Regulatory Arbitrage of Risk Measures

Ruodu Wang*

November 29, 2015

Abstract

We introduce regulatory arbitrage of risk measures as one of the key considerations in choosing a suitable risk measure to use in banking regulation. When a risk measure is applied to calculate regulatory capital requirement, the magnitude of regulatory arbitrage is the amount of possible capital requirement reduction through splitting a financial risk into several fragments. Coherent risk measures by definition are free of regulatory arbitrage; dividing risks will not reduce the total capital requirement under a coherent risk measure. However, risk measures in practical use, such as the Value-at-Risk (VaR), are often not coherent and the magnitude of their regulatory arbitrage is then of significant importance. We quantify regulatory arbitrage of risk measures in a rigorous mathematical framework, and categorize risk measures into three classes: free of regulatory arbitrage, of limited regulatory arbitrage, and of infinite regulatory arbitrage. We provide explicit results to characterize regulatory arbitrage for general classes of risk measures, including distortion risk measures and convex risk measures. Several examples of risk measures of limited regulatory arbitrage are illustrated, as possible alternatives for coherent risk measures.

Key-words: risk measures; regulatory arbitrage, subadditivity; Value-at-Risk; regulatory capital

*Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada. Email: wang@uwaterloo.ca. The author would like to thank the Editor, two anonymous reviewers, Paul Embrechts, Andreas Tsanakas, Giovanni Puccetti and Tiantian Mao for helpful comments and discussions on an earlier version of this paper. The author acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC Grant No. 435844). This version contains some corrections made in April 2016, thanks to Jennifer Braun for pointing these errors out.

1 Introduction

In the past few decades, risk measures have been a prominent tool for risk management in both academic research and industrial practice. A risk measure maps a financial risk, usually represented by a random variable, to a real number which can be interpreted differently within various contexts. In banking, it represents the capital requirement to regulate a risk; in insurance, it calculates the premium for an insurance contract; in economics, it ranks the preference of a risk for a market participant. For instance, the class of distortion risk measures, with various names and slightly different definitions, can be found in Artzner et al. (1999) and Acerbi (2002) for banking regulation, in Wang et al. (1997) for insurance premium principles, and in Yaari (1987) and Schmeidler (1989) for economic risk preference.

In this paper, we interpret risk measures as regulatory capital requirements for financial risks in a fixed future time period, as treated in Artzner et al. (1999). That is, $\rho : \mathcal{X} \to \mathbb{R}$, where $\mathcal{X}$ is a set of random variables in a pre-specified probability space. Since its introduction in the 1980s, the Value-at-Risk (VaR) has been the most popular risk measure in use for the purpose of capital requirements. The Expected Shortfall (ES, also known as TVaR in insurance regulation), as an alternative to VaR, has drawn increasing attention during the last decade. There have been extensive debates on "*VaR versus ES in regulation*", stimulated by the potential move from VaR to ES proposed by the Basel Committee on Banking Supervision (BCBS) in two recent consultative documents (BCBS, 2012, 2013). The same discussion takes place in insurance regulation; see SCOR (2008), Sandström (2010) and IAIS (2014). We refer to the academic papers Embrechts et al. (2014) and Emmer et al. (2014) for comprehensive discussions on recent issues related to the two risk measures.

One key property that distinguishes coherent risk measures, such as ES, from non-coherent risk measures, such as VaR, is *subadditivity* (or a similar notion of *convexity*). More precisely, a subadditive risk measure ρ satisfies that for any two risks X and Y, $\rho(X + Y) \leqslant \rho(X) + \rho(Y)$ always holds. More detailed theory and financial interpretation of subadditivity can be found in Delbaen (2012).

Subadditivity was proposed in the seminal paper Artzner et al. (1999) as a desirable property for a reasonable capital requirement, based on arguments of "a merger does not create extra risk". Translating this property into the language of finance, a subadditive risk measure suggests that putting two risks together in one portfolio would not increase the total capital requirement, compared to as if the two risks were run in two separate portfolios. Similarly, when calculating the guarantee for a business of multiple desks, a supervisor can count the sum of the evaluated individual guarantees to cover the total risk. Subadditivity also creates convenience to allocate the overall risk in portfolios to individual risk factors; see discussions in Delbaen (2000, Section 8).

Although subadditivity is supported by many academics in mathematical finance, it is sometimes criticized by some financial statisticians and practitioners from the perspectives of practicality, robustness, and backtesting. In particular, coherent risk measures are criticized for lack of robustness (Cont et al., 2010) and elicitability (Gneiting, 2011). For recent developments, we refer to Kou and Peng (2014); Krätschmer et al. (2014); Embrechts et al. (2015) on robustness issues and Davis (2013); Ziegel (2015); Bellini and Bignozzi (2015) on elicitability and backtesting issues. For the backtesting of ES, see McNeil and Frey (2000) and Kerkhof and Melenberg (2004).

In this paper, we do not intend to jump to a conclusion for the debates on subadditive and non-subadditive risk measures, but would like to provide some insight from another perspective. Let us look at one of the original arguments in Artzner et al. (1999) supporting subadditivity; similar arguments can be found in many other papers.

> *"If a firm were forced to meet a requirement of extra capital which did not satisfy this property* [subadditivity]*, the firm might be motivated to break up into two separately incorporated affiliates, a matter of concern for the regulator."*

Thus, a *regulatory arbitrage*[1] exists for ρ if $\rho(X) > \rho(Y) + \rho(Z)$ for $Y + Z = X$. It is obvious that a subadditive risk measure excludes the above regulatory arbitrage. In this paper, we propose to quantify this type of regulatory arbitrage. That is, to answer the following question:

If a non-subadditive risk measure ρ is employed for calculating capital requirements, how large is the possible regulatory arbitrage?

In this paper, we define a *regulatory arbitrage* associated with a risk measure ρ at a risk X as the largest possible difference between $\rho(X)$ and $\rho(X_1) + \cdots + \rho(X_n)$ where $X_1, \ldots, X_n$ are fragments adding up to X, i.e. $\sum_{i=1}^n X_i = X$, and n is any positive integer. Such problems for finite n appear in the context of risk sharing or risk transfer under convex risk measures; see Barrieu and El Karoui (2005) and Jouini et al. (2008). We carry out a systematic research of this quantity as a property of the

[1] As a reviewer kindly pointed out, the term "regulatory arbitrage" in common finance parlance has a much broader meaning. When we study regulatory arbitrage of risk measures, a particular partial and narrow interpretation of the concept is used in this paper.

underlying risk measure. We categorize risk measures into three classes: free of regulatory arbitrage, of limited regulatory arbitrage, and of infinite regulatory arbitrage. A risk measure of infinite regulatory arbitrage could be problematic as a capital requirement principle. An important observation is that VaR at any confidence level is of infinite regulatory arbitrage, whereas ES or any coherent risk measures are always free of regulatory arbitrage.

We further extend our study of regulatory arbitrage for two most commonly used and studied classes of risk measures: distortion risk measures (Wang et al., 1997; Acerbi, 2002), and convex risk measures (Föllmer and Schied, 2002; Frittelli and Rosazza Gianin, 2002). It turns out that for these two classes of risk measures ρ, regulatory arbitrage can be quantified as the difference between ρ and the largest coherent risk measure dominated by ρ. We observe that many (non-coherent) convex risk measures and distortion risk measures have limited regulatory arbitrage, and hence are more resistant to a manipulation of risks than VaR. We also provide some examples of risk measures used in practice, for example the worst- and best-case risk measures, mixture of risk measures, and risk-adjusted capital (see FOPI, 2006, in the context of the Swiss Stress Test).

Most study on the non-subadditive phenomena of risk measures is from the viewpoint of risk aggregation (see for instance Embrechts et al., 2013), whereas our study is based on the perspective of risk fragmentation. Through the study of regulatory arbitrage, we highlight the importance of some properties of risk measures, which, to some extend, have been neglected in finance. The concept of regulatory arbitrage is not confined to law-invariant risk measures, although all examples are law-invariant due to statistical tractability in practice.

The rest of the paper is organized as follows. In Section 2 we put some preliminaries, introduce regulatory arbitrage and minimal measures induced by risk measures, and discuss their theoretical properties. In Section 3 we study regulatory arbitrage of two special classes of commonly-used risk measures: distortion risk measures and convex risk measures. In both cases, explicit quantifications of regulatory arbitrage and minimal measure are available. In Section 4, regulatory arbitrage with transaction costs is studied. Some examples are given in Section 5. Section 6 draws a conclusion.

2 Regulatory arbitrage

In this section, we study the major subject of this paper: regulatory arbitrage of risk measures. Before introducing the concept of regulatory arbitrage, we first review some preliminaries on risk measures.

2.1 Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be an atomless probability space and denote by $L^0 := L^0(\Omega, \mathcal{F}, \mathbb{P})$, the set of all random variables in the probability space. In this paper, $X \in L^0$ represents a financial loss when $X > 0$, and a profit when $X < 0$. We denote by L^∞ (resp. L^1) the set of all bounded (resp. integrable) random variables in $(\Omega, \mathcal{F}, \mathbb{P})$.

