

RESEARCH ARTICLE

Distributionally robust reinsurance with expectile*

Xinqiao Xie¹, Haiyan Liu^{2,*}, Tiantian Mao³ and Xiao Bai Zhu⁴

¹Department of Finance and Statistics, School of Management, University of Science and Technology of China, Hefei, China,

²Department of Mathematics and Department of Statistics and Probability, Michigan State University, East Lansing, USA,

³Department of Finance and Statistics, School of Management, University of Science and Technology of China, Hefei, China
and ⁴Department of Finance, Chinese University of Hong Kong, China

*Corresponding author. E-mail: hliu@math.msu.edu

Received: 31 January 2022; **Revised:** 25 October 2022; **Accepted:** 6 December 2022

Keywords: Expectile; distributional robustness; uncertainty; three-point distribution; stop-loss

Abstract

We study a distributionally robust reinsurance problem with the risk measure being an expectile and under expected value premium principle. The mean and variance of the ground-up loss are known, but the loss distribution is otherwise unspecified. A minimax problem is formulated with its inner problem being a maximization problem over all distributions with known mean and variance. We show that the inner problem is equivalent to maximizing the problem over three-point distributions, reducing the infinite-dimensional optimization problem to a finite-dimensional optimization problem. The finite-dimensional optimization problem can be solved numerically. Numerical examples are given to study the impacts of the parameters involved.

1. Introduction

The optimal reinsurance problem has been a popular topic since the seminal work of Borch (1960) and Arrow (1963). Especially after the introduction of coherent risk measure in Artzner *et al.* (1999) and convex risk measure in Frittelli and Rossaza Gianin (2002) and Föllmer and Schied (2002), the classical optimal reinsurance problem based on a risk measure has been widely studied under various choices of risk measures and different constraints on premiums; see for example Cai and Tan (2007), Chi and Tan (2011), Cui *et al.* (2013), Cheung *et al.* (2014) and Cai *et al.* (2016). See Cai and Chi (2020) for a review of optimal reinsurance designs based on risk measures.

In a classical reinsurance problem, the distribution of a loss is assumed to be precisely known. However, oftentimes only partial information is available for the loss distribution due to the lack of data and estimation error in practice. Recently, including model uncertainty in evaluation of a risk and reinsurance design has drawn increasing attention. Generally, model uncertainty is described by an uncertainty set. Two common ways are the moment-based uncertainty set and the distance-based uncertainty set. The former considers distributions satisfying certain constraints on moments while the latter considers distributions that are within a distance from a reference distribution. The introduction of uncertainty in evaluation of a risk motivates the study of the worst-case risk measure. For instance, El Ghaoui *et al.* (2003) studied the worst-case Value-at-Risk (VaR) and obtained a closed-form solution for the worst-case VaR over an uncertainty set that contains distributions with known mean and variance. Natarajan *et al.* (2010) showed that the worst-case Conditional Value-at-Risk (CVaR) for the same uncertainty set in El Ghaoui *et al.* (2003). In addition, Li (2018) extended those results to a general class of law invariant coherent risk measures. See Schied *et al.* (2009) for a review of robust preferences as a robust approach to the problem of model uncertainty. In reinsurance design, Hu *et al.* (2015) studied

*All authors contributed equally to this work.

optimal reinsurance with stop-loss contracts and incomplete information on the loss distribution in the sense that only the first two moments of the loss are known. See Pflug *et al.* (2017), Birghila and Pflug (2019) and Gavagan *et al.* (2022) for the design of an optimal insurance policy with the uncertainty set defined by the Wasserstein distance. Asimit *et al.* (2017) considered model uncertainty in insurance contract design by maximizing over a finite set of probability measures.

In this paper, we study a distributionally robust optimal reinsurance problem with a risk measure called expectile. Expectiles, introduced in Newey and Powell (1987) as the minimizers of an asymmetric quadratic loss function in the context of regression, are gaining increasing popularity in econometric literature (e.g. Kuan *et al.* 2009) and actuarial science (e.g. Bellini *et al.* 2014 and Cai and Weng 2016). Bellini *et al.* (2014) showed that expectile is a coherent risk measure under certain conditions, and it is robust in the sense of Lipschitzianity with respect to the Wasserstein metric. We assume that the distribution of a loss is partially known in the sense that the mean and variance of the loss are known. The distributionally robust optimal reinsurance problem we study is a minimax problem, where the inner problem is a maximization of the total retained loss over all distributions with known mean and variance and the outer problem is a minimization problem over all possible stop-loss reinsurance contracts. The main idea of solving the inner problem is to show that the inner problem is equivalent to optimization over all three-point distributions with known mean and variance, thus we reduce the infinite-dimensional optimization problem to a finite-dimensional optimization problem. At first glance, this conclusion seems similar to the one obtained in Liu and Mao (2022) for solving a distributionally robust reinsurance problem with VaR and CVaR. However, the proof of the main result is different from that in Liu and Mao (2022) because an expectile with levels different from 1/2 does not admit an explicit formula based on the distribution function as for VaR or CVaR. In addition, in contrast to Liu and Mao (2022), we did not obtain a closed-form solution to the reinsurance problem based on expectile, but came up with a finite-dimensional optimization problem. The main contribution of this paper is that we show that the worst-case distribution is among three-point distributions, which reduces the infinite-dimensional optimization problem to a finite-dimensional optimization problem. We emphasize that our main results appear nontrivial as the classical minimax theorem or duality cannot apply directly to the problem and a new technique is needed to obtain the main result.

The rest of the paper is organized as follows. In Section 2, the definition and properties of an expectile are given, and we present our distributionally robust reinsurance problem as a minimax problem. Section 3 aims to tackle the inner problem of the minimax problem. Proofs of the main results are given in Section 4. Numerical examples are given in Section 5 to study the impacts of the parameters on the optimal solution. Concluding remarks are given in Section 6.

2. Expectile and problem formulation

2.1. Expectile

Expectile, first introduced in Newey and Powell (1987) as the minimizer of an asymmetric quadratic loss function in the context of regression, is defined as follows.

Definition 1. *The α -expectile of a loss random variable X with $\mathbb{E}[X^2] < \infty$ at a confidence level $\alpha \in (0, 1)$, denoted by $e_\alpha(X)$, is defined as the unique minimizer of the following problem:*

$$e_\alpha(X) = \arg \min_{x \in \mathbb{R}} \left\{ \alpha \mathbb{E}[(X - x)_+^2] + (1 - \alpha) \mathbb{E}[(x - X)_+^2] \right\}, \quad (2.1)$$

where $(x)_+ := \max\{x, 0\}$.

Being the minimizer of a weighted mean squared error, an expectile has the property of elicibility, which is desirable as a risk measure has to be estimated from historical data and an elicitable risk measure makes it possible to verify and compare competing estimation procedures (e.g. Gneiting 2011; Kratz *et al.* 2018 and Bettels *et al.* 2022). Built on Weber (2006), Bellini and Bignozzi (2015) provided a full characterization of all elicitable monetary risk measures. See Bellini *et al.* (2014), Ziegel (2016),

Embrechts *et al.* (2021) and references therein for more discussions on the elicibility of risk measures and related properties. The following proposition is a list of properties of expectiles given in Bellini *et al.* (2014) and Cai and Weng (2016).

Proposition 1. *Let X be a loss random variable with $\mathbb{E}[X^2] < \infty$ and $e_\alpha(X)$ be the α -expectile of X , $\alpha \in (0, 1)$. Then*

(i) *A number $e_\alpha(X) \in \mathbb{R}$ solves optimization problem (2.1) if and only if*

$$\alpha \mathbb{E}[(X - e_\alpha(X))_+] = (1 - \alpha) \mathbb{E}[(e_\alpha(X) - X)_+]. \tag{2.2}$$

(ii) *The expectile $e_\alpha(X)$ is a coherent risk measure if $\alpha \geq 1/2$.*

(iii) *$e_\alpha(X) \leq \text{ess-sup } X$.*

(iv) *$e_\alpha(X) = \mathbb{E}[X] + \beta \mathbb{E}[(X - e_\alpha(X))_+]$ with $\beta = \frac{2\alpha - 1}{1 - \alpha}$.*

Proposition 1(iv) implies that $e_\alpha(X) \leq \mathbb{E}[X]$ for $\alpha \leq 1/2$ and $e_\alpha(X) \geq \mathbb{E}[X]$ for $\alpha \geq 1/2$. For the purpose of risk management in insurance and finance, a risk measure of a loss random variable, as a tool of calculating premium or regulatory capital requirement, is normally required to be larger than the expected loss. In addition, the expectile is a coherent risk measure for $\alpha \geq 1/2$, possessing the sub-additivity property, which is a natural requirement to meet that “a merger does not create extra risk”. Therefore, throughout this paper, we are interested in the case of $\alpha > 1/2$, and we will also show later that the reinsurance problem is trivial for $\alpha \leq 1/2$ (see Proposition 2).

