR_{\mu}\left(-\phi\right)+CVaR_{\mu}\left(\phi\right)}CVaR_{\mu}\left(\phi\right)\right)x\end{gathered} \tag{16}$$

for $y\in L^1$ *and* $x\in\mathbb{R}$*. Expression (16) provides us with a continuous, sub-additive, homogeneous and one-translation-invariant risk measure that extends* $CVaR_{\mu}$ *and applies to the Pareto distribution* w *and its linear combinations with risks of bounded expectation. Furthermore, according to Remark* 2*, (4) and (5), and bearing in mind that the* $CVaR_{\mu}$*-sub-gradient is (Rockafellar et al. [3]):*

$$\left\{z\in L^1;\mathbb{E}\left(z\right)=1,\ 0\leq z\leq\frac{1}{1-\mu}\right\}$$

(16) may be represented by its convex and $weak^*$*-compact sub-gradient, composed of those couples* $(z,\xi)\in L^1\times\mathbb{R}$*, such that:*

$$\begin{cases}\mathbb{E}\left(z\right)=1 \\ 0\leq z\leq\frac{1}{1-\mu} \\ \xi=-VaR_{\mu}\left(w\right)+\dfrac{VaR_{\mu}\left(-w\right)+VaR_{\mu}\left(w\right)}{CVaR_{\mu}\left(-\phi\right)+CVaR_{\mu}\left(\phi\right)}\left(\mathbb{E}\left(\phi z\right)+CVaR_{\mu}\left(\phi\right)\right)\end{cases}. \tag{17}$$

As already said, we have a choice for $\phi\in L^1$*, and the selection of this risk affects the final extension* $\tilde{C}VaR_{\mu}$ *in (16). In practice,* ϕ *may be chosen in such a manner that the risk of some selected heavy tailed distributions still matches their value at risk, since* VaR_{μ} *has a nice economic interpretation in terms of potential capital losses. Formally, one can select a collection of risks* $\{y_1,y_2,...,y_k\}\subset L^1$ *with bounded expectation and then choose* $\phi\in L^1$*, so as to satisfy:*

$$\tilde{C}VaR_{\mu}\left(y_i+w\right)=VaR_{\mu}\left(y_i+w\right)$$

$i=1,2,...,k$*.* □

Example 2. *According to Remark* 3*, Example* 1 *may be easily extended for more than one Pareto (or other heavy tailed) distribution. In other words, if* $\{w_1,w_2,...,w_m\}$ *are independent Pareto distributions whose non-trivial linear combinations have unbounded expectations, then one can construct a continuous, sub-additive, homogeneous and one-translation-invariant* $\tilde{C}VaR_{\mu}$*, such that* $\tilde{C}VaR_{\mu}(y)=CVaR_{\mu}(y)$ *if the expectation of* y *is bounded, and* $\tilde{C}VaR_{\mu}(w_i)=VaR_{\mu}(w_i)$*,* $i=1,2,...,m$*. Expression (15) provides us with the effective construction in a recursive manner. Besides, the role of CVaR may be played by many other coherent risk measures (Wang measure, dual power transform, risk measures given by concave distortions,* $WCVaR$*, etc.), and the role of VaR may be played by alternative selections of every risk* $\rho_L(w_i)$ *(ad hoc selections based on expert* *opinions, such that* $-\rho_L(w_i) \leq \rho_L(-w_i)$*, VaR with a much higher level of confidence than* μ *if the tails are too fat, etc.). Finally, if the role of the coherent risk measure is replaced by a deviation measure (absolute deviation, standard deviation, semi-deviations, etc.), then we will be creating its continuous, sub-additive, homogeneous and zero-translation-invariant extensions (deviations) that also apply for some risks with unbounded expectation (and therefore, unbounded first order central moment, unbounded second order central moment, etc.) and their linear combinations with risks with finite mean value.* □

Example 3. *Actuarial application: extending the expected value premium principle: Many classical financial and actuarial problems (portfolio choice, optimal reinsurance, operational risk,* etc.*) have been revisited with coherent risk measures. Our extension permits us to involve risks with unbounded expectation.*