A *risk measure* is a function $\rho : \mathcal{X} \to \mathbb{R} \cup \{-\infty, \infty\}$, which assigns a number or infinity to each financial risk in a set $\mathcal{X}$ of random variables, $\mathcal{X} \supset L^\infty$ and $\mathcal{X}$ is closed under addition and positive scalar multiplication. We naturally assume $\rho(X) < \infty$ for all $X \in L^\infty$, i.e. a finite capital requirement for bounded risks. Below we list some standard properties considered for risk measures: for all $X, Y \in \mathcal{X}$,

(A1) *Monotonicity*: if $X \leqslant Y$ $\mathbb{P}$-a.s, then $\rho(X) \leqslant \rho(Y)$;

(A2) *Cash-invariance*: for any $m \in \mathbb{R}$, $\rho(X - m) = \rho(X) - m$;

(A3) *Positive homogeneity*: for any $\alpha > 0$, $\rho(\alpha X) = \alpha\rho(X)$;

(A4) *Subadditivity*: $\rho(X + Y) \leqslant \rho(X) + \rho(Y)$;

(A5) *Convexity*: for any $\lambda \in [0, 1]$, $\rho(\lambda X + (1 - \lambda)(Y)) \leqslant \lambda\rho(X) + (1 - \lambda)\rho(Y)$;

(A6) *Loadedness*: $\rho(X) \geqslant \mathbb{E}[X]$;

(A7) *Comonotonic additivity*: if X, Y are comonotonic, then $\rho(X + Y) = \rho(X) + \rho(Y)$;

(A8) *Law-invariance*: if X and Y have the same distribution F, then $\rho(X) = \rho(Y)$.

Two random variables X and Y are comonotonic (Yaari, 1987) if

$$(X(\omega) - X(\omega'))(Y(\omega) - Y(\omega')) \geqslant 0 \ \text{ for a.s. } (\omega, \omega') \in \Omega^2.$$

Different sets of properties are typically chosen in different contexts. Note that some of the above properties may be implied by a set of other properties. We refer to Föllmer and Schied (2011) for a general introduction to risk measures in mathematical finance. A *monetary risk measure* satisfies (A1, A2), a *convex risk measure* satisfies (A1, A2, A5), and a *coherent risk measure* satisfies (A1-A5). The loadedness (A6) is essential in insurance premium pricing (in which, a risk measure is usually interpreted as a premium principle), as the insurance industry relies highly on laws of large numbers. Although the loadedness receives less attention as a property of capital requirements in finance, we will notice its importance later in this paper. Law-invariance (A8) is a desirable property from a statistical point of view, and commonly-used risk measures in practice are all law-invariant.

The most popular risk measure in banking and finance is the *Value-at-Risk* (VaR) at confidence level $p \in (0,1)$, defined as $\mathrm{VaR}_p : L^0 \to \mathbb{R}$,

$$\mathrm{VaR}_p(X) = \inf\{x : \mathbb{P}(X \leqslant x) \geqslant p\}, \quad X \in L^0,$$

which satisfies (A1-A3, A7-A8). It is mostly criticized for not being subadditive. In this paper we will see that it has another significant drawback related to regulatory arbitrage. A coherent alternative to VaR is *Expected Shortfall* (ES) at confidence level $p \in (0,1)$, defined as $\mathrm{ES}_p : L^0 \to (-\infty, \infty]$,

$$\mathrm{ES}_p(X) = \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(X)\mathrm{d}q, \quad X \in L^0,$$

satisfying properties (A1-A8). Note that if $X \in L^1$, then $\mathrm{ES}_p(X) \in \mathbb{R}$, $p \in (0,1)$. In addition, we let $\mathrm{ES}_0 : L^1 \to \mathbb{R}$, $\mathrm{ES}_0(X) = \mathbb{E}[X]$ and $\mathrm{ES}_1 : L^0 \to (-\infty, \infty]$, $\mathrm{ES}_1(X) = \text{ess-sup}(X)$, the essential supremum of X. Throughout the paper, we follow the convention $\sup \varnothing = -\infty$.

2.2 Minimal measure and regulatory arbitrage

A regulatory arbitrage occurs when dividing a risk $X \in \mathcal{X}$ into several fragments results in a reduced capital requirement. A rigorous quantification of this phenomenon calls for the following definition of a minimal measure Ψ_ρ.

Definition 2.1. The *minimal measure* of a risk measure ρ, $\Psi_\rho : \mathcal{X} \to [-\infty, \infty]$, is defined as

$$\Psi_\rho(X) = \inf\left\{\sum_{i=1}^n \rho(Y_i) : n \in \mathbb{N},\ Y_i \in \mathcal{X},\ i = 1, \dots, n,\ \sum_{i=1}^n Y_i = X\right\}.$$

Ψ_ρ has a clear financial interpretation: it represents the minimum amount of capital requirement among all possible ways of dividing a risk $X \in \mathcal{X}$ into fragments. It is immediate that $\Psi_\rho \leqslant \rho$, and Ψ_ρ itself can be considered as a risk measure. For the ease of discussion, we assume that there is no cost that occurs in dividing risks. We will study cases when a transaction cost is applied to dividing risks in Section 4. The minimal measure Ψ_ρ can be seen as the limit of *inf-convolutions* of ρ; see (2.3) and Remark 2.1 below.

With the notion of the minimal measure, we are now able to quantify regulatory arbitrage by the most natural way: the risk measure minus its minimal measure. In almost all practical cases, risk measures take values in $\mathbb{R}$, for its interpretation as the capital requirement. Therefore, we can safely assume $\rho(X) \in \mathbb{R}$ for all $X \in \mathcal{X}$ throughout the rest of this paper to avoid ill-posed situations such as $\infty - \infty$. Note that the definition of Ψ_ρ does not require that $\rho(X) \in \mathbb{R}$.

Definition 2.2. The *regulatory arbitrage* of a risk measure ρ, $\Phi_\rho : \mathcal{X} \to [0, \infty]$, is defined as

$$\Phi_\rho(X) = \rho(X) - \Psi_\rho(X).$$

Obviously, for some risk measures, there is a relatively large regulatory arbitrage, whereas for some other risk measures, regulatory arbitrage can be small or zero. Thus, we naturally categorize risk measures into four classes.

Definition 2.3. A risk measure ρ is

(i) *free of regulatory arbitrage* if $\Phi_\rho(X) = 0$ for all $X \in \mathcal{X}$,

(ii) *of limited regulatory arbitrage* if $\Phi_\rho(X) < \infty$ for all $X \in \mathcal{X}$,

(iii) *of unlimited regulatory arbitrage* if $\Phi_\rho(X) = \infty$ for some $X \in \mathcal{X}$,

(iv) *of infinite regulatory arbitrage* if $\Phi_\rho(X) = \infty$ for all $X \in \mathcal{X}$.

Later we will see that (iii) and (iv) are equivalent from a more detailed mathematical analysis. For now, we first confirm the straightforward fact that subadditivity is equivalent to $\Phi_\rho = 0$.

Proposition 2.1. *A risk measure is free of regulatory arbitrage if and only if it is subadditive.*

Proof. Suppose that ρ is free of regulatory arbitrage. This implies that $\Psi_\rho(X) = \rho(X)$ for all $X \in \mathcal{X}$. By the definition of Ψ_ρ, one has $\rho(Y_1) + \rho(Y_2) \geqslant \Psi_\rho(X) = \rho(X)$ for all $X, Y_1, Y_2 \in \mathcal{X}$ such that $Y_1 + Y_2 = X$. This implies that ρ is subadditive.

Now suppose that ρ is subadditive. Then for any positive integer n, $\sum_{i=1}^n \rho(Y_i) \geqslant \rho(X)$ for all $X, Y_1, Y_2, \ldots, n \in \mathcal{X}$ such that $\sum_{i=1}^n Y_i = X$. This shows that $\Psi_\rho(X) \geqslant \rho(X)$. Since $\Psi_\rho(X) \leqslant \rho(X)$, we obtain that $\Psi_\rho = \rho$ on $\mathcal{X}$, i.e. ρ is free of regulatory arbitrage. □

Coherent risk measures, such as ES, are free of regulatory arbitrage as expected from their motivation in Artzner et al. (1999). In light of the criticisms of subadditivity, a limited regulatory arbitrage may become desirable if one needs to move away from subadditivity. However, an infinite regulatory arbitrage could lead to serious issues and such situations should be taken with much care. It is thus of our interest to classify risk measures with limited, unlimited or infinite regulatory arbitrage, and quantify the magnitude of this regulatory arbitrage.

In the following we look at the case of VaR, the prominent risk measure in banking practice for the recent few decades. Many academics and practitioners are aware of the fact that VaR is problematic in risk aggregation; it would be of great interest to see how vulnerable this risk measure behaves when a risk can be divided into fragments.