2.2. Distributionally robust reinsurance with expectile

Let X be a non-negative ground-up loss faced by an insurer. The insurer transfers part of the loss, say $I(X)$, to a reinsurer at the cost of paying reinsurance premium. The reinsurance premium is considered as a function of the reinsurance contract $I(X)$, denoted by $\pi(I(X))$. In a reinsurance contract, the function $I(\cdot)$ is called a ceded loss function. After purchasing the reinsurance contract $I(X)$, the total retained risk exposure of the insurer is $X - I(X) + \pi(I(X))$. In this paper, we determine the optimal ceded loss function or reinsurance contract from the insurer’s perspective instead of the reinsurer’s standpoint.

In a classical reinsurance problem, the distribution of the ground-up loss X is assumed precisely known. The aim of a classical reinsurance problem is to find an optimal reinsurance contract so that the risk measurement of the total retained risk exposure of the insurer is minimized, that is

$$\text{minimize } \rho(X - I(X) + \pi(I(X))) \quad \text{over } I \in \mathcal{I}, \tag{2.3}$$

where ρ is a risk measure and \mathcal{I} is a set of candidate reinsurance contracts. See Cai and Chi (2020) for a review of classical optimal reinsurance designs with risk measures.

In this paper, we consider a distributionally robust optimal reinsurance problem in which the cumulative distribution function (cdf) of the ground-up loss X is not completely known. Throughout the paper, we assume that the distribution of the ground-up loss is partially known in the sense that only the mean and variance of X are known. Given a pair of non-negative mean and standard deviation (μ, σ) of X , define the uncertainty set:

$$S(\mu, \sigma) = \left\{ F \text{ is a cdf on } [0, \infty): \int_0^\infty x dF(x) = \mu, \int_0^\infty x^2 dF(x) = \mu^2 + \sigma^2 \right\}.$$

Let \mathcal{I} be the class of stop-loss reinsurance contracts. A stop-loss reinsurance contract $I(X)$ is defined as $I(X) = (X - d)_+$, $d \in [0, \infty]$, where d is called a *deductible*. By convention, $(X - \infty)_+ = 0$. Borch (1960) showed that a stop-loss reinsurance is optimal when the insurer minimizes the variance of its total retained risk exposure with the premium computed under the expected value premium principle. Besides, Arrow (1963) showed that a stop-loss reinsurance is also optimal if the insurer maximizes his/her expected utility of its terminal wealth under the expected value premium principle. A similar conclusion was also obtained in Cheung *et al.* (2014) under law-invariant convex risk measures.

Furthermore, stop-loss reinsurance is popular in practice. Thus, we consider stop-loss reinsurance contracts as our candidate reinsurance contracts. A common premium principle is the expected value premium principle, which is defined as $\pi(I(\cdot)) = (1 + \theta)\mathbb{E}[I(\cdot)]$ for a reinsurance contract $I \in \mathcal{I}$, where $\theta > 0$ is called a safety loading factor. We are interested in the following distributionally robust reinsurance problem with the risk measure expectile and under expected value premium principle:

$$\min_{I \in \mathcal{I}} \sup_{F \in \mathcal{S}(\mu, \sigma)} e_\alpha^F(X - I(X) + \pi(I(X))), \tag{2.4}$$

where $\alpha > 1/2$ and the superscript F indicates that the expectile and the premium are calculated with X following the distribution F . With $I(X) = (X - d)_+$, the total retained risk exposure of the insurer is $X - I(X) + \pi(I(X)) = X \wedge d + (1 + \theta)\mathbb{E}[(X - d)_+]$, where $x \wedge y := \min\{x, y\}$. Furthermore, by the translation invariance of a coherent risk measure, problem (2.4) can be reduced to

$$\min_{d \geq 0} \sup_{F \in \mathcal{S}(\mu, \sigma)} \{e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]\}. \tag{2.5}$$

A distribution $F \in \mathcal{S}(\mu, \sigma)$ that solves the inner problem of (2.5) is called the *worst-case distribution*. Notably, if $\alpha \leq 1/2$, we can show that the objective function $e_\alpha^F[X \wedge d] + (1 + \theta)\mathbb{E}^F[(X - d)_+]$ is always decreasing in d , and thus, the optimal deductible of problem (2.5) is $d^* = \infty$.

Proposition 2. *For $\alpha \leq 1/2$, we have the optimal deductible of problem (2.5) is $d^* = \infty$.*

Proof. Denote by $g^F(d) := e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]$, $d \in \mathbb{R}$, and it suffices to show that $g^F(d)$ is decreasing in $d \geq 0$, which obviously implies $\sup_{F \in \mathcal{S}(\mu, \sigma)} g^F(d)$ is decreasing in $d \geq 0$. By definition of $e_\alpha^F[X \wedge d] =: x_d$ in (2.2), we have x_d satisfies

$$\alpha \int_{x_d}^d \bar{F}(y)dy = (1 - \alpha) \int_0^{x_d} F(y)dy.$$

where $\bar{F}(y) = 1 - F(y)$. Taking (left-)derivative with respect to d yields

$$\frac{\partial x_d}{\partial d} = \frac{\alpha \bar{F}(d)}{\alpha + (1 - 2\alpha)F(x_d)}.$$

Noting that $\partial \mathbb{E}[(X - d)_+] / \partial d = -\bar{F}(d)$, we have

$$\frac{\partial g^F(d)}{\partial d} = \frac{-(1 - 2\alpha)F(x_d)\bar{F}(d)}{\alpha + (1 - 2\alpha)F(x_d)} - \theta \bar{F}(d) \leq 0,$$

where the inequality follows from that $\alpha \leq 1/2$. Thus, we have $g^F(d)$ is decreasing in $d \geq 0$ which completes the proof. \square

The distributional robust reinsurance problem is trivial for $\alpha \leq 1/2$ and the optimal deductible is $d^* = \infty$ by Proposition 2. Therefore, in the rest of this paper, we only need to consider the case $\alpha > 1/2$. In the next section, we will first solve the inner problem of (2.5) for $\alpha > 1/2$, that is, we work on the worst-case distribution of the inner problem of (2.5).

Remark 1. It is tempting to use the minimax theorem to tackle problem (2.5) since both $e_\alpha^F(X)$ and $\mathbb{E}^F[(X - d)_+]$ are quasi-linear in F (which does not imply that the objective function as a whole is quasi-linear in F). However, the quasi-convexity or quasi-concavity of the objective function with respect to (X, d) or (F, d) for problem (2.5) cannot be established since e_α is convex in d but the functional $X \wedge d$ is concave in d . Therefore, the minimax theorem and duality of the optimization problem cannot apply directly to problem (2.5).

3. Main results

3.1. The worst-case distribution

In this section, we focus on tackling the inner problem of (2.5) for $\alpha > 1/2$, that is,

$$\sup_{F \in S(\mu, \sigma)} \{e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]\} \tag{3.1}$$

We aim to show that the worst-case distribution of the optimization problem (3.1) must be a three-point distribution if it exists, that is, it belongs to the following uncertainty set

$$S_3(\mu, \sigma) := \{F \in S(\mu, \sigma) : F \text{ is a three-point cdf}\}. \tag{3.2}$$

Here, we make the convention that two-point distribution and point mass distributions are special cases of three-point distributions. The following theorem states that the worst-case distribution to problem (3.1) is among three-point distributions.

Theorem 1. *For $d \geq 0$ and $\alpha > 1/2$, the problem (3.1) is equivalent to*

$$\sup_{F \in S_3(\mu, \sigma)} \{e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]\} \tag{3.3}$$

in the sense that the two problems have the same optimal value. Moreover, the worst-case distribution of the problem (3.1) exists if and only if the worst-case distribution of the problem (3.3) exists, and any worst-case distribution of the problem (3.3) must be that of the problem (3.1).

Theorem 1 states that we can work on the set of three-point distributions without loss of generality. We next give an example to illustrate that in general the worst-case distribution of the problem (3.1) is not unique and we may find distributions outside the set $S_3(\mu, \sigma)$ to attain the supremum.

Remark 2. Generally speaking, the worst-case distribution of the problem (3.1) is not unique. For example, letting $d = 0$, the problem (3.1) reduces to

$$\sup_{F \in S(\mu, \sigma)} (1 + \theta)\mathbb{E}^F[X].$$

In this special case, the optimal value is $(1 + \theta)\mu$ and the worst-case distribution is any feasible distribution. We also point out that the case $d = 0$ means full reinsurance, which is a common reinsurance treaty in practice; see numerical results in Section 5.

From Theorem 1, we know that the worst-case distribution of problem (3.1) is among three-point distributions with a more specific form. Denote by

$$[x_1, p_1; x_2, p_2; x_3, p_3]$$

a three-point distribution of a random variable X with $\mathbb{P}(X = x_i) = p_i \geq 0$, $i = 1, 2, 3$, where $0 \leq x_1 \leq x_2 \leq x_3$, $p_1 + p_2 + p_3 = 1$, and $p_i \in [0, 1]$, $i = 1, 2, 3$. More specifically, we can get the following result from the proof of Theorem 1.