Let us deal with a particular application whose purpose is just illustrative. Consider a random wealth w *whose expectation is infinite.* w *endows* $\mathbb{R}$ *with a probability measure* $\mathbb{P}$ *on its Borel* σ*-algebra* $\mathcal{F}$*. Obviously, here:*

$$\mathbb{E}(|w|) = \int_{-\infty}^{\infty} |u| \, d\mathbb{P}(u) = \infty \tag{18}$$

and, therefore, $\mathbb{E}(w)$ *and* $CVaR_\mu(w)$ *do not exist. According to Example* 1*, we can consider the risk measure* $\tilde{C}VaR_\mu$*, which satisfies* $\tilde{C}VaR_\mu(w) = VaR_\mu(w)$*.*

There are many derivatives $f(w)$ *of* w *belonging to* L^1*, and therefore, the expected value premium principle (*$EVPP$*) applies for them all. If* $\lambda \geq 0$ *is the loading rate, then the price of* $f(w)$ *will be given by:*

$$EVPP_\lambda(f(w)) = (1+\lambda)\,\mathbb{E}(f(w)) = (1+\lambda)\int_{-\infty}^{\infty} f(u)\, d\mathbb{P}(u) \tag{19}$$

Interesting particular cases for $f(w)$ *may be the "call-spreads with thresholds* $a < b$*", given by:*

$$f(w) = \begin{cases} 0, \; if \; w < a \\ w - a, \; if \; a \leq w \leq b \\ b - a, \; if \; w > b \end{cases}$$

but there are many more examples. Since (19) applies to price for every $f(w) \in L^1$*, we can use the pricing method of Balbás et al. [16] in order to overcome the caveat implied by (18), and therefore, we can extend (19) and create the pricing rule* $EVPP_{(\lambda,\mu,\phi)}$ *for risks being linear combinations of* w *and* L^1*. Actually,* $EVPP_{(\lambda,\mu,\phi)}(w)$ *will be the optimal value of the following optimization problem:*

$$\begin{cases} Min \;\; (1+\lambda)\left[\tilde{C}VaR_\mu(y-w) + \mathbb{E}(y)\right] \\ \\ y = f(w) \in L^1 \end{cases}$$

where y *is the decision variable. According to Balbás et al. [16],* $EVPP_\lambda$ *will be extended to* $EVPP_{(\lambda,\mu,\phi)}$ *in such a manner that it is still continuous on* $L^1 + \{xw; x \in \mathbb{R}\}$*, sub-additive and* *homogeneous. Moreover, bearing in mind (10), (17) and the duality results of Balbás et al. [16], it may be proven that:*

$$EVPP_{(\lambda,\mu,\phi)}(w) =$$

$$(1+\lambda)\left[-VaR_{\mu}(w) + \frac{VaR_{\mu}(-w) + VaR_{\mu}(w)}{CVaR_{\mu}(-\phi) + CVaR_{\mu}(\phi)}\left(\mathbb{E}(\phi) + CVaR_{\mu}(\phi)\right)\right]$$

$\square$

5. Conclusions

This paper proposes a constructive way to extend risk measures beyond L^1 and simultaneously preserve good mathematical properties, such as continuity, sub-additivity, homogeneity and translation invariance. This may be very useful to integrate risks with unbounded expectation (for instance, some Pareto distributions) with more standard risks in a global framework. The good properties of the extended risk measure have favorable implications, in the sense that many classical actuarial and financial problems (pricing, hedging, portfolio selection, equilibrium, optimal reinsurance, operational risk, *etc.*) may be revisited even when some expectations are infinite. Illustrative practical examples are presented, such as extensions for the conditional value at risk or for the expected value premium principle.

Acknowledgments

The authors sincerely thank the anonymous reviewers for their useful comments and suggestions.

Research partially supported by "Ministerio de Economía"(Spain, Grant ECO2012-39031-C02-01) and the Natural Sciences and Engineering Research Council (NSERC) of Canada, Grant 36860-2012.

Author Contributions

The ideas of this paper are the result of many discussions among the authors. Therefore, the three authors have all contributed to and are responsible for the whole content.