Theorem 2.2. VaR_p *for* $p \in (0,1)$ *is of infinite regulatory arbitrage.*

Proof. Let k be an integer such that $1 - 1/k > p$, and $\{A_i,\ i = 1,\dots,k\}$ be a partition of Ω, with $\mathbb{P}(A_i) = 1/k$, $i = 1,\dots,k$. For each $X \in L^0$, and any number $m > 0$, write

$$Y_i = \mathrm{I}_{A_i} m - \frac{1}{k-1}(1 - \mathrm{I}_{A_i})m.$$

It is easy to see that $\sum_{i=1}^k Y_i = 0$. Note that for each $i = 1,\dots,k$, $\mathbb{P}(Y_i < 0) = \mathbb{P}(A_i^c) = 1 - 1/k > p$. Hence, $\mathrm{VaR}_p(Y_i) = -m/(k-1) < 0$. It follows from $X = X + \sum_{i=1}^k Y_i$ that

$$\Psi_\rho(X) \leqslant \mathrm{VaR}_p(X) + \sum_{i=1}^k \mathrm{VaR}_p(Y_i) = \mathrm{VaR}_p(X) - \frac{mk}{k-1}.$$

Therefore, $\Phi_{\mathrm{VaR}_p}(X) \geqslant \frac{mk}{k-1}$. Since m is arbitrary, we conclude that $\Phi_{\mathrm{VaR}_p} = \infty$, and thus VaR_p is of infinite regulatory arbitrage. $\square$

Theorem 2.2 shows that VaR_p is indeed problematic in view of regulatory arbitrage. Actually, in the proof we have seen that any risky position X can be divided into $k+1$ random variables, yielding an arbitrarily smaller sum of VaR_p evaluated at each fragment, $p < 1 - 1/k$, compared to $\mathrm{VaR}_p(X)$.

Example 2.1 (The vanishing VaR)**.** We give an example to illustrate how a simple splitting vanishes the sum of VaRs regardless of the choice of the risky position X. We take $p = 0.99$ but this value is totally arbitrary. Let X be any positive and continuously distributed random variable. Denote by x_i the $(i/101)$-quantile of the distribution of X, $i = 1,\dots,100$, and in addition let $x_0 = 0$ and $x_{101} = \infty$. Take

$$Y_i = X\mathrm{I}_{\{X \in [x_{i-1}, x_i)\}},\ i = 1,\dots,101.$$

It is clear that $\sum_{i=1}^{101} Y_i = X$ and $\mathrm{VaR}_{0.99}(Y_i) = 0$, which implies that $\sum_{i=1}^{101} \mathrm{VaR}_{0.99}(Y_i) = 0$.

2.3 Theoretical properties

In the next we move to a more general framework, and systematically investigate properties of the two functionals Φ_ρ and Ψ_ρ. In light of the simple relationship $\Phi_\rho = \rho - \Psi_\rho$, we mainly focus our discussions on Ψ_ρ, as the latter is naturally another risk measure with nice properties; see Theorem 2.4 below.

Lemma 2.3. *Suppose that $\rho(0) = 0$. We have that*

$$\Psi_\rho(X) = \lim_{n\to\infty} \inf\left\{\sum_{i=1}^n \rho(Y_i) : Y_i \in \mathcal{X},\ i = 1,\dots,n,\ \sum_{i=1}^n Y_i = X\right\}. \tag{2.1}$$

Proof. It suffices to see that

$$\Psi_\rho^{(n)}(X) := \inf\left\{\sum_{i=1}^n \rho(Y_i) : Y_i \in \mathcal{X},\ i = 1,\dots,n,\ \sum_{i=1}^n Y_i = X\right\} \tag{2.2}$$

is decreasing in n. This can be seen from

$$\Psi_\rho^{(n+1)}(X) \leqslant \inf\left\{\sum_{i=1}^{n+1} \rho(Y_i) : Y_i \in \mathcal{X},\ i = 1,\dots,n,\ \sum_{i=1}^n Y_i = X,\ Y_{n+1} = 0\right\} = \Psi_\rho^{(n)}(X).$$

Thus, in Definition 2.2 , the infimum with respect to $n \in \mathbb{N}$ in Ψ_ρ can be replaced by a limit in n. □

Theorem 2.4. *If a risk measure ρ on $\mathcal{X}$ satisfies any properties (A1-A3, A6, A8) in Section 2.1, then Ψ_ρ inherits the corresponding properties. Moreover,*

(i) Ψ_ρ is always subadditive;

(ii) if ρ is convex and $\rho(0) = 0$, then Ψ_ρ is convex;

(iii) Ψ_ρ is the largest subadditive risk measure dominated by ρ;

(iv) if $\rho \leqslant \tau$ for another risk measure τ on $\mathcal{X}$, then $\Psi_\rho \leqslant \Psi_\tau$.

Proof. It can be directly checked that if ρ is monotone, cash-invariant, positive homogeneous, loaded or law-invariant, Ψ_ρ satisfies the corresponding properties. (iv) is also straightforward. In the following we show (i), (ii) and (iii).

(i) Denote for $k \in \mathbb{N}$, and $X \in \mathcal{X}$,

$$C_k(X) = \{(X_1,\dots,X_k) \in \mathcal{X}^k : X_1 + \cdots + X_k = X\}.$$

Then for any $X, Y \in \mathcal{X}$,

$$
\begin{aligned}
&\Psi_\rho(X+Y) \\
&= \inf\left\{\sum_{i=1}^n \rho(Z_i) : n \in \mathbb{N},\ (Z_1,\dots,Z_n) \in \mathcal{C}_n(X+Y)\right\} \\
&\leqslant \inf\left\{\sum_{i=1}^n \rho(X_i) + \sum_{i=1}^k \rho(Y_i) : n,k \in \mathbb{N}, (X_1,\dots,X_n) \in \mathcal{C}_n(X),\ (Y_1,\dots,Y_k) \in \mathcal{C}_k(Y)\right\} \\
&= \inf\left\{\sum_{i=1}^n \rho(X_i) : n \in \mathbb{N},\ (X_1,\dots,X_n) \in \mathcal{C}_n(X)\right\} \\
&\quad + \inf\left\{\sum_{i=1}^k \rho(Y_i) : k \in \mathbb{N},\ (Y_1,\dots,Y_k) \in \mathcal{C}_k(Y)\right\} \\
&= \Psi_\rho(X) + \Psi_\rho(Y).
\end{aligned}
$$

Hence, Ψ_ρ is subadditive.

(ii) Since $\rho(0) = 0$, we use (2.1). For any $X, Y \in \mathcal{X}$,

$$
\begin{aligned}
&\Psi_\rho(\lambda X + (1-\lambda)Y) \\
&= \lim_{n\to\infty} \inf\left\{\sum_{i=1}^n \rho(Z_i) :\ (Z_1,\dots,Z_n) \in \mathcal{C}_n(\lambda X + (1-\lambda)Y)\right\} \\
&\leqslant \lim_{n\to\infty} \inf\left\{\sum_{i=1}^n \rho(\lambda X_i + (1-\lambda)Y_i) : (X_1,\dots,X_n) \in \mathcal{C}_n(X),\ (Y_1,\dots,Y_n) \in \mathcal{C}_n(Y)\right\} \\
&\leqslant \lim_{n\to\infty} \inf\left\{\sum_{i=1}^n (\lambda\rho(X_i) + (1-\lambda)\rho(Y_i)) : (X_1,\dots,X_n) \in \mathcal{C}_n(X),\ (Y_1,\dots,Y_n) \in \mathcal{C}_n(Y)\right\} \\
&= \lim_{n\to\infty} \inf\left\{\sum_{i=1}^n \lambda\rho(X_i) : (X_1,\dots,X_n) \in \mathcal{C}_n(X)\right\} \\
&\quad + \lim_{n\to\infty} \inf\left\{\sum_{i=1}^n (1-\lambda)\rho(Y_i) : (Y_1,\dots,Y_n) \in \mathcal{C}_n(Y)\right\} \\
&= \lambda\Psi_\rho(X) + (1-\lambda)\Psi_\rho(Y).
\end{aligned}
$$

Hence, Ψ_ρ is convex.

(iii) We first observe that $\rho_0 : \mathcal{X} \to \{-\infty\}$ is a subadditive risk measure, hence there exist subadditive risk measures dominated by ρ. As a consequence, there exists a largest subadditive risk measure ρ_+ dominated by ρ since the supremum of a collection of subadditive risk measures is still subadditive. It follows that $\Psi_\rho \leqslant \rho_+ \leqslant \rho$. Since Ψ_ρ and ρ_+ are both subadditive, we have that $\Psi_{\Psi_\rho} = \Psi_\rho$ and $\Psi_{\rho_+} = \rho_+$. Then we have by (iv) that $\Psi_\rho \leqslant \rho_+ \leqslant \Psi_\rho$, which implies $\Psi_\rho = \rho_+$. □

As the largest subadditive functional dominated by ρ, the minimal measure Ψ_ρ is called a *sub-additive hull* of ρ. Such a construction and its properties are used in convex analysis; see Rockafeller (1970).