Corollary 1. *For $d \geq 0$, the problem (3.3) and thus, the problem (3.1) is equivalent to*

$$\sup_{F \in S_3^*(\mu, \sigma)} \{e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]\}, \tag{3.4}$$

where

$$S_3^*(\mu, \sigma; d) = \{F = [x_1, p_1; x_2, p_2; x_3, p_3] \in S_3(\mu, \sigma) : x_1 \leq e_\alpha^F(X \wedge d) \leq x_2 \leq d \leq x_3\}.$$

3.2. Transformations of the main problem

In this subsection, we aim to transform problem (2.5) as a finite-dimensional tractable problem based on Theorem 1 and Corollary 1. We first make the following observations. For any $F \in S_3^*(\mu, \sigma; d)$, by

Proposition 1(i), one can verify that

$$e_\alpha^F(X \wedge d) = \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} \tag{3.5}$$

and hence,

$$\begin{aligned} f^F(d, X) &:= e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+] \\ &= \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3. \end{aligned} \tag{3.6}$$

Combining with Theorem 1, we conclude that the infinite-dimensional optimization problem (3.1) can be reduced to a finite-dimensional optimization problem:

$$\sup \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3 \tag{3.7}$$

$$\text{s.t.} \begin{cases} p_1 + p_2 + p_3 = 1, p_i \geq 0, i = 1, 2, 3, \\ p_1x_1 + p_2x_2 + p_3x_3 = \mu, \\ p_1x_1^2 + p_2x_2^2 + p_3x_3^2 = \mu^2 + \sigma^2, \\ \frac{d - x_2}{x_2 - x_1} \leq \frac{(1 - \alpha)p_1}{\alpha p_3}, \\ 0 \leq x_1 \leq x_2 \leq d \leq x_3. \end{cases} \tag{3.8}$$

The fourth constraint in (3.8) guarantees that $e_\alpha^F(X \wedge d) \leq x_2$. For any three-point distribution $G = [x_1, p_1; x_2, p_2; x_3, p_3] \in \mathcal{S}_3(\mu, \sigma)$ satisfying $x_1 \leq x_2 \leq e_\alpha^G(X \wedge d) \leq d \leq x_3$, by Proposition 1(i), we obtain

$$e_\alpha^G(X \wedge d) = \frac{(1 - \alpha)p_1x_1 + (1 - \alpha)p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + (1 - \alpha)p_2 + \alpha p_3}.$$

The condition $x_2 \leq e_\alpha(X_G \wedge d)$ is equivalent to $(1 - \alpha)p_1(x_1 - x_2) + \alpha p_3(d - x_2) \geq 0$, which implies

$$\begin{aligned} f^G(d, X) &= \frac{(1 - \alpha)p_1x_1 + (1 - \alpha)p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + (1 - \alpha)p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3 \\ &\geq \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3. \end{aligned}$$

Together with Corollary 1, dropping the fourth constraint in (3.8) still leads to the same maximum of (3.7) subject to all constraints in (3.8). Hence, we have the following theorem.

Theorem 2. *The optimization problem (3.1) is equivalent to*

$$\sup \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3 \tag{3.9}$$

$$\text{s.t.} \begin{cases} p_1 + p_2 + p_3 = 1, p_i \geq 0, i = 1, 2, 3, \\ p_1x_1 + p_2x_2 + p_3x_3 = \mu, \\ p_1x_1^2 + p_2x_2^2 + p_3x_3^2 = \mu^2 + \sigma^2, \\ 0 \leq x_1 \leq x_2 \leq d \leq x_3 \end{cases} \tag{3.10}$$

in the sense that the two problems have the same optimal value. Moreover, the worst-case distribution of the problem (3.1) exists if and only if the optimal solution of the problem (3.9) exists, and the worst-case distribution of the problem (3.1) is $F^ = [x_1^*, p_1^*; x_2^*, p_2^*; x_3^*, p_3^*]$ if $(x_i^*, p_i^*, i = 1, 2, 3)$ is the optimal solution of the problem (3.9).*

With Theorem 2, we reduce the infinite-dimensional optimization problem (3.1), or the inner problem of (2.4), to a finite-dimensional optimization problem, which can be solved numerically. In Section 5, we will solve problem (2.4) numerically, where the inner problem is solved by the Matlab build-in function “fmincon” and the outer problem is solved via a grid search.

Lemma 1 and Theorem 1 imply that the worst-case value of problem (3.1) is increasing in σ ; that is, the optimal value of the problem (3.1) is equivalent to the optimal value of the following problem

$$\sup_{F \in \mathcal{U}_{X \leq \sigma} S(\mu, \sigma)} \{e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]\}.$$

Therefore, we have the following reformulation of the problem (3.1).

Proposition 3. *The problem (3.1) is equivalent to the following problem*

$$\sup \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3 \tag{3.11}$$

$$\text{s.t.} \begin{cases} p_1 + p_2 + p_3 = 1, p_i \geq 0, i = 1, 2, 3, \\ p_1x_1 + p_2x_2 + p_3x_3 = \mu, \\ p_1x_1^2 + p_2x_2^2 + p_3x_3^2 \leq \mu^2 + \sigma^2, \\ 0 \leq x_1 \leq x_2 \leq d \leq x_3. \end{cases} \tag{3.12}$$

in the sense that they have the same optimal value.

It is worth noting that the result in Proposition 3 is not easy to check due to the non-negativity assumption. In what follows, we explain why the non-negativity of the loss risk plays an essential role in the problem and makes the problem more complicated. If we drop the assumption of non-negativity of the loss, by the translation invariance and positive homogeneity of e_α , it follows that for any risk X with mean 0 and variance 1, $Y := \mu + \sigma X$ has mean μ and variance σ^2 , and

$$e_\alpha(Y \wedge d_1) + (1 + \theta)\mathbb{E}[(Y - d_1)_+] = \mu + \sigma (e_\alpha(X \wedge d) + (1 + \theta)\mathbb{E}[(X - d)_+]),$$

where $d_1 = \mu + \sigma d$. Therefore, it suffices to consider the special case of the uncertainty set with $\mu = 0$ and $\sigma = 1$. If the optimal deductible in the case of $\mu = 0$ and $\sigma = 1$ is d^* , then the optimal deductible for the general case (μ, σ) is $\mu + \sigma d^*$. However, with the constraint of non-negativity, the above observations do not hold anymore.

The following proposition discusses the attainability of problems (3.11) and (3.9).

Proposition 4.

- (i) *The supremum value of the problem (3.11) is always attainable.*
- (ii) *The supremum value of the problem (3.9) is attainable if one of the following conditions*

$$(1 + \theta)[(1 - \alpha)(1 - \mu/d) + \alpha\mu/d]^2 \leq \alpha(1 - \alpha), \mu < d, \text{ and } \mu d < \mu^2 + \sigma^2 \tag{3.13}$$

is violated.

From Proposition 4, we have that for any $d \geq 0$, there exists an F^* with $\mathbb{E}^{F^*}[X] = \mu$ and $\text{VaR}^{F^*}(X) \leq \sigma^2$ such that

$$\sup_{F \in S(\mu, \sigma)} f^F(d, X) = f^{F^*}(d, X),$$

where $f^{F^*}(d, X) := e_\alpha^{F^*}(X \wedge d) + (1 + \theta)\mathbb{E}^{F^*}[(X - d)_+]$. We also point out that if $F^* \neq [0, 1 - \mu/d; d, \mu/d]$, then the problem (3.9) is attainable.

We close this section by showing that the main results in the paper can be generalized to the case with a higher order moment condition. That is, if we replace the constraint on variance by a constraint of a higher order moment, then the results that are parallel to Theorems 1 and 2 still hold. To be more

specific, if the uncertainty set is replaced by

$$S_k(\mu, m) = \left\{ F \text{ is a cdf on } [0, \infty) : \int_0^\infty x dF(x) = \mu, \int_0^\infty x^k dF(x) = m \right\},$$

where $k > 1$, then similarly, we can show that Theorem 1 still holds. The following problem

$$\sup_{F \in S_k(\mu, m)} e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+], \tag{3.14}$$

is equivalent to

$$\sup \frac{(1 - \alpha)p_1x_1 + \alpha p_2x_2 + \alpha p_3d}{(1 - \alpha)p_1 + \alpha p_2 + \alpha p_3} + (1 + \theta)(x_3 - d)p_3 \tag{3.15}$$

$$\text{s.t. } \begin{cases} p_1 + p_2 + p_3 = 1, p_i \geq 0, i = 1, 2, 3, \\ p_1x_1 + p_2x_2 + p_3x_3 = \mu, \\ p_1x_1^k + p_2x_2^k + p_3x_3^k = m, \\ 0 \leq x_1 \leq x_2 \leq d \leq x_3, \end{cases} \tag{3.16}$$

in the sense that the two problems have the same optimal value and optimal solution. More specifically, the worst-case distribution of the problem (3.14) exists if and only if the optimal solution of the problem (3.15) exists, and the worst-case distribution of the problem (3.14) is $F^* = [x_1^*, p_1^*; x_2^*, p_2^*; x_3^*, p_3^*]$ if $(x_i^*, p_i^*, i = 1, 2, 3)$ is the optimal solution of the problem (3.15).

4. Proofs of the main results in Section 3

To prove Theorem 1, we need the following lemma.