Appendix: Proof of Theorem 1

Consider the family $\mathcal{C}$ of convex and $weak^*$-compact subsets $C \subset L^q \times \mathbb{R}^m$, such that the natural projections $\Pi_q : L^q \times \mathbb{R}^m \longrightarrow L^q$ and $\Pi_m : L^q \times \mathbb{R}^m \longrightarrow \mathbb{R}^m$ satisfy $\Pi_q(C) = \Delta_\rho$ and $\Pi_m(C) = \Delta_L$. $\mathcal{C}$ is non-void since:

$$\Delta_\rho \times \Delta_L = \{(z,\xi);\ z \in \Delta_\rho,\ \xi \in \Delta_L\} \in \mathcal{C}.$$

If we show that $\mathcal{C}$ is inductive, then Zorn's lemma will guarantee the existence of a minimal element $\tilde{\Delta}_L \in \mathcal{C}$ (Kelly [17]). In order to see that $\mathcal{C}$ is inductive, consider a totally ordered chain $\{C_i; i \in I\} \subset \mathcal{C}$ and let us show that:

$$C = \cap_{i \in I} C_i$$

is a lower bound of $\{C_i; i \in I\}$ that belongs to $\mathcal{C}$. Since C is obviously convex and $weak^*$-compact, it is sufficient to see that $\Pi_q(C) = \Delta_\rho$ and $\Pi_m(C) = \Delta_L$. Let us show that $\Pi_q(C) = \Delta_\rho$, and the second equality will be analogous (and therefore omitted). $\Pi_q(C) \subset \Pi_q(C_i) = \Delta_\rho$ for every $i \in I$ is obvious, so let us show the opposite inclusion. Suppose that $z \in \Delta_\rho$, and take the net $(h_i)_{i \in I}$, such that $h_i \in C_i$ and $\Pi_q(h_i) = z$ for every $i \in I$. Fix $i_0 \in I$. Since C_{i_0} is $weak^*$-compact, there exists an agglomeration point $h \in C_{i_0}$ of $(h_i)_{i \geq i_0}$ (Kelly [17]), and $\Pi_q(h) = z$ is obvious, because Π_q is $weak^*$-continuous. It only remains to see that $h \in C_{i_1}$ for every $i_1 \geq i_0$, which is a trivial consequence of $(h_i)_{i \geq i_1} \subset C_{i_1}$, because C_{i_1} is $weak^*$-compact and, therefore, $weak^*$-closed.

Once the existence of a minimal element $\tilde{\Delta}_L \in \mathcal{C}$ has been proven, define $\tilde{\rho}_L : L^p + L \longrightarrow \mathbb{R}$ according to (5). The results of Balbás *et al.* [9] show that $\tilde{\rho}_L$ satisfies Conditions a, d and e of Theorem 1. Let us show Condition b. We will show that $\tilde{\rho}_L(y) = \rho(y)$ if $y \in L^p$, and the second equality will be omitted because it is similar. Bearing in mind $\Pi_q\left(\tilde{\Delta}_L\right) = \Delta_\rho$, (5) and (2), we have:

$$\begin{aligned}\tilde{\rho}_L(y) &= Max\left\{-\mathbb{E}(yz);\ (z,\xi) \in \tilde{\Delta}_L\right\} \\ &= Max\left\{-\mathbb{E}(y\Pi_q(z,\xi));\ (z,\xi) \in \tilde{\Delta}_L\right\} \\ &= Max\left\{-\mathbb{E}(yz);\ z \in \Delta_\rho\right\} = \rho(y).\end{aligned}$$

Let us finally show Condition c. According to Balbás *et al.* [9], there is a bijection:

$$\mathcal{D} \leftrightarrow \Gamma \tag{20}$$

between the family $\mathcal{D}$ of convex and $weak^*$-compact subsets $C \subset L^q \times \mathbb{R}^m$ and the family Γ of continuous, sub-additive and homogeneous functionals $f : L^p + L \longrightarrow \mathbb{R}$. Moreover, if $C \in \mathcal{D}$, then its associated functional of Γ is given by:

$$f(y+w) = Max\left\{-\mathbb{E}(yz) - \sum_{i=1}^{m} x_i\xi_i;\ (z,\xi_1,\xi_2,...,\xi_m) \in C\right\},$$

whereas for $f \in \Gamma$, the associated set of $\mathcal{D}$ is given by:

$$\begin{aligned}&C = \left\{(z,\xi);-\mathbb{E}(yz) - \textstyle\sum_{i=1}^m x_i\xi_i \leq f\left(y + \sum_{i=1}^m x_i w_i\right)\ \forall (y,x) \in L^p \times \mathbb{R}^m\right\} \\ &\subset L^q \times \mathbb{R}^m.\end{aligned} \tag{21}$$