Delbaen (2000) introduced the *inf-convolution* of risk measures

$$\rho_1 \Box \rho_2(X) := \inf\{\rho_1(Y_1) + \rho_2(Y_2) : Y_1, Y_2 \in \mathcal{X},\ Y_1 + Y_2 = X\}.$$

One can check that

$$\begin{aligned}
&\rho\Box\rho\Box\rho(X)\\
&= \inf\{\rho(Y_1) + \rho\Box\rho(Z) : Y_1, Z \in \mathcal{X},\ Y_1 + Z = X\}\\
&= \inf\{\rho(Y_1) + \inf\{\rho(Y_2) + \rho(Y_3) : Y_2 + Y_3 = Z,\ Y_2, Y_3 \in \mathcal{X}\} : Y_1, Z \in \mathcal{X},\ Y_1 + Z = X\}\\
&= \inf\{\rho(Y_1) + \rho(Y_2) + \rho(Y_3) : Y_1, Y_2, Y_3 \in \mathcal{X},\ Y_1 + Y_2 + Y_3 = X\}\\
&= \Psi_\rho^{(3)}(X),
\end{aligned}$$

and by induction, we obtain $\Psi_\rho^{(n)}$ in (2.2),

$$\Psi_\rho^{(n)}(X) = \rho \underbrace{\Box \cdots \Box}_{n \text{ times}} \rho. \tag{2.3}$$

By Lemma 2.3, for a risk measure ρ with $\rho(0) = 0$, $\Psi_\rho = \lim_{n\to\infty} \Psi_\rho^{(n)} = \rho\Box\rho\Box\cdots$ is the *self-convolution limit* of ρ.

Remark 2.1. More discussions on the inf-convolution of different convex risk measures and its financial interpretation can be found for instance in Barrieu and El Karoui (2005) and Jouini et al. (2008), in the context of optimal risk transfer. Tsanakas (2009) considered the problem of optimal capital allocation in an insurance group, where the self-convolution of a convex risk measure is used to determine when to stop a portfolio splitting. As far as we are aware, there is no existing research in the literature addressing self-convolution of a general (not necessary convex) risk measure or its limit.

The next example studies the regulatory arbitrage of the *entropic risk measure* ρ, defined as $\rho(X) = \frac{1}{\beta}\log \mathbb{E}[\exp(\beta X)]$, $\beta > 0$, $X \in L^\infty$. The entropic risk measure comes from *indifference utility principle* (Gerber, 1974), and is well-known to be a convex risk measure.

Example 2.2 (Entropic risk measures)**.** Let $\rho(X) = \frac{1}{\beta}\log \mathbb{E}[\exp(\beta X)]$, $\beta > 0$, $X \in L^\infty$. By Jensen's inequality, we have that $\rho(X) \geqslant \mathbb{E}[X]$ for all $X \in L^\infty$, which by Theorem 2.4 (iv) implies that $\Psi_\rho(X) \geqslant$

$\Psi_{\mathbb{E}[\cdot]}(X) = \mathbb{E}[X]$. On the other hand, we have from Lemma 2.3 that

$$\Psi_\rho(X) \leqslant \lim_{n\to\infty} n\rho(X/n) = \lim_{n\to\infty} \frac{1}{\beta} \log((\mathbb{E}[\exp(\beta X/n)])^n) = \lim_{n\to\infty} \frac{1}{\beta} \log((M_X(\beta/n))^n) = \mathbb{E}[X],$$

where $M_X(\cdot)$ is the moment generating function of X. Therefore, we have $\Psi_\rho(X) = \mathbb{E}[X]$ for all $X \in L^\infty$. As a consequence, the regulatory arbitrage $\Phi_\rho(X) = \rho(X) - \mathbb{E}[X] < \infty$. That is, ρ is of limited regulatory arbitrage, but not free of regulatory arbitrage.

A consequence of Theorem 2.4 is that (iii) and (iv) in Definition 2.3 are equivalent.

Proposition 2.5. *A risk measure ρ has unlimited regulatory arbitrage if and only if it has infinite regulatory arbitrage.*

Proof. By definition, an infinite regulatory arbitrage of ρ implies an unlimited regulatory arbitrage. We show the other direction. Suppose that $\Phi_\rho(X) = \infty$ for some $X \in \mathcal{X}$. This implies that $\Psi_\rho(X) = -\infty$ as $\rho(X) < \infty$. For any $Y \in \mathcal{X}$, by Theorem 2.4 (i) and the fact that $\Psi_\rho \leqslant \rho$, we have that $\Psi_\rho(Y) \leqslant \Psi_\rho(X) + \Psi_\rho(Y - X) \leqslant \Psi_\rho(X) + \rho(Y - X) = -\infty$. As a consequence, $\Phi_\rho(Y) = \infty$ and ρ has infinite regulatory arbitrage. □

The following proposition is useful to identify types of regulatory arbitrage of risk measures.

Proposition 2.6. *Suppose that ρ_1 and ρ_2 are two risk measures on $\mathcal{X}$, and $\rho_1(X) \leqslant \rho_2(X)$ for all $X \in \mathcal{X}$.*

(i) If ρ_1 is of limited regulatory arbitrage, then ρ_2 is also of limited regulatory arbitrage.

(ii) If ρ_2 is of infinite regulatory arbitrage, then ρ_1 is also of infinite regulatory arbitrage.

Proof. (i) We have that $\Psi_{\rho_1}(X) > -\infty$ for all $X \in \mathcal{X}$. By Theorem 2.4 (iv), $\Psi_{\rho_2}(X) \geqslant \Psi_{\rho_1}(X) > -\infty$. This implies that ρ_2 is not of infinite regulatory arbitrage, and by Proposition 2.5, ρ_2 is of limited regulatory arbitrage.

(ii) We have that $\Psi_{\rho_2}(X) = -\infty$ for all $X \in \mathcal{X}$. By Theorem 2.4 (iv), $\Psi_{\rho_1}(X) \leqslant \Psi_{\rho_2}(X) = -\infty$ for all $X \in \mathcal{X}$. This implies that ρ_1 is of infinite regulatory arbitrage. □

Proposition 2.6 suggests that any risk measure that is dominated by a VaR_p for some $p \in (0, 1)$ is of infinite regulatory arbitrage. Note that if ρ_1 is free of regulatory arbitrage, and $\rho_2 \geqslant \rho_1$, then ρ_2 is of limited regulatory arbitrage but not necessarily free of regulatory arbitrage in general.

In Artzner et al. (1999), four properties were proposed for a coherent risk measure: monotonicity (A1), cash-invariance (A2), positive homogeneity (A3) and subadditivity (A4). Note that Ψ_ρ is subadditive and it inherits (A1)-(A3) from ρ; hence Ψ_ρ is coherent if ρ is monetary and positive homogeneous, and Ψ_ρ is the largest coherent risk measure dominated by ρ. The following proposition gives a simple criterion for an infinite regulatory arbitrage of ρ.

Proposition 2.7. *A positive homogeneous risk measure ρ is of infinite or unlimited regulatory arbitrage if and only if $\Phi_\rho(0) > 0$.*

Proof. In Proposition 2.5 we have seen that infinite regulatory arbitrage and unlimited regulatory arbitrage are equivalent. Obviously, $\Phi_\rho(0) > 0$ is required for ρ to be of infinite regulatory arbitrage. Next we show the other direction. Suppose that $\Phi_\rho(0) > 0$ and note that $\rho(0) = 0$ for positive homogeneous risk measures. Since $\Psi_\rho(0) < 0$, there exist $n \in \mathbb{N}$ and $Y_i \in \mathcal{X}$, $i = 1, \dots, n$, such that $\sum_{i=1}^n Y_i = 0$ and $\Psi_\rho(0) \leqslant \sum_{i=1}^n \rho(Y_i) < 0$. For any $c > 0$, we have that

$$\Psi_\rho(0) \leqslant \sum_{i=1}^n \rho(cY_i) \leqslant c \sum_{i=1}^n \rho(Y_i).$$

Thus, $\Psi_\rho(0) = -\infty$ and ρ is of unlimited regulatory arbitrage. □

3 Regulatory arbitrage of commonly-used risk measures

In this section, we study regulatory arbitrage of distortion risk measures and convex risk measures, two most commonly studied and most practical classes of risk measures. To use classic duality results in risk measures, we restrict $\mathcal{X} = L^\infty$ as in most of the literature. For risk measures on sets beyond L^∞, see Kaina and Rüschendorf (2009) and Filipović and Svindland (2012).

3.1 Regulatory arbitrage of distortion risk measures

A *distortion risk measure* is defined as

$$\rho : L^\infty \to \mathbb{R}, \quad \rho(X) = \int_{\mathbb{R}} x \mathrm{d}h(\mathbb{P}(X \leqslant x)), \tag{3.1}$$

where h is a non-decreasing function with $h(0) = 0$ and $h(1) = 1$. h is called a *distortion function*. The functional (3.1) was introduced in economic decision theory by Yaari (1987), in insurance pricing by Wang et al. (1997) and in banking risk measurement by Acerbi (2002) (who focuses on the case of convex h) under the name *spectral risk measures*. A distortion risk measure is always monetary, positive homogeneous, comonotonic additive and law-invariant; further more it is coherent if and only if its distortion function h is convex; see Yaari (1987, Theorem 2). It is also the only class of monetary and law-invariant risk measures that are comonotonic additive.