Lemma 1. For $d \geq 0, \sigma_1 < \sigma_2$ and $\alpha \geq 1/2$, let $F \in S_3(\mu, \sigma_1)$ be a distribution such that $\mathbb{P}^F(X = x_i) = p_i, i = 1, 2, 3$, with $x_1 < x_2 < x_3$ and

$$x_1 \leq e_\alpha^F(X \wedge d) \leq x_2 \leq d < x_3. \tag{4.1}$$

Then there exists $F^* \in S_3(\mu, \sigma_2)$ such that (4.1) holds,

$$e_\alpha^{F^*}(X \wedge d) = e_\alpha^F(X \wedge d) \text{ and } \mathbb{E}^{F^*}[(X - d)_+] \geq \mathbb{E}^F[(X - d)_+].$$

Proof. Our proof mainly involves two steps: We first define a three-point random variable Y_c such that $\mathbb{E}[Y_c] = \mathbb{E}^F[X], \mathbb{E}[(Y_c - d)_+] \geq \mathbb{E}^F[(X - d)_+]$, and $e_\alpha(Y_c \wedge d) = e_\alpha^F(X \wedge d)$. Here in general it does not hold $\text{Var}(Y_c) = \sigma_2^2$, and thus, the next step is to modify the definition of Y_c to obtain a desired distribution. To show it, for $c \in [0, p_2]$, define a random variable

$$Y_c = \begin{cases} x_1, & \text{with probability } p_1 + a, \\ x_2, & \text{with probability } p_2 - c, \\ y, & \text{with probability } p_3 + b, \end{cases} \tag{4.2}$$

where

$$a := \frac{\alpha(d - x_2)c}{m} \geq 0, b := c - a \geq 0, y := \frac{\mu - (p_1 + a)x_1 - (p_2 - c)x_2}{p_3 + b},$$

and

$$m := (1 - \alpha)[e_\alpha^F(X \wedge d) - x_1] + \alpha[d - e_\alpha^F(X \wedge d)].$$

One can easily verify that $\mathbb{E}[Y_c] = \mu$. We claim that $y > d$. Indeed, as

$$y > d \Leftrightarrow p_3x_3 - ax_1 + cx_2 > d(p_3 + b) \Leftrightarrow p_3x_3 - ax_1 + cx_2 > d(p_3 + c - a) \\ \Leftrightarrow p_3(x_3 - d) > a(x_1 - d) + c(d - x_2),$$

it remains to show $p_3(x_3 - d) > a(x_1 - d) + c(d - x_2)$. Note that

$$a(x_1 - d) + c(d - x_2) = \frac{\alpha(d - x_2)c}{m}(x_1 - d) + c(d - x_2) \\ = c(d - x_2) \left[\frac{\alpha(x_1 - d)}{m} + 1 \right] \\ = c(d - x_2) \frac{\alpha(x_1 - d) + (1 - \alpha)[e_\alpha^F(X \wedge d) - x_1] + \alpha[d - e_\alpha^F(X \wedge d)]}{m} \\ = c(d - x_2) \frac{(1 - 2\alpha)[e_\alpha^F(X \wedge d) - x_1]}{m} \leq 0, \tag{4.3}$$

where the last inequality follows from $x_1 \leq e_\alpha^F(X \wedge d) \leq x_2 \leq d$ and $\alpha \geq 1/2$. As $x_3 > d$, we have $p_3(x_3 - d) > 0 \geq a(x_1 - d) + c(d - x_2)$. Hence,

$$y > d. \tag{4.4}$$

Moreover, by standard computation, we have

$$\mathbb{E}[(Y_c - d)_+] = (y - d)(p_3 + b) \\ = \mu - (p_1 + a)x_1 - (p_2 - c)x_2 - d(p_3 + b) \\ = p_3(x_3 - d) + a(d - x_1) + c(x_2 - d) \\ \geq p_3(x_3 - d) = \mathbb{E}^F[(X - d)_+], \tag{4.5}$$

where the inequality follows from (4.3). Next, we show that $e_\alpha(X_F \wedge d) = e_\alpha(Y_c \wedge d)$. By Proposition 1(i), we know

$$\alpha \mathbb{E}^F [(X \wedge d - e_\alpha^F(X \wedge d))_+] = (1 - \alpha) \mathbb{E}^F [(e_\alpha^F(X \wedge d) - X \wedge d)_+],$$

which is equivalent to

$$\alpha [(x_2 - e_\alpha^F(X \wedge d))p_2 + (d - e_\alpha^F(X \wedge d))p_3] = (1 - \alpha)(e_\alpha^F(X \wedge d) - x_1)p_1. \tag{4.6}$$

It then follows from standard computation that

$$\alpha \mathbb{E} [(Y_c \wedge d - e_\alpha^F(X \wedge d))_+] \\ = \alpha [(x_2 - e_\alpha^F(X \wedge d))(p_2 - c) + (d - e_\alpha^F(X \wedge d))(p_3 + b)] \\ = (1 - \alpha)(e_\alpha^F(X \wedge d) - x_1)p_1 - \alpha(x_2 - e_\alpha^F(X \wedge d))c + \alpha(d - e_\alpha^F(X \wedge d))b \\ = (1 - \alpha)(e_\alpha^F(X \wedge d) - x_1)p_1 + \alpha c(d - x_2) \frac{(1 - \alpha)(e_\alpha^F(X \wedge d) - x_1)}{m} \\ = (1 - \alpha)(e_\alpha^F(X \wedge d) - x_1)p_1 + a(1 - \alpha)(e_\alpha^F(X \wedge d) - x_1) \\ = (1 - \alpha) \mathbb{E} [(e_\alpha^F(X \wedge d) - Y_c \wedge d)_+],$$

where the first equality follows from $e_\alpha^F(X \wedge d) \leq x_2 \leq d$ and the second equality follows from (4.6). By Proposition 1(i), we know that $e_\alpha^F(X \wedge d)$ is the α -expectile of $Y_c \wedge d$, that is,

$$e_\alpha^F(X \wedge d) = e_\alpha(Y_c \wedge d). \tag{4.7}$$

Now we consider the following two cases:

- (i) If there exists a $c^* \in [0, p_2]$ such that $\text{Var}(Y_{c^*}) = \sigma_2^2$, combining with (4.5) and (4.7), we know that the distribution of Y_{c^*} is the desired three-point distribution.
- (ii) Otherwise, $\text{Var}(Y_c) < \sigma_2^2$ for all $c \in [0, p_2]$ and we need to define a new random variable. Note that for $c = p_2$, the random variable Y_c (defined in (4.2)) reduces to

$$Y = \begin{cases} x_1, & \text{with probability } q, \\ \frac{\mu - qx_1}{1 - q}, & \text{with probability } 1 - q, \end{cases}$$

where $q = p_1 + \frac{\alpha(d-x_2)p_2}{m}$. By (4.4), we know $\frac{\mu - qx_1}{1 - q} > d$. For $h \in [0, 1 - q]$, define

$$Z_h = \begin{cases} x_1, & \text{with probability } q, \\ d, & \text{with probability } h, \\ z, & \text{with probability } 1 - q - h, \end{cases} \tag{4.8}$$

where $z := \frac{\mu - x_1q - dh}{1 - q - h}$. It is straightforward to verify that $z > d$, $\mathbb{E}(Z_h) = \mu$, $\text{Var}(Z_0) = \text{Var}(Y)$ and $e_\alpha(Z_h \wedge d) = e_\alpha(Y \wedge d) = e_\alpha(Y_c \wedge d) = e_\alpha^F(X \wedge d)$. Moreover, note that

$$\begin{aligned} \mathbb{E}[(Z_h - d)_+] &= (z - d)(1 - q - h) \\ &= (1 - q) \left[\frac{\mu - x_1q}{1 - q} - d \right] \\ &= \mathbb{E}[(Y - d)_+] = \mathbb{E}[(Y_c - d)_+] \geq \mathbb{E}^F[(X - d)_+], \end{aligned}$$

where the last inequality follows from (4.5). As $\text{Var}(Y_c) < \sigma_2^2$ for all $c \in [0, p_2]$, by the continuity of $\text{Var}(Y_c)$ with respect to c , we have $\text{Var}(Y_{p_2}) = \text{Var}(Y) < \sigma_2^2$. Noting that $\lim_{h \rightarrow 1 - q} \text{Var}(Z_h) = \infty$ and $\text{Var}(Z_0) = \text{Var}(Y) < \sigma_2^2$, there exists an $h^* \in [0, 1 - q]$ such that $\text{Var}(Z_{h^*}) = \sigma_2^2$. In this case, the distribution of Z_{h^*} is the desired distribution F^* .

Combining the above two cases, we complete the proof. □

To better understand the main steps in Lemma 1, we give the following example to illustrate the two-step procedure involved in the proof.