Suppose that $f : L^p + L \longrightarrow \mathbb{R}$ satisfies Conditions a and b of Theorem 1 and $f(y+w) \leq \tilde{\rho}_L(y+w)$ for every $y + w \in L^p + L$. Since the identification (20) is increasing, the set C of (21) satisfies $C \subset \tilde{\Delta}_L$. If we show that $\Pi_q(C) = \Delta_\rho$ and $\Pi_L(C) = \Delta_L$, then we will have $C \in \mathcal{C}$, and $\tilde{\Delta}_L$ being minimal, we will have that $C = \tilde{\Delta}_L$; and, therefore, $f = \tilde{\rho}_L$. Hence, it remains to see that $\Pi_q(C) = \Delta_\rho$ and $\Pi_m(C) = \Delta_L$, and let us prove the first equality only, since the second one is similar. Condition b implies that:

$$\begin{aligned}&Max\left\{-\mathbb{E}(y\Pi_q(z,\xi));\ (z,\xi) \in C\right\} \\ =\ &Max\left\{-\mathbb{E}(yz);\ (z,\xi) \in C\right\} = f(y) \\ =\ &\rho(y) = Max\left\{-\mathbb{E}(yz);\ z \in \Delta_\rho\right\}\end{aligned}$$

for every $y \in L^p$. Hence, the identification (20) applying for L^p rather than $L^p + L$ implies that $\Pi_q(C) = \Delta_\rho$, because both $\Pi_q(C)$ and Δ_ρ are obviously convex and $weak^*$-compact subsets of L^q. □

Conflicts of Interest

The authors declare no conflicts of interest.

References

1. Aumann, R.J.; Serrano, R. An economic index of riskiness. *J. Polit. Econ.*. **2008**, *116*, 810–836.
2. Artzner, P.; Delbaen, F.; Eber, J.M.; Heath, D. Coherent measures of risk. *Math. Financ.* **1999**, *9*, 203–228.
3. Rockafellar, R.T.; Uryasev, S.; Zabarankin, M. Generalized deviations in risk analysis. *Financ. Stoch.* **2006**, *10*, 51–74.
4. Kupper, M.; Svindland, G. Dual representation of monotone convex functions on L^0. *Proc. Am. Math. Soc.* **2011**, *139*, 4073–4086.
5. Ogryczak, W.; Ruszczynski, A. Dual stochastic dominance and related mean risk models. *SIAM J. Optim*. **2002**, *13*, 60–78.
6. Mansini, R.; Ogryczak, W.; Speranza, M.G. Conditional value at risk and related linear programming models for portfolio optimization. *Ann. Oper. Res.* **2007**, *152*, 227–256.
7. Konno, H.; Akishino, K.; Yamamoto, R. Optimization of a long-short portfolio under non-convex transaction costs. *Comput. Optim. Appl.* **2005**, *32*, 115–132.
8. Rudin, W. *Functional Analysis*; McGraw-Hill, Inc: New-York, USA, 1973.
9. Balbás, A.; Balbás, B.; Balbás, R. Good deals in markets with friction. *Quantit. Financ.* **2013**, *13*, 827–836.
10. Balbás, A.; Balbás, B.; Balbás, R. Minimizing measures of risk by saddle point conditions. *J. Comput. Appl. Math.* **2010**, *234*, 2924–2931.
11. Föllmer, H.; Schied, A. Convex measures of risk and trading constraints. *Financ. Stoch.* **2002**, *6*, 429–447.
12. Delbaen, F. Risk measures for non-integrable random variables. *Math. Financ.* **2009**, *19*, 329–333.
13. Nešlehová, J.; Embrechts, P.; Chavez-Demoulin, V. Infinite-mean models and the LDA for operational risk. *J. Oper Risk.* **2006**, *1*, 3–25.
14. Chavez-Demoulin, V.; Embrechts, P.; Nešlehová. J. Quantitative models for operational risk: Extremes, dependence and aggregation. *J. Bank. Financ.* **2006**, *30*, 2635–2658.
15. Embrechts, P.; Puccetti, G.; Rüschendorf, L.; Wang, R.; Beleraj, A. An academic response to Basel 3.5. *Risks* **2014**, *2*, 25–48.
16. Balbás, A.; Balbás, R.; Garrido, J. Extending pricing rules with general risk functions. *Eur. J. Oper. Res.* **2010**, *201*, 23–33.
17. Kelly, J. *General Topology*; Graduate Texts in Mathematics; Springer: New-York, USA, 1975.