The following theorem characterizes the regulatory arbitrage of distortion risk measures.

Theorem 3.1. *A distortion risk measure ρ with distortion function h is*

(i) free of regulatory arbitrage if and only if it is coherent;

(ii) is of limited regulatory arbitrage if and only if it is loaded;

(iii) is of infinite regulatory arbitrage if and only if it is not loaded.

Moreover, the minimal measure Ψ_ρ is given by

$$\Psi_\rho(X) = \sup_{\mu \in \mathcal{Q}} \rho_\mu(X), \quad X \in L^\infty, \tag{3.2}$$

where ρ_μ is a distortion risk measure with distortion function μ, $\mathcal{Q}$ is the set of all increasing convex functions μ on $[0, 1]$ such that $\mu(0) = 0$, $\mu(1) = 1$, and $\mu(x) \geqslant h(x)$ for all $x \in [0, 1]$.

Proof. We first show (3.2) and then (i)-(iii). By Theorem 2.4, Ψ_ρ is a law-invariant coherent risk measure. By Kusuoka's representation of law-invariant coherent risk measures on L^∞ (Kusuoka, 2001; Jouini et al., 2006), there exists a compact convex set $\mathcal{P}$ of probability measures on $[0, 1]$, such that for all $X \in L^\infty$,

$$\Psi_\rho(X) = \sup_{\mu \in \mathcal{P}} \int_0^1 \mathrm{ES}_p(X) \mathrm{d}\mu(p).$$

Since $\Psi_\rho \leqslant \rho$, we have $\mathcal{P} \subset \mathcal{Q}$, and hence, for all $X \in L^\infty$,

$$\Psi_\rho(X) \leqslant \sup_{\mu \in \mathcal{Q}} \int_0^1 \mathrm{ES}_p(X) \mathrm{d}\mu(p).$$

On the other hand, by Theorem 2.4 (iii), Ψ_ρ is the largest coherent risk measure that is dominated by ρ, and therefore $\sup_{\mu \in \mathcal{Q}} \int_0^1 \mathrm{ES}_p(X) \mathrm{d}\mu(p) \leqslant \Psi_\rho(X)$. In summary, we obtain $\sup_{\mu \in \mathcal{Q}} \int_0^1 \mathrm{ES}_p(X) \mathrm{d}\mu(p) = \Psi_\rho(X)$, that is, $\Psi_\rho(X)$ is the supremum of all coherent distortion risk measures dominated by ρ. Note that $\rho_\mu \leqslant \rho$ on L^∞ is equivalent to $\mu \geqslant h$ on $[0, 1]$ (see Lemma A.1 of Wang et al., 2014). Thus (3.2) follows.

(i) is obvious in light of Proposition 2.1. (ii) and (iii) imply each other; as such we will only show (ii). Suppose that ρ is loaded. It follows that

$$0 = \rho(0) \geqslant \Psi_\rho(0) \geqslant \mathbb{E}[0] = 0.$$

It follows that $\Phi(0) = \rho(0) - \Psi_\rho(0) = 0$, and by Proposition 2.7, ρ is of limited regulatory arbitrage. Now suppose that ρ is of limited regulatory arbitrage, and consequently Q in (3.2) is not empty. Since $\int_0^1 \mathrm{ES}_p(X)\mathrm{d}\mu(p) \geqslant \mathrm{ES}_0(X) = \mathbb{E}[X]$ for all probability measures μ on $[0,1]$, we have that $\rho(X) \geqslant \Psi_\rho(X) \geqslant \mathbb{E}[X]$ for all $X \in L^\infty$, and hence ρ is loaded. □

Note that in the proof of Theorem 3.1, we can see that (3.2) holds for all law-invariant, positive homogeneous monetary risk measures. Theorem 3.1 suggests that there are serious issues with respect to regulatory arbitrage for non-loaded distortion risk measures, such as VaR_p. If for some reason one wants to weaken the assumption of subadditivity, then a loaded non-coherent risk measure is a reasonable alternative. We hereby highlight the importance of loadedness, which we believe have been possibly neglected in the literature of risk measures interpreted as capital requirement principles, although it has always been a key issue to consider in insurance pricing. An example of loaded but not coherent distortion risk measure can be found in a mixture of VaR_p and ES_p; see Section 5.3.

As an immediate corollary, we have the following characterization using distortion functions.

Corollary 3.2. *Let ρ be a distortion risk measure with distortion function h.*

(i) ρ is free of regulatory arbitrage if and only if h is convex on $[0,1]$.

(ii) ρ is of limited regulatory arbitrage if and only if $h(x) \leqslant x$ for all $x \in$ $[0,1]$.

(iii) ρ is of infinite regulatory arbitrage if and only if $h(x) > x$ for some $x \in$ $[0,1]$.

Remark 3.1. Even when ρ is a distortion risk measure, Ψ_ρ is not necessarily a distortion risk measure, as illustrated by the following simple example. Suppose h_1 and h_2 are two convex distortion functions with continuous derivatives, such that $h_1(1/2) = h_2(1/2)$, $h_1'(1/2) > h_2'(1/2)$, $h_1(t) < h_2(t)$ for $t \in [0,1/2)$ and $h_1(t) > h_2(t)$ for $t \in (1/2,1]$. Let $h(t) = \min\{h_1(t), h_2(t)\}$ for $t \in [0,1]$. It is easy to see that h is dominated by h_1 and h_2, and h is no longer convex. Let ρ_1, ρ_2 be two distortion risk measures with distortion functions h_1 and h_2, respectively. Since $h \leqslant h_1$ and $h \leqslant h_2$, we have that $\rho_1 \leqslant \rho$ and $\rho_2 \leqslant \rho$ (see Wang et al., 2014, Lemma A.1 (b)). Since Ψ_ρ is the largest coherent risk measure dominated by ρ, $\rho_1 \leqslant \Psi_\rho \leqslant \rho$ and $\rho_2 \leqslant \Psi_\rho \leqslant \rho$ holds. If Ψ_ρ is a distortion risk measure, we would have that the distortion function h^* of Ψ_ρ satisfies $h_1 \geqslant h^* \geqslant h$ and $h_2 \geqslant h^* \geqslant h$, leading to $h^* = h$, which contradicts the coherence of Ψ_ρ.

3.2 Regulatory arbitrage of convex risk measures

Following the terminology used in Föllmer and Schied (2011), a *convex risk measure* is a risk measure which is monotone, cash-invariant, and convex. The following additional property is often imposed for a robust representation (Föllmer and Schied, 2002) of convex risk measures ρ.

(A9) Fatou property: $\liminf_{n\to\infty} \rho(X_n) \geqslant \rho(X)$ if $X, X_1, X_2, \cdots \in L^\infty$, $X_n \to X$ $\mathbb{P}$-almost surely and $\sup_{n\in\mathbb{N}} |X_n| < Y$ for some $Y \in L^\infty$.

A risk measure ρ is a convex risk measure with the Fatou property if and only if it has the following representation (Föllmer and Schied, 2002; Frittelli and Rosazza Gianin, 2002): there exists a convex and lower-semicontinuous (with respect to the total variation distance on $\mathcal{P}$) function $v : \mathcal{P} \to [0, \infty]$, where $\mathcal{P}$ is the set of all probability measures on Ω absolutely continuous with respect to $\mathbb{P}$, such that

$$\rho(X) = \sup_{Q\in\mathcal{P}} \{\mathbb{E}^Q[X] - v(Q)\}, \quad X \in L^\infty. \tag{3.3}$$

The convex function v is called the *penalty function* of ρ and can be chosen as the dual of (3.3):

$$v(Q) = \sup_{X\in L^\infty} \{\mathbb{E}^Q[X] - \rho(X)\}, \quad Q \in \mathcal{P}. \tag{3.4}$$

Below we state a useful result about the inf-convolution of two convex risk measures, given in Barrieu and El Karoui (2005).

Lemma 3.3 (Theorem 3.6 of Barrieu and El Karoui (2005))**.** *Suppose ρ_1, ρ_2 are two convex risk measures on L^∞, with penalty functions v_1, v_2, respectively. Then, the penalty function of $\rho_1 \square \rho_2$ is given by $v = v_1 + v_2$.*

Using Lemma 3.3, we obtain the following result on the minimal measure of convex risk measures, where we impose a condition $\rho(0) = 0$ which is common when risk measures are interpreted as capital requirements.