Example 1. Suppose $d = 6$, $\alpha = 0.9$. Consider $S_3(\mu, \sigma_1)$ with $\mu = \frac{10}{3}$, $\sigma_1 = \frac{\sqrt{35}}{3}$. Let $\sigma_2 > \sigma_1$ and F be the distribution function of a discrete random variable X , where

$$X = \begin{cases} 2, & \text{with probability } \frac{2}{3}, \\ 5, & \text{with probability } \frac{1}{6}, \\ 7, & \text{with probability } \frac{1}{6}. \end{cases}$$

Obviously, $F \in S_3(\mu, \sigma_1)$. Moreover, $2 < e_\alpha^F(X \wedge d) = \frac{107}{22} < 5 < 6 < 7$. Hence, X satisfies the conditions of Lemma 1. For $c \in [0, \frac{1}{6}]$, define

$$Y_c = \begin{cases} 2, & \text{with probability } \frac{2}{3} + a, \\ 5, & \text{with probability } \frac{1}{6} - c, \\ y, & \text{with probability } \frac{1}{6} + b, \end{cases} \tag{4.9}$$

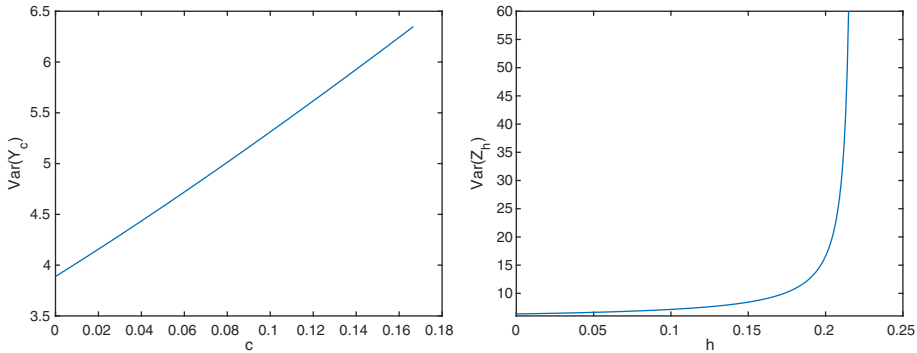


Figure 1. Trend of random variable's variance with respect to c and h .

where

$$a = \frac{99c}{144} \geq 0, \quad b = \frac{45c}{144} \geq 0 \quad \text{and} \quad y = \frac{168 + 522c}{24 + 45c}.$$

It is easy to verify that $\mathbb{E}[Y_c] = \mu$, $y > d$ and $e_\alpha(Y_c \wedge d) = \frac{107}{22}$.

- (i) If $\sigma_2^2 \leq \frac{400}{63} = 6.3492$, we can always find a $c^* \in [0, 1/6]$ such that $\text{Var}(Y_{c^*}) = \sigma_2^2$; see the left graph in Figure 1. One can verify that

$$\mathbb{E}[(Y_c - d)_+] \geq \mathbb{E}^F[(X - d)_+], \quad e_\alpha^F(X \wedge d) = e_\alpha(Y_c \wedge d).$$

- (ii) If $\sigma_2^2 > 6.3492$, for all $c \in [0, 1/6]$, we have $\text{Var}(Y_c) < \sigma_2^2$. For $c = \frac{1}{6}$, the random variable Y_c defined by (4.9) reduces to

$$Y = \begin{cases} 2, & \text{with probability } \frac{25}{32}, \\ \frac{170}{21}, & \text{with probability } \frac{7}{32}. \end{cases}$$

For $h \in [0, 7/32]$, define

$$Z_h = \begin{cases} 2, & \text{with probability } q, \\ 6, & \text{with probability } h, \\ z, & \text{with probability } 1 - q - h, \end{cases}$$

where $q = \frac{25}{32}$, $z = \frac{\frac{10}{3} - 2q - 6h}{1 - q - h}$. The right graph in Figure 1 plots the relationship between $\text{Var}(Z_h)$ and h . One can verify that $\mathbb{E}[Z_h] = 10/3 = \mu$,

$$\mathbb{E}[(Z_h - d)_+] \geq \mathbb{E}^F[(X - d)_+] \quad \text{and} \quad e_\alpha^F(X \wedge d) = e_\alpha(Z_h \wedge d).$$

Now we are ready to prove Theorem 1.

Proof of Theorem 1. Let $f^F(d, X) = e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]$ for each $F \in S(\mu, \sigma)$. Note that $e_\alpha^F(X \wedge d) \leq d$ as $X \wedge d \leq d$. Let $A_1 = \{X \leq e_\alpha^F(X \wedge d)\}$, $A_2 = \{e_\alpha^F(X \wedge d) < X \leq d\}$, and $A_3 = \{X > d\}$.

Denote $\mathbb{E}^F[X|A_i]$ by $x_i, i = 1, 2, 3$. Define a discrete random variable

$$\tilde{X}(\omega) = \begin{cases} x_1, & \text{if } \omega \in A_1, \\ x_2, & \text{if } \omega \in A_2, \\ x_3, & \text{if } \omega \in A_3. \end{cases} \tag{4.10}$$

Denote the distribution of \tilde{X} by \tilde{F} . Obviously \tilde{F} is a three-point distribution satisfying (4.1), that is, $x_1 \leq e_\alpha^F(X \wedge d) \leq x_2 \leq d < x_3$. It follows that $\mathbb{E}^{\tilde{F}}[\tilde{X}] = \mu$ and

$$\mathbb{E}^{\tilde{F}}[(\tilde{X} - d)_+] = (x_3 - d)\mathbb{P}(A_3) = \mathbb{E}^F[(X - d)_+].$$

By Hölder’s inequality, we have $(\int_{A_i} x dF(x))^2 \leq (\int_{A_i} x^2 dF(x))\mathbb{P}(A_i), i = 1, 2, 3$. As a result,

$$\mathbb{E}^{\tilde{F}}[\tilde{X}^2] = \sum_{i=1}^3 \frac{(\int_{A_i} x dF(x))^2}{\mathbb{P}(A_i)} \leq \sum_{i=1}^3 \int_{A_i} x^2 dF(x) = \mathbb{E}^F[X^2].$$

Therefore, $\text{Var}^{\tilde{F}}(\tilde{X}) \leq \sigma^2 = \text{Var}^F(X)$. Note that

$$\begin{aligned} \alpha \mathbb{E}^{\tilde{F}}[(\tilde{X} \wedge d - e_\alpha^F(X \wedge d))_+] &= \alpha [(x_2 - e_\alpha^F(X \wedge d))\mathbb{P}(A_2) + (d - e_\alpha^F(X \wedge d))\mathbb{P}(A_3)] \\ &= \alpha \mathbb{E}^F[(X \wedge d - e_\alpha^F(X \wedge d))_+], \end{aligned}$$

and

$$(1 - \alpha)\mathbb{E}^{\tilde{F}}[(e_\alpha^F(X \wedge d) - \tilde{X} \wedge d)_+] = (1 - \alpha)\mathbb{E}^F[(e_\alpha^F(X \wedge d) - X \wedge d)_+].$$

By Proposition 1(i), we have $e_\alpha^{\tilde{F}}(\tilde{X} \wedge d) = e_\alpha^F(X \wedge d)$. Hence, for any $F \in S(\mu, \sigma)$, there exists a random variable \tilde{X} defined in (4.10) following a three-point distribution such that $f^{\tilde{F}}(d, \tilde{X}) = f^F(d, X)$. Next, we consider the following two cases.

- (i) If $\text{Var}^{\tilde{F}}(\tilde{X}) = \sigma^2$, then $\tilde{F} \in S_3(\mu, \sigma)$ and $f^{\tilde{F}}(d, \tilde{X}) = f^F(d, X)$. The result follows.
- (ii) If $\text{Var}^{\tilde{F}}(\tilde{X}) < \sigma^2$, by Lemma 1, there exists a three-point distribution $F^* \in S_3(\mu, \sigma)$ such that $\mathbb{E}^{F^*}[X] = \mathbb{E}^{\tilde{F}}[\tilde{X}]$, $\text{Var}^{F^*}(X) = \sigma^2$, $e_\alpha^{F^*}(X \wedge d) = e_\alpha^{\tilde{F}}(\tilde{X} \wedge d)$ and $\mathbb{E}^{F^*}[(X - d)_+] \geq \mathbb{E}^{\tilde{F}}[(\tilde{X} - d)_+]$. Then we have $f^{F^*}(d, X) \geq f^{\tilde{F}}(d, \tilde{X}) = f^F(d, X)$.

Hence, for any $F \in S(\mu, \sigma)$, we can find a three-point distribution $F^* \in S_3(\mu, \sigma)$ such that $f^{F^*}(d, X) \geq f^F(d, X)$. The proof is complete. □

Remark 3.¹ It is worth noting that if we extend the stop-loss contract to $I(x) = c(x - d)_+$, where $c \in [0, 1]$, with one more parameter c introduced, the difficulty of solving the problem increases significantly. With $c = 1$, we have $X - I(X) = X \wedge d$, but for a $c \in [0, 1)$, $X - I(X) = X - c(X - d)_+ = X \wedge d + (1 - c)(X - d)_+$, and the objective function is

$$e_\alpha^F(X \wedge d + (1 - c)(X - d)_+) + c(1 + \theta)\mathbb{E}^F[(X - d)_+]$$

as opposed to

$$e_\alpha^F(X \wedge d) + (1 + \theta)\mathbb{E}^F[(X - d)_+]$$

in the case of stop-loss contract. Generally speaking, we could not establish the same result in Theorem 1. However, we can show that for $I(x) = c(x - d)_+$, we can confine the worst-case distribution to the four-point distributions set $S_4(\mu, \sigma)$ with similar method in Theorem 1 and Lemma 1.