Proposition 3.4. *Suppose that ρ is a convex risk measure on L^∞ with the Fatou property and penalty function v, $\rho(0) = 0$, and $\mathcal{Q} = \{Q \in \mathcal{P} : v(Q) = 0\}$ is non-empty. Then Ψ_ρ is a coherent risk measure given by*

$$\Psi_\rho(X) = \sup_{Q\in\mathcal{Q}} \mathbb{E}^Q[X], \quad X \in L^\infty. \tag{3.5}$$

Proof. First, observe that $\rho(0) = 0$ implies that $v(Q) \geqslant 0$ for all $Q \in \mathcal{P}$. By (2.1), we have that $\Psi_\rho = \lim_{n\to\infty} \Psi_\rho^{(n)}$. From Lemma 3.3, it follows that $\Psi_\rho^{(n)}$ has penalty function nv. By a simple monotone convergence argument, we can verify that Ψ_ρ has penalty function v_0 where for $Q \in \mathcal{P}$,

$$v_0(Q) = \lim_{n\to\infty} nv(Q) = \begin{cases} 0, & Q \in \mathcal{Q}, \\ \infty, & Q \notin \mathcal{Q}. \end{cases}$$

The proposition follows. □

Note that the optimal splitting for convex risk measures in this context is the equally splitting; see for instance Tsanakas (2009, Section 4.3). We can see that by splitting the risky position X into n equal pieces, the "super-linear behavior" in convex risk measures is weakened and eventually disappears as $n \to \infty$.

Remark 3.2. Example 2.2 fits into Proposition 3.4. The result on the largest coherent risk measure dominated by a convex risk measure can also be found in Föllmer and Schied (2011, Corollary 4.19).

In order to guarantee a convex risk measure is of limited regulatory arbitrage, we use the following Lebesgue property, which is stronger than the Fatou property.

(A10) Lebesgue property: $\lim_{n\to\infty} \rho(X_n) = \rho(X)$ if $X, X_1, X_2, \dots \in L^\infty$, $X_n \to X$ $\mathbb{P}$-almost surely and $\sup_{n\in\mathbb{N}} |X_n| < Y$ for some $Y \in L^\infty$.

Corollary 3.5. *A convex risk measures ρ on L^∞, with the Lebesgue property and $\rho(0) = 0$, is of limited regulatory arbitrage, and it is free of regulatory arbitrage if and only if it is coherent.*

Proof. To show that ρ is of limited regulatory, it suffices to show that it is not of infinite regulatory arbitrage, that is, $\Psi_\rho(X) > -\infty$ for some $X \in L^\infty$. This is equivalent to that $\mathcal{Q}$ in (3.5) is non-empty. Note that $\inf\{v(Q) : Q \in \mathcal{P}\} = -\rho(0) = 0$. By Corollary 4.35 of Föllmer and Schied (2011), the Lebesgue property implies that the above infimum is a minimum, and hence $\mathcal{Q}$ in is non-empty. If ρ is free of arbitrage, then it is equal to Ψ_ρ, which is a coherent risk measure. □

Remark 3.3. If ρ is a law-invariant convex risk measure on L^∞ and $\rho(0) = 0$, then a law-invariant representation is available, leading to the Kusuoka's representation of Ψ_ρ:

$$\Psi_\rho(X) = \sup_{\mu\in\mathcal{Q}} \int_0^1 \mathrm{ES}_p(X)\mathrm{d}\mu(p), \quad X \in L^\infty.$$

where $\mathcal{Q} = \{Q \in \mathcal{P}([0,1]) : v(Q) = 0\}$, $\mathcal{P}([0,1])$ is the set of all probability measures on $[0,1]$, and $v : \mathcal{P}([0,1]) \to [0,\infty]$ is the law-invariant penalty function of ρ; see Frittelli and Rosazza Gianin (2005). The Fatou property is automatically satisfied by law-invariant convex risk measures on L^∞; see Jouini et al. (2006). More interestingly, since a law-invariant convex risk measure ρ with $\rho(0) = 0$ is always loaded, by Proposition 2.6 we know that ρ is of limited regulatory arbitrage. Clearly, the Lebesgue property in Corollary 3.5 is a sufficient but not necessary condition.

Remark 3.4. The results obtained in this section also are valid for risk measures on spaces beyond L^∞, as long as $\rho(X) \in \mathbb{R}$ is satisfied and a proper continuity condition (the Fatou property in L^p space) is imposed so that the representation (3.3) holds. We refer to Kaina and Rüschendorf (2009) and Filipović and Svindland (2012) for convex risk measures on L^p.

4 Regulatory arbitrage with transaction cost

In practice, financial institutions may not be allowed to divide its risks arbitrarily without being penalized from either a regulatory consideration or an operational consideration. Suppose that a penalty (or fee) of $f(n) \in [0,\infty]$, $f(1) = 0$ (typically increasing in n) is added to the capital requirement of the financial institution for a division of n fragments. In this case, we can define the *minimal measure with cost f*

$$\Psi_{\rho,f}(X) = \inf\left\{\sum_{i=1}^n \rho(Y_i) + f(n) : n \in \mathbb{N},\ Y_i \in \mathcal{X},\ i = 1,\dots,n,\ \sum_{i=1}^n Y_i = X\right\},$$

and its *regulatory arbitrage with cost f*

$$\Phi_{\rho,f}(X) = \rho(X) - \Psi_{\rho,f}(X).$$

With cost f added, we use the same terminologies for different classes of regulatory arbitrage as in Definition 2.3.

Theorem 4.1. *Suppose that ρ is a convex risk measure with the Fatou property and penalty function v.*

(i) We have that

$$\Psi_{\rho,f}(X) = \inf_{n\in\mathbb{N}}\{\rho_n(X)\}, \tag{4.1}$$

where ρ_n is a convex risk measure with penalty function $nv - f(n)$.

(ii) ρ is of limited regulatory arbitrage with cost f if $\rho(0) \geqslant 0$.

(iii) ρ is free of regulatory arbitrage with cost f if and only if $f(n) \geqslant (n-1)v(Q)$ for all $n \in \mathbb{N}$ and all $Q \in \mathcal{P}$ with $v(Q) < \infty$.

Proof. (i) It is straightforward that

$$\Psi_{\rho,f}(X) = \inf_{n\in\mathbb{N}}\{\Psi_{\rho}^{(n)}(X) + f(n)\}.$$

By Lemma 3.3, $\Psi_{\rho}^{(n)}$ is a convex risk measure with penalty function nv, and (4.1) follows.

(ii) We observe that $\Psi_{\rho,f}(X) \geqslant \Psi_{\rho}(X)$, and

$$\Psi_{\rho}^{(n)}(0) = \sup_{Q\in\mathcal{P}}\{-nv(Q)\} = n\rho(0).$$

As a consequence, $\Psi_{\rho}(0) \geqslant 0$. By Proposition 2.5, we obtain that ρ is of limited regulatory arbitrage, and hence $\Psi_{\rho}(X) > -\infty$ for all $X \in L^{\infty}$. As a consequence, $\Psi_{\rho,f}(X) > -\infty$ for all $X \in L^{\infty}$, and ρ has limited regulatory arbitrage.

(iii) That ρ is free of regulatory arbitrage with cost f is equivalent to that $\Psi_{\rho,f}(X) \geqslant \rho(X)$ for all $X \in \mathcal{L}^{\infty}$. That is, $\rho_n(X) \geqslant \rho(X)$ for all $X \in L^{\infty}$ and $n \in \mathbb{N}$. Note that in general, for two convex risk measures ρ' and ρ'' on L^{∞} with penalties v' and v'', respectively, $\rho' \leqslant \rho''$ is equivalent to $v' \geqslant v''$; this can be seen, for instance, from Föllmer and Schied (2011, Theorem 4.16). Since ρ_n and ρ are both convex risk measures,

$$\rho_n(X) \geqslant \rho(X) \text{ for all } X \in L^{\infty} \Leftrightarrow nv(Q) - f(n) \leqslant v(Q) \text{ for all } Q \in \mathcal{P}.$$

Thus, we obtain that

$$f(n) \geqslant (n-1)v(Q) \text{ for all } Q \in \mathcal{P} \text{ with } v(Q) < \infty \Leftrightarrow \rho_n(X) \geqslant \rho(X) \text{ for all } X \in L^{\infty},$$

and (iii) follows. □

Theorem 4.1 implies that if a penalty with a proper magnitude is applied to capital requirement with a convex risk measure, then regulatory arbitrage can be eliminated. This is consistent with our intuition: a financial institution would not have the incentive to split a risk if a relatively large cost is applied to the splitting. Note that here this cost only depends on ρ, but not on the underlying risk X. This creates convenience in the design of an external regulation. In particular, if ρ is a coherent risk measure, then $v(Q)$ is either 0 or ∞, and therefore $f(n) \geqslant (n-1)v(Q)$ always holds for all $n \in \mathbb{N}$ and all $Q \in \mathcal{P}$ with $v(Q) < \infty$. That is, if a coherent risk measure is implemented for calculating capital requirement, no transaction cost is necessary to eliminate regulatory arbitrage.

In practice, a financial institution may be confined to an operation of at most k subsidies. In this case, we have a special case of cost f: $f(n) = \infty$ for $n > k$ and $f(n) = 0$ for $n \leqslant k$. The corresponding minimal measure is given by

$$\Psi_{\rho,f}(X) = \inf_{1 \leqslant n \leqslant k} \Psi_{\rho}^{(n)}(X).$$

Remark 4.1. In Theorem 4.1 (ii), an if-and-only-if statement can also be given with more detailed analysis. Since $\rho(0) \geqslant 0$ is assumed in all practical cases, we skip the technical discussions on the cases when $\rho(0) < 0$.

One may wonder what would happen if a transaction cost is applied to capital requirement with VaR. Since $\Psi^{(k)}_{\mathrm{VaR}_p}(X) = -\infty$ for all $X \in L^0$ and $k > 1/(1-p)$, we simply obtain $\Psi_{\mathrm{VaR}_p,f}(X) = -\infty$ for all f with $f(k) < \infty$. That is, as long as f is finite, VaR_p is of infinite regulatory arbitrage with cost f. If $f(n) = \infty$ for n larger than some k, VaR_p may have limited regulatory arbitrage. We illustrate this fact in the following example.

Example 4.1. Suppose that $f(1) = f(2) = 0$ and $f(n) = \infty$ for $n > 2$, that is, splitting to more than two fragments is not allowed. For simplicity we consider $\mathcal{X} = L^\infty$. With this cost function f, VaR_p is of limited regulatory arbitrage for $p > 1/2$. To see this, note that for any random variable X, it holds that for all $\varepsilon > 0$,

$$\mathrm{VaR}_p(-X) = \inf\{x : \mathbb{P}(-X \leqslant x) \geqslant p\} = -\inf\{y : \mathbb{P}(X < y) > 1-p\} \geqslant -\mathrm{VaR}_{1-p+\varepsilon}(X)$$

Thus, by taking ε small enough, we have that $p > 1 - p + \varepsilon$ and

$$\mathrm{VaR}_p(X) + \mathrm{VaR}_p(-X) = \mathrm{VaR}_p(X) - \mathrm{VaR}_{1-p+\varepsilon}(X) \geqslant 0.$$

This implies that

$$\Psi_{\mathrm{VaR}_p,f}(0) = \inf\{\mathrm{VaR}_p(X) + \mathrm{VaR}_p(-X) : X \in L^\infty\} \geqslant 0.$$

It is obvious that $\Psi_{\mathrm{VaR}_p,f}$ is cash-invariant and monotone. For any $K \in \mathbb{R}$, it follows that $\Psi_{\mathrm{VaR}_p,f}(X) \geqslant K$ for all $X \geqslant K$. Hence, $\Phi_{\mathrm{VaR}_p,f}(X) < \infty$ for all $X \in L^\infty$.

Remark 4.2. The setup of regulatory arbitrage with transaction cost is worth further exploration. For a general non-convex risk measure ρ that is of infinite regulatory arbitrage (without transaction cost), it is unclear how one can design a cost function f such that ρ is of limited regulatory arbitrage with this cost function. Moreover, the determination of regulatory arbitrage with transaction cost for general distortion risk measures raises considerable technical challenges, and it is left for future work.

Remark 4.3. The idea of transaction costs in this context is briefly explored by Tsanakas (2009). Impediments to fragmentation do not only exist in terms of transaction costs discussed here but also in that some particular forms of splitting may be inadmissible for regulatory or operational reasons. For example, in an insurance group, liabilities can be split across legal entities via intra-group risk transfers. However, the split cannot be arbitrary, in order to avoid moral hazard issues; see Asimit et al. (2013) and Gatzert and Schmeiser (2011).

5 Some examples

In this section we discuss a few examples of non-coherent risk measures with limited regulatory arbitrage. Many of the examples arise naturally in the context of regulatory capital, risk pricing or insurance premiums.

5.1 Worst-case risk measures

Motivated by the debates on the advantages and shortcomings of both VaR and ES, it is natural to consider a risk measure that is a combination of VaR and ES. The first way to combine two risk measures is to consider the maximum of the two, that is, for $0 < q < p < 1$, define

$$W_{p,q}(X) = \max\{\mathrm{VaR}_p(X), \mathrm{ES}_q(X)\}, \quad X \in L^1.$$

The acceptance set (Artzner et al., 1999) of $W_{p,q}$ is the intersection of those of $\mathrm{VaR}_p(X)$ and $\mathrm{ES}_q(X)$; a simple explanation is that a risky position is accepted by a company only if it is accepted by several evaluators (for instance, a manager and a shareholder), each using a different risk measure. More generally, $W_{p,q}$ is a *worst-case* risk measure, defined as the maximum (or supremum) over a class of risk measures. See Delbaen (2000, Section 4.3) for more details on the maximum of risk measures.

We will see that $W_{p,q}$ is of limited regulatory arbitrage. In fact, we have the following stronger result, directly obtained from Proposition 2.6.

Corollary 5.1. *Suppose that $\rho_1, \dots, \rho_n$ are n risk measures on $\mathcal{X}$, and let*

$$W_n(X) = \max\{\rho_1(X), \dots, \rho_n(X)\}, \quad X \in \mathcal{X}.$$

If one of $\rho_1, \ldots, \rho_n$ is of limited regulatory arbitrage, then W_n is of limited regulatory arbitrage.

This observation justifies that, although $W_{p,q}$ loses the subadditivity of ES_q, the statistical advantages (such as elicitability and robustness with respect to the weak topology) of VaR_p, and is no longer a distortion (hence loses comonotonic additivity as well), it has limited regulatory arbitrage and is sometimes argued as a reasonable choice for quantifying capital requirement.

5.2 Best-case risk measures

As opposed to the worst-case risk measures, *best-case* risk measure is the minimum (or infimum) over a collection of risk measures. Although the intuition behind the best-case risk measures does not fit exactly well the context of regulatory capital requirements, it can be interpreted as the pricing of risks: suppose n agents are evaluating an acceptable price for a short position in a risky asset X using different risk measures $\rho_1, \ldots, \rho_n$, then the lowest value among $\rho_1(X), \ldots, \rho_n(X)$ would be the price of risk. A best-case risk measure can be defined as

$$B_n(X) = \min\{\rho_1(X), \ldots, \rho_n(X)\}, \quad X \in \mathcal{X}.$$

For simplicity, we consider the case when $\rho_1, \ldots, \rho_n$ are distortion risk measures.

Corollary 5.2. *Suppose that $\rho_1, \ldots, \rho_n$ are distortion risk measures, then B_n is of limited regulatory arbitrage if and only if all of $\rho_1, \ldots, \rho_n$ are loaded.*

Proof. First, that all of $\rho_1, \ldots, \rho_n$ are loaded is equivalent to that B_n is loaded. Suppose that B_n is loaded, and so is Ψ_{B_n} from Theorem 2.4. It follows that $\Psi_{B_n}(0) \geqslant 0 = B_n(0)$ and hence $\Phi(0) = 0$. By Proposition 2.7, B_n is of limited regulatory arbitrage.

Suppose that B_n is of limited regulatory arbitrage. Then by Proposition 2.6, each of $\rho_1, \ldots, \rho_n$ is of limited regulatory arbitrage. Hence they are loaded by Theorem 3.1. □

5.3 Mixture risk measures

Another way to combine two risk measures is through a mixture. We defined the Mixture-Value-at-Risk (MVaR) for $p, q \in (0, 1)$, as

$$\mathrm{MVaR}_p^q(X) = q\mathrm{VaR}_p(X) + (1-q)\mathrm{ES}_p(X), \quad X \in L^1.$$

The risk measure MVaR is a mixture of $\mathrm{VaR}_p(X)$ and $\mathrm{ES}_p(X)$, and hence it is also a distortion risk measure. Similar to $W_{p,q}$, MVaR_p^q loses both the subadditivity of ES_p and the statistical advantages of VaR_p. However, MVaR_p^q is often loaded and hence it has limited regulatory arbitrage, as shown below.

Proposition 5.3. MVaR_p^q *is loaded if and only if* $q \leqslant p$*.*

Proof. The distortion function of MVaR_p^q is given by

$$h(t) = q\mathrm{I}_{\{t \geqslant p\}} + (1-q)\mathrm{I}_{\{t \geqslant p\}}\frac{t-p}{1-p}, \quad t \in [0,1].$$

Since $h(t)$ is linear on $[p, 1]$, it is easy to check that $h(t) \leqslant t$, $t \in [0, 1]$ if and only if $q \leqslant p$. □

As a combination of VaR and ES at the same level, MVaR_p^q generally provides a balance between the two most popular risk measures. Since p is typically close to 1, the condition $q \leqslant p$ is usually satisfied. For more flexibility, the two confidence levels of VaR and ES in the Mixture-Value-at-Risk may be chosen differently.