Proof of Proposition 4. (i) Denote by f^* the optimal value of the problem (3.11). There exist feasible distributions $F_n = [x_1^n, p_1^n; x_2^n, p_2^n; x_3^n, p_3^n] \in S_3(\mu, \sigma), n \in \mathbb{N}$, such that $f^{F_n}(d, X) \rightarrow f^*$, where $f^F(d, X)$ is defined by (3.6). The constraints in (3.10) imply that $p_i^n \in [0, 1], i = 1, 2, 3$ and $x_1^n, x_2^n \in [0, d]$ for $k \in \mathbb{N}$. We consider the following two cases.

¹This is suggested by one anonymous referee.

- (a) Suppose that $\{x_3^n, n \in \mathbb{N}\}$ is bounded. By Bolzano–Weierstrass Theorem, there exists a subsequence of $(x_1^n, x_2^n, x_3^n, p_1^n, p_2^n, p_3^n)$ converging to $(x_1^*, x_2^*, x_3^*, p_1^*, p_2^*, p_3^*)$. By the uniform convergence theorem, one can verify that $(x_1^*, x_2^*, x_3^*, p_1^*, p_2^*, p_3^*)$ is also a feasible solution and $f^{F^*}(d, X) = f^*$, where $F^* = [x_1^*, p_1^*; x_2^*, p_2^*; x_3^*, p_3^*]$, and thus $(x_1^*, x_2^*, x_3^*, p_1^*, p_2^*, p_3^*)$ is an optimal solution.
- (b) Suppose that $\{x_3^k, k \in \mathbb{N}\}$ is unbounded. There exists a subsequence of $(x_1^{n_j}, x_2^{n_j}, x_3^{n_j}, p_1^{n_j}, p_2^{n_j}, p_3^{n_j})$ such that it converges to $(x_1^*, x_2^*, \infty, p_1^*, p_2^*, p_3^*)$. As $p_3^{n_j} \leq (\mu^2 + \sigma^2)/(x_3^{n_j})^2$, letting $n_j \rightarrow \infty$ yields $p_3^* = 0$. Denote by $F^* = [x_1^*, p_1^*; x_2^*, p_2^*]$. Note that $\lim_{n_j \rightarrow \infty} x_3^{n_j} p_3^{n_j} = 0$ because $x_3^{n_j} p_3^{n_j} \leq (\mu^2 + \sigma^2)/x_3^{n_j}$. One can verify that $f^{F^*}(d, X) = \lim_{n_j \rightarrow \infty} f^{F^{n_j}}(d, X) = f^*$. It also implies

$$\sum_{i=1}^2 p_i^* x_i^* = \lim_{n_j \rightarrow \infty} \sum_{i=1}^2 p_i^{n_j} x_i^{n_j} = \lim_{n_j \rightarrow \infty} \sum_{i=1}^3 p_i^{n_j} x_i^{n_j} = \mu$$

and

$$\sum_{i=1}^2 p_i^* (x_i^*)^2 = \lim_{n_j \rightarrow \infty} \sum_{i=1}^2 p_i^{n_j} (x_i^{n_j})^2 \leq \lim_{n_j \rightarrow \infty} \sum_{i=1}^3 p_i^{n_j} (x_i^{n_j})^2 = \mu^2 + \sigma^2.$$

Therefore, $(x_1^*, x_2^*, d, p_1^*, p_2^*, 0)$ is a feasible solution of the problem (3.11) and the optimal value is attained.

Combining the above two cases, we complete the proof of (i).

- (ii) We consider the same two cases (a) and (b) as in the proof of (i). For case (a), it is obvious that $(x_1^*, x_2^*, x_3^*, p_1^*, p_2^*, p_3^*)$ is an optimal solution. For case (b), we have $f^{F^*}(d, X) = f^*$ and $F^* = [x_1^*, p_1^*; x_2^*, p_2^*]$ where $(x_1^*, x_2^*, d, p_1^*, p_2^*, 0)$ is a feasible solution of the problem (3.11). This implies $\sum_{i=1}^2 p_i^* (x_i^*)^2 \leq \mu^2 + \sigma^2$. We will show $\sum_{i=1}^2 p_i^* (x_i^*)^2 = \mu^2 + \sigma^2$ by contradiction, which implies $(x_1^*, x_2^*, d, p_1^*, p_2^*, 0)$ is an optimal solution. Suppose $\sum_{i=1}^2 p_i^* (x_i^*)^2 < \mu^2 + \sigma^2$, we consider the following five cases.
 - (b.i) If $0 \leq x_1^* < x_2^* < d$, define $G = [x_1^*, p_1^* + \delta; x_2^* + \varepsilon, p_2^* - \delta]$, where $\varepsilon = (x_2^* - x_1^*)\delta/(p_2^* - \delta)$ and $\delta \in (0, p_2^*)$ is small enough such that $\varepsilon \in (0, d - x_2^*)$ and $\mathbb{E}^G[X^2] \leq \mu^2 + \sigma^2$. In this case, one can verify that $f^G(d, X) > f^*$ which yields a contradiction to that $(x_1^*, x_2^* + \varepsilon, d, p_1^* + \delta, p_2^* - \delta, 0)$ is a feasible solution of the problem (3.11).
 - (b.ii) If $0 \leq x_1^* = x_2^* < d$, then it's a degenerate distribution and $0 < \mu < d$. Define $G = [\mu - \varepsilon, \frac{1}{2}; \mu + \varepsilon, \frac{1}{2}]$, where $\varepsilon \in (0, \mu)$ is small enough such that $\mu + \varepsilon \leq d$ and $\mathbb{E}^G[X^2] \leq \mu^2 + \sigma^2$. In this case, one can verify that $f^G(d, X) > f^*$ which yields a contradiction to that $(\mu - \varepsilon, \mu + \varepsilon, d, \frac{1}{2}, \frac{1}{2}, 0)$ is a feasible solution of the problem (3.11).
 - (b.iii) If $0 < x_1^* < x_2^* = d$, define $G = [x_1^* - \varepsilon, p_1^* - \delta; x_2^*, p_2^* + \delta]$, where $\varepsilon = (x_2^* - x_1^*)\delta/(p_1^* - \delta)$ and $\delta \in (0, p_1^*)$ are small enough such that $\varepsilon \in (0, x_1^*)$ and $\mathbb{E}^G[X^2] \leq \mu^2 + \sigma^2$. In this case, one can verify that $f^G(d, X) > f^*$ which yields a contradiction.
 - (b.iv) If $0 < x_1^* = x_2^* = d$, then it's a degenerate distribution and $0 < \mu = d$. Define $G = [\mu - \varepsilon_1, q_1; \mu + \varepsilon_2, q_2]$, which satisfies $y_1 q_1 + y_2 q_2 = \mu = d, q_i > 0, i = 1, 2$ and $\varepsilon_1 \in (0, \mu)$, then $\varepsilon_2 = \frac{q_1}{1 - q_1} \varepsilon_1$. There exist an ε_1 and a q_1 that are small enough such that $\mathbb{E}^G[X^2] \leq \mu^2 + \sigma^2$ and $f^G(d, X) > f^*$, a contradiction.
 - (b.v) If $0 = x_1^* < x_2^* = d$, then $p_1^* = 1 - \mu/d$ and $p_2^* = \mu/d$. One can calculate that

$$f^* = e_\alpha^{F^*}(X) = \frac{\alpha \mu}{(1 - \alpha)(1 - \mu/d) + \alpha \mu/d}.$$

In this case, if $(1 + \theta)[(1 - \alpha)(1 - \mu/d) + \alpha \mu/d]^2 > \alpha(1 - \alpha)$, define $G = [0, p_1^* + \delta; d, p_2^* - 2\delta; 2d, \delta]$, where $\delta \in [0, p_2^*/2]$ is small enough such that

$$\delta < \frac{\alpha(1 - \alpha) - (1 + \theta)[(1 - \alpha)(1 - \mu/d) + \alpha \mu/d]^2}{(1 + \theta)(1 - 2\alpha)[(1 - \alpha)(1 - \mu/d) + \alpha \mu/d]}$$

Table 1. Optimal deductibles and optimal values: $\mu = 15, \sigma = 5$.

		$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$
$\theta = 0.8$	d^*	15.56	17.79	∞	∞
	Optimal value	19.00	18.32	17.18	16.02
$\theta = 0.6$	d^*	14.62	16.23	21.01	∞
	Optimal value	18.51	18.00	17.17	16.02
$\theta = 0.4$	d^*	13.37	14.57	17.14	∞
	Optimal value	17.91	17.55	17.00	16.02
$\theta = 0.2$	d^*	11.16	12.09	13.68	18.33
	Optimal value	17.08	16.87	16.56	16.00
$\theta = 0.1$	d^*	8.54	9.46	10.89	13.85
	Optimal value	16.48	16.35	16.15	15.82

and $\mathbb{E}^G[X^2] \leq \mu^2 + \sigma^2$. We have

$$\begin{aligned}
 f^G(d, X) &= e_a^G(X \wedge d) + (1 + \theta)\mathbb{E}^G[(X - d)_+] \\
 &= \frac{\alpha\mu - \alpha d\delta}{(1 - \alpha)(1 - \mu/d + \delta) + \alpha(\mu/d - \delta)} + (1 + \theta)d\delta \\
 &> \frac{\alpha\mu}{(1 - \alpha)(1 - \mu/d) + \alpha\mu/d} = f^*.
 \end{aligned}$$

This yields a contradiction.

Combining the above five cases, we complete the proof. □

5. Numerical examples

This section provides numerical analyses of the problem (3.1). We study the impacts of the parameters θ, α , and (μ, σ) on the optimal reinsurance design. After that, we compare our robust results with those obtained in the classical reinsurance model when the loss distributions are assumed to be Gamma, Lognormal, and Pareto distributions, respectively. To have more insights into model uncertainty, we further compare our results with the robust reinsurance design with VaR and CVaR in Liu and Mao (2022).

5.1. Impacts of parameters

Tables 1–3. give the optimal deductibles and the optimal values of the distributionally robust reinsurance problem (2.4) for three pairs of (μ, σ) , where μ is 15 and σ is 5, 10, 20.

Recall that the expected value premium principle is defined as $\pi(I(X)) = (1 + \theta)\mathbb{E}[I(X)]$. This implies that for the same reinsurance coverage, the larger the θ is, the more expensive the reinsurance will be. In other words, larger θ would motivate the insurers to retain more risks by themselves instead of entering a reinsurance contract. Hence, the optimal deductible d^* should increase in the same direction with θ . In addition, the confidence level α of an α -expectile represents the risk tolerance of the insurer. The higher the α is, the more risk-sensitive the insurer is. Thus, the insurer would like to transfer more risk to the reinsurer by choosing a smaller deductible. The observations we made in Tables 1 and 2 align with our intuitions. The same logic applies to Table 3, but it is interesting to notice that when the insurers face significant uncertainty (large σ), they would prefer to transfer all risks to the reinsurer regardless of the price (see columns $\alpha = 0.9$ and $\alpha = 0.8$ of Table 3). Moreover, when the optimal contract is a zero-deductible plan ($d^* = 0$), then the corresponding objective function value reduces to $(1 + \theta)\mu$, which is

Table 2. Optimal deductibles and optimal values: $\mu = 15, \sigma = 10$.

		$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$
$\theta = 0.8$	d^*	16.11	20.58	∞	∞
	Optimal value	23.00	21.63	19.36	17.04
$\theta = 0.6$	d^*	14.24	17.50	26.90	∞
	Optimal value	22.02	21.00	19.34	17.04
$\theta = 0.4$	d^*	11.74	14.13	19.28	∞
	Optimal value	20.81	20.10	19.00	17.04
$\theta = 0.2$	d^*	0	0	0	21.69
	Optimal value	18	18	18	17.00
$\theta = 0.1$	d^*	0	0	0	0
	Optimal value	16.50	16.50	16.50	16.50

Table 3. Optimal deductibles and optimal values: $\mu = 15, \sigma = 20$.

		$\alpha = 0.9$	$\alpha = 0.8$	$\alpha = 0.7$	$\alpha = 0.6$
$\theta = 0.8$	d^*	0	0	∞	∞
	Optimal value	27	27	23.72	19.07
$\theta = 0.6$	d^*	0	0	38.92	∞
	Optimal value	24	24	23.68	19.07
$\theta = 0.4$	d^*	0	0	0	∞
	Optimal value	21	21	21.00	19.07
$\theta = 0.2$	d^*	0	0	0	0
	Optimal value	18	18	18	18
$\theta = 0.1$	d^*	0	0	0	0
	Optimal value	16.50	16.50	16.50	16.50

only relevant to the safety loading factor (θ) and the expected loss (e.g. the two rows with $\theta = 0.1$ and $\theta = 0.2$ in Table 3).

Intuitively, when both the price of the reinsurance is expensive (large θ) and the insurer is not risk-sensitive (small α), then the insurer would prefer not to purchase any reinsurance. We can verify this result by looking at the right upper corner of all three tables where $d^* = \infty$. Numerically, we set d^* to be ∞ when the plot of the objective function values exhibits a decreasing yet converging trend as d increases. We verified our results by examining each of such scenarios with d up to 1000, which should be sufficient since the probability of a positive payoff for $d = 1000$ would be less than 1%.

5.2. Comparison with classical reinsurance model

Here, we compare the optimal deductibles and the optimal objective function values obtained in our distributionally robust model with those obtained in the classical reinsurance model. We assume the loss random variable in the classical reinsurance model follows the commonly used distributions in insurance: Gamma, Lognormal, and Pareto distributions.

- (i) (Lognormal distribution) Suppose that X follows a lognormal distribution with $\ln(X) \sim N(\mu, \sigma^2)$. Then $E[X] = e^{\mu + \sigma^2/2}$ and $\text{Var}(X) = e^{\sigma^2 + 2\mu}(e^{\sigma^2} - 1)$.

Table 4. Comparison with nonrobust case with $\mu = 15, \alpha = 0.9, \theta = 0.2$.

		Robust	Lognormal	Pareto	Gamma
$\sigma = 3$	d^*	12.70	12.38	12.98	12.35
	Optimal value	16.25	15.78	15.45	15.81
	Premium	3.65	3.43	2.48	3.49
	Premium/Risk ratio	0.2246	0.2173	0.1605	0.2207
$\sigma = 5$	d^*	11.16	10.71	11.99	10.58
	Optimal value	17.08	16.20	15.67	16.28
	Premium	6.09	5.53	3.68	5.75
	Premium/Risk ratio	0.3566	0.3414	0.2348	0.3532
$\sigma = 10$	d^*	0	7.34	10.40	6.44
	Optimal value	18.00	16.92	16.00	17.18
	Premium	18	9.63	5.61	10.80
	Premium/Risk ratio	1	0.5691	0.3506	0.6286
$\sigma = 20$	d^*	0	3.7	9.15	1.16
	Optimal value	18.00	17.54	16.26	17.92
	Premium	18	13.87	7.12	16.78
	Premium/Risk ratio	1	0.7907	0.4379	0.9364

(ii) (Pareto distribution) Suppose that X follows a Pareto distribution with cumulative distribution function $F(x) = 1 - (\tau/x)^\beta$ for $x \geq \tau$, where $\beta > 1$. Then $\mathbb{E}[X] = \frac{\beta\tau}{\beta - 1}$ and $\text{Var}(X) = \frac{\beta\tau^2}{(\beta - 1)^2(\beta - 2)}$.

(iii) (Gamma distribution) Suppose that X follows a gamma distribution with density function

$$f(x) = \frac{\tau^\gamma x^{\gamma-1} e^{-\tau x}}{\Gamma(\gamma)}, \quad x > 0,$$

where $\gamma, \tau > 0$, and Γ is the Gamma function defined by

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt.$$

Then $\mathbb{E}[X] = \gamma/\tau$ and $\text{Var}(X) = \gamma/\tau^2$.

Recall that the classical reinsurance model corresponding to our distributionally robust model (2.5) is as follows:

$$\min_{d \geq 0} \{e_\alpha(X \wedge d) + (1 + \theta)\mathbb{E}[(X - d)_+]\},$$

where X follows a precise distribution. In order to make comparisons with our robust results, for each pair of (μ, σ) studied in the previous section, the parameters of the aforementioned models are set such that $\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$. Table 4 gives the results for different values of σ , and in order to have a better illustration on the effect of model uncertainty, we also include the premium/risk ratio under each model. The premium is the price of the reinsurance contract, and the risk retained by the insurer is simply the optimal values of our objective functions. It is not surprising to observe that under both robust and non-robust cases, the premium/risk ratio is increasing with the riskiness of loss variable, measured by sigma. For $\sigma = 3, 5$, the optimal deductible is strictly positive, and the premium/risk ratio in the robust case is moderately larger than those in the non-robust case. However, if σ is large, the optimal deductible is zero and the premium/risk ratio is much larger than those in the non-robust case. The large differences in the premium/risk ratios between robust and non-robust models illustrate the catastrophic

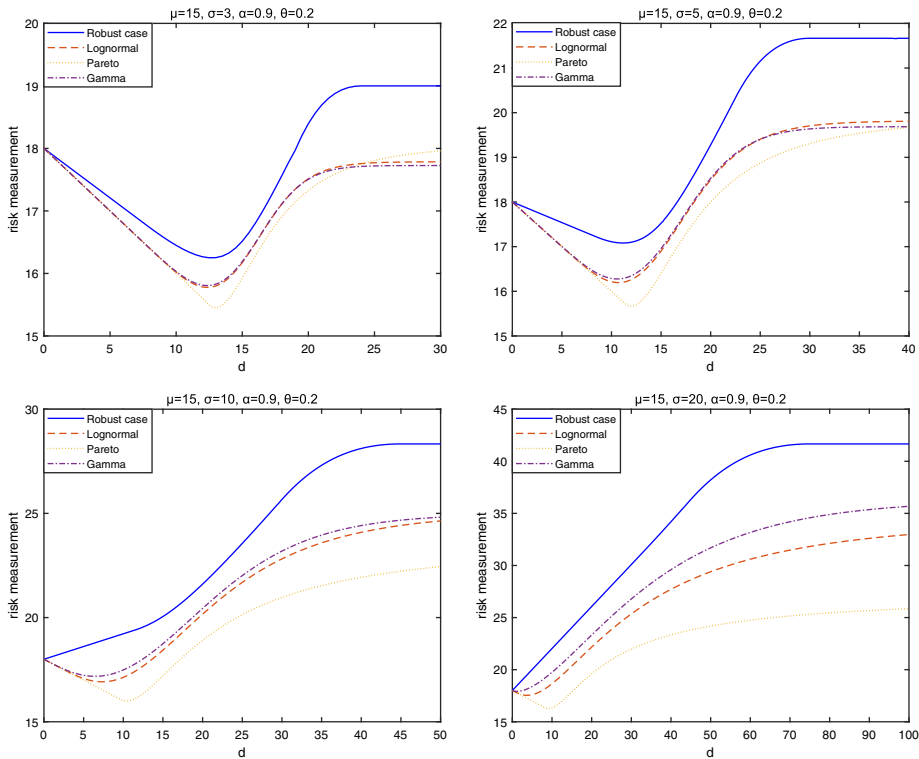


Figure 2. Comparison of risk measurement values in robust case and non-robust case.

consequences if the insurer failed to select the correct model, and the insurers need to be alert to model uncertainty. However, due to the limited data in the tail of the loss distribution, determining the true model is rather difficult, if not impossible, and our results under the robust design suggest the insurer should take more cautious actions.

Figure 2 plots the values of the objective function with respect to different deductibles d . The robust-case curve corresponds to the objective function in our distributionally robust reinsurance model, and the other three curves correspond to the classical reinsurance model when the distribution function is precisely known as Gamma, Lognormal, and Pareto distribution. The objective value in the robust case is consistently larger due to the model setting, and we do confirm that the risk may be underestimated if the distributional uncertainty is ignored. From the graphs in the first row of Figure 2, it is interesting to observe that the optimal reinsurance contract under the robust case is not necessarily the most conservative one in terms of the amount of loss been preserved by the insurer. However, when σ becomes large, the optimal contract under the robust case eventually becomes the zero-deductible plan, whereas the optimal contracts under other three cases may still have moderate deductibles. This implies that a significant portion of the risk may be unintentionally held by the insurer if the optimal design is determined by a misspecified loss distribution. For example, as illustrated in the bottom right graph, if Pareto distribution is misused, the value of the expectile function will be underestimated by up to 25%. Model uncertainty plays a crucial role when the underlying risk is significant, and the distributionally robust optimization is able to provide a conservative benchmark.

5.3. Comparison with distributionally robust reinsurance model with VaR/CVaR

In this section, we compare optimal deductibles and optimal values obtained in the distributionally robust model with expectile with the model in Liu and Mao (2022) under VaR/CVaR. Tables 5 and 6 give the

Table 5. Comparison of optimal deductibles and optimal values: $\mu = 15$ and $\theta = 0.2$.

		CVaR _{0.99}	$e_{0.99}$	$e_{0.991}$	$e_{0.993}$	$e_{0.995}$
$\sigma = 1$	d^*	14.106	14.116	14.115	14.113	14.109
	Optimal value	15.447	15.444	15.445	15.445	15.446
$\sigma = 3$	d^*	12.317	12.349	12.347	12.338	12.330
	Optimal value	16.342	16.334	16.334	16.336	16.338
$\sigma = 5$	d^*	10.528	10.584	10.577	10.563	10.522
	Optimal value	17.236	17.222	17.224	17.227	17.229
$\sigma = 10$	d^*	0	0	0	0	0
	Optimal value	18	18	18	18	18

Table 6. Comparison of optimal deductibles and optimal values: $\mu = 15$ and $\theta = 0.2$.

		CVaR _{0.8}	$e_{0.8}$	$e_{0.81}$	$e_{0.85}$	$e_{0.9}$
$\sigma = 1$	d^*	14.106	14.418	14.395	14.313	14.232
	Optimal value	15.447	15.374	15.379	15.397	15.416
$\sigma = 3$	d^*	12.317	13.254	13.186	12.944	12.697
	Optimal value	16.342	16.123	16.137	16.191	16.249
$\sigma = 5$	d^*	10.528	12.089	11.973	11.576	11.162
	Optimal value	17.236	16.871	16.895	16.985	17.082
$\sigma = 10$	d^*	0	0	0	0	0
	Optimal value	18	18	18	18	18

results for four different values of σ with $\mu = 15$ and $\theta = 2$. Following the conclusion in Bellini and Bernardino (2017) that “for the most common distributions, expectiles are closer to the center of the distribution than the corresponding quantiles,” we choose a series of larger α 's for e_α than for VaR $_\alpha$ to make the comparison. Table 5 compares the optimal deductibles and optimal values when $\alpha = 0.99$ for VaR and $\alpha = 0.99, 0.991, 0.993, 0.995$ for the expectile. The results suggest that for the same level α , the optimal deductible based on VaR is always smaller than that based on expectile, which means that the VaR users are more conservative and they prefer to transfer more risk to a reinsurer than the expectile user with the same level. Table 6 compares the optimal deductibles and optimal values when $\alpha = 0.8$ for VaR and $\alpha = 0.8, 0.81, 0.85, 0.9$ for the expectile. In this case, we have similar observations, but the optimal deductibles for VaR user and expectile user vary more significantly as the difference between the two risk measures, expectile and CVaR, gets larger for a smaller level α . Both tables suggest that when σ is large, the optimal deductibles are zeros and the optimal values are identical. Intuitively, a large σ means more uncertainty, and the insurer would rather transfer all risk to the reinsurer concerning the remarkable uncertainty. In addition, we also want to mention that the optimal deductibles and the optimal values of CVaR at level $\alpha = 0.8$ and 0.99 are identical since the result based on CVaR satisfies the hybrid property. This is because that VaR/CVaR-user in distributionally robust reinsurance is indifferent to the level α satisfying $\sigma^2/\mu^2 < \theta \leq \alpha/(1 - \alpha)$. In sharp contrast, the optimal deductible based on expectile is continuous in the parameter α . From this perspective, the problem based on expectile is more reasonable.

6. Concluding remarks

In this paper, we investigate a distributionally robust reinsurance problem with expectile under model uncertainty, where the distribution of loss is partially known in the sense that only its mean and variance

are known. By showing that the worst-case distribution must belong to the set of three-point distributions, we reduced the infinite-dimensional minimax problem to a finite-dimensional optimization problem, which is a tractable optimization problem. By comparing the results with the classical reinsurance problem, the importance of including model uncertainty is presented, and we demonstrate the consequence of model misspecification may be severe if the underlying risk is significant. In the end, we want to point out that the characterization of the explicit solution of the worst-case distribution is very challenging, and we leave it for future work.

Acknowledgments. The authors thank the Editor and three anonymous referees for their insightful comments, which helped improve the paper.

References

- Arrow, K.J. (1963) Uncertainty and the welfare economics of medical care. *American Economic Review*, **53**(3), 941–973.
- Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999) Coherent measures of risk. *Mathematical Finance*, **9**(3), 203–228.
- Asimit, A.V., Bignozzi, V., Cheung, K.C., Hu, J. and Kim, E.-S. (2017) Robust and pareto optimality of insurance contracts. *European Journal of Operational Research*, **262**, 720–732.
- Bellini, F. and Bignozzi, V. (2015) On elicitable risk measures. *Quantitative Finance*, **15**(5), 725–733.
- Bellini, F. and Bernardino, E. (2017) Risk management with expectiles. *The European Journal of Finance*, **23**(6), 487–506.
- Bellini, F., Klar, B., Müller, A. and Gianin, E.R. (2014) Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, **54**, 41–48.
- Bettels, S., Kim, S. and Weber, S. (2022) Multinomial backtesting of distortion risk measures.
- Birghila, C. and Pflug, G.Ch. (2019) Optimal XL-insurance under Wasserstein-type ambiguity. *Insurance: Mathematics and Economics*, **88**, 30–43.
- Borch, K. (1960) An attempt to determine the optimum amount of stop loss reinsurance. *Transactions of the 16th International Congress of Actuaries I*, pp. 597–610.
- Cai, J. and Chi, Y. (2020) Optimal reinsurance designs based on risk measures: A review. *Statistical Theory and Related Fields*, **4**(1), 1–13.
- Cai, J., Lemieux, C. and Liu, F. (2016) Optimal reinsurance from the perspectives of both an insurer and a reinsurer. *ASTIN Bulletin*, **46**(3), 815–849.
- Cai, J. and Tan, K.S. (2007) Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. *ASTIN Bulletin*, **37**(1), 93–112.
- Cai, J. and Weng, C. (2016) Optimal reinsurance with expectile. *Scandinavian Actuarial Journal*, **2016**(7), 624–645.
- Cheung, K.C., Sung, K.C.J., Yam, S.C.P. and Yung, S.P. (2014) Optimal reinsurance under general law-invariant risk measures. *Scandinavian Actuarial Journal*, **2014**, 72–91.
- Chi, Y. and Tan, K.S. (2011) Optimal reinsurance under VaR and CVaR risk measures: A simplified approach. *ASTIN Bulletin*, **41**(2), 487–509.
- Cui, W., Yang, J. and Wu, L. (2013) Optimal reinsurance minimizing the distortion risk measure under general reinsurance premium principles. *Insurance: Mathematics and Economics*, **53**(1), 74–85.
- El Ghaoui, L., Oks, M. and Oustry, F. (2003) Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, **51**(4), 543–556.
- Embrechts, P., Mao, T., Wang, Q. and Wang, R. (2021) Bayes risk, elicibility, and the Expected