5.4 Risk adjusted value

We give yet another example in the context of evaluating regulatory requirements under the Swiss Solvency Test and Solvency II (FOPI, 2006; European Commission, 2009). Define the *risk adjusted capital* (RAC) of a risk X using a distortion risk measure ρ, as

$$\mathrm{RAC}(X) = \rho(X) - \mathbb{E}[X].$$

The RAC is the least amount of additional capital needed to prevent a company's insolvency. It is often interpreted as non-negative, and hence $\max\{\rho(X) - \mathbb{E}[X], 0\}$ is used instead in practice. The taker of a liability is compensated by receiving the expected value of future claims and funds equal to the cost of raising the necessary regulatory capital to support the liability, that is, one uses the *risk adjusted value*

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

for some $\lambda > 0$ (sometimes called the *cost-of-capital rate*) to determine the total necessary capital for a liability X. A detailed analysis of the cost-of-capital approach for insurance can be found in Wüthrich et al. (2010). It is immediate that ρ dominates $\mathbb{E}[X]$, which is free of regulatory arbitrage, and therefore by Proposition 2.6 ρ is of limited regulatory arbitrage.

5.5 Esscher premium

The last example we would like to mention is the Esscher premium (EP), introduced by Bühlmann (1980) with a motivation from an economic equilibrium, as a Pareto-optimal solution to a market situation with independent risks and agents with exponential utilities. The Esscher premium is defined as

$$\mathrm{EP}_h(X) = \frac{\mathbb{E}[Xe^{hX}]}{\mathbb{E}[e^{hX}]}, \quad X \in L^\infty, \quad h > 0.$$

As a risk measure, EP_h is not monotone, positive homogeneous or convex. We can see that

$$\frac{\mathbb{E}[Xe^{hX}]}{\mathbb{E}[e^{hX}]} \geqslant \frac{\mathbb{E}[X]\mathbb{E}[e^{hX}]}{\mathbb{E}[e^{hX}]} = \mathbb{E}[X],$$

and hence EP_h is of limited regulatory arbitrage, if it is interpreted as a capital requirement principle.

6 Conclusion

We introduced regulatory arbitrage of risk measures for calculating capital requirements. Risk measures generally fall into three categories: free of regulatory arbitrage, of limited regulatory arbitrage, and of infinite regulatory arbitrage. If a risk measure is of infinite regulatory arbitrage, it is vulnerable with respect to manipulation of risks, a concern raised in Artzner et al. (1999) and discussed extensively since then. We showed that VaR is of infinite regulatory arbitrage, while many distortion risk measures and convex risk measures are of limited regulatory arbitrage. Our results partially support the use of ES in financial risk management; in particular, they add to the discussion on VaR and ES raised by the Basel Committee on Banking Supervision.

References

Acerbi, C. Spectral measures of risk: A coherent representation of subjective risk aversion. *J. Bank. Financ.*, 2002, **26**(7), 1505–1518.

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

Asimit, A. V., Badescu, A. M. and Tsanakas, A. Optimal risk transfers in insurance groups. *Eur. Actuar. J.* 3(1), 2013, 159–190.

Barrieu, P. and El Karoui, N. Inf-convolution of risk measures and optimal risk transfer. *Finance Stoch.*, 2005, **9**, 269–298.

BCBS. *Consultative Document May 2012. Fundamental review of the trading book.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements, 2012. Available at http://www.bis.org/publ/bcbs219.htm (accessed 20 December 2014).

BCBS. *Consultative Document October 2013. Fundamental review of the trading book: A revised market risk framework.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements, 2013. Available at http://www.bis.org/publ/bcbs265.htm (accessed 20 December 2014).

Bellini, F. and Bignozzi, V. On elicitable risk measures. Forthcoming in *Quant. Finance*, 2015.

Bühlmann, H. An economic premium principle. *ASTIN Bull.*, 1980, **11**, 52–60.

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

Davis, M. H. A. Consistency of risk measures estimates. *Preprint*, 2013. Available at SSRN http://ssrn.com/abstract=2342279 (accessed 20 December 2014).

Delbaen, F. *Coherent risk measures*. Lecture notes, Scuola Normale Superiore di Pisa, 2000. Available at http://www.math.ethz.ch/~delbaen (accessed 20 December 2014).

Delbaen, F. *Monetary utility functions*, 2012. (Osaka University Press).

Embrechts, P., Puccetti, G. and Rüschendorf, L. Model uncertainty and VaR aggregation. *J. Bank. Financ.*, 2013, **37**(8), 2750–2764.

Embrechts, P., Puccetti, G., Rüschendorf, L., Wang, R. and Beleraj, A. An academic response to Basel 3.5. *Risks*, 2014, **2**(1), 25–48.

Embrechts, P., Wang, B. and Wang, R. Aggregation-robustness and model uncertainty of regulatory risk measures. Forthcoming in *Finance Stoch.*, 2015.

Emmer, S., Kratz, M. and Tasche, D. What is the best risk measure in practice? A comparison of standard measures. *Preprint*, 2014. Available at http://ssrn.com/abstract=2370378 (accessed 21 February 2015).

European Commission. Directive 2009/138/EC of the European Parliament and of the Council of 25 November 2009 on the taking-up and pursuit of the business of insurance and reinsurance (Solvency II), *Official Journal of the European Union*, 2009, L335.

FOPI (2006). *The Swiss experience with market consistent technical provisions - the cost of capital approach*, Federal Office of Private Insurance, 2006. Available at http://www.finma.ch/archiv/bpv/e/themen/00506/00552/00727/index.html (accessed 20 December 2014).

Frittelli M. and Rosazza Gianin E. Putting order in risk measures. *J. Bank. Financ.*, 2002, **26**, 1473–1486.

Frittelli M. and Rosazza Gianin E. Law invariant convex risk measures. *Adv. Mathe. Econ.*, 2005, **7**, 33–46.

Filipović, D. and Svindland, G. The canonical model space for law-invariant convex risk measures is L^1. *Math. Finance*, 2012, **22**(3), 585–589.

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

Föllmer, H. and Schied, A. *Stochastic Finance: An Introduction in Discrete Time,* 2011 (Walter de Gruyter, Third Edition).

Gatzert, N. and Schmeiser H. On the risk situation of financial conglomerates: does diversification matter? *Financial Markets and Portfolio Management*, **25**(1), 2011, 3–26.

Gerber, H. U. On additive premium calculation principles. *ASTIN Bull.*, 1974, **7**(3), 215–222.

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

IAIS (2014). *Consultation Document December 2014. Risk-based global insurance capital standard*. International Association of Insurance Supervisors. Available at http://iaisweb.org/index.cfm?event=getPage&nodeId=25319 (accessed 21 February 2015).

Jouini, E., Schachermayer, W. and Touzi, N. Law invariant risk measures have the Fatou property. *Adv. Math. Econ.*, 2006, **9**, 49–71.

Jouini, E., Schachermayer, W. and Touzi, N. Optimal risk sharing for law invariant monetary utility functions. *Math. Finance*, 2008, **18**(2), 269–292.

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

Kerkhof, J. and Melenberg, B. Backtesting for risk-based regulatory capital. *J. Bank. Financ.*, 2004, **28**(8), 1845-1865.

Kou, S. and Peng, X. On the measurement of economic tail risk. *Preprint*, 2014. Available at http://arxiv.org/abs/1401.4787 (accessed 21 February 2015).

Krätschmer, V., Schied, A. and Zähle, H. Comparative and quantitiative robustness for law-invariant risk measures. *Finance Stoch.*, 2014, **18**(2), 271–295.

Kusuoka, S. On law invariant coherent risk measures. *Adv. in Math. Econ.*, 2001, **3**, 83–95.

McNeil, A. J. and Frey, R. Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *J. Empir. Financ.*, 2000, **7**(3), 271–300.

Rockafeller, R. T. *Convex analysis*, 1970 (Princeton University Press, Princeton, NJ).

Sandström, A. *Handbook of solvency for actuaries and risk managers: theory and practice*, 2010 (Taylor & Francis US, Florida).

Schmeidler, D. Subject probability and expected utility without additivity, *Econometrica*, 1989, **57**(3), 571–587.

SCOR. *From principle-based risk management to solvency requirements*. Second ed., SCOR, Zurich. SCOR Switzerland AG, 2008.

Tsanakas, A. To split or not to split: Capital allocation with convex risk measures. *Insur. Math. Econ.*, 2009, **44**(2), 268–277.

Wang, R., Bignozzi, V. and Tsakanas, A. How superadditive can a risk measure be? Forthcoming in *SIAM J. Finan. Math,*, 2015. Available at SSRN http://ssrn.com/abstract=2373149 (accessed 23 February 2015).

Wang, S. S., Young, V. R. and Panjer, H. H. Axiomatic characterization of insurance prices. *Insur. Math. Econ.*, 1997, **21**(2), 173–183.

Wüthrich, M. V., Bühlmann, H. and Furrer, H. *Market-consistent actuarial valuation*, 2010 (Springer).

Yaari, M. E. The dual theory of choice under risk. *Econometrica*, 1987, **55**(1), 95–115.

Ziegel, J. Coherence and elicitability. Forthcoming in *Math. Finance*, 2015. DOI: 10.1111/mafi.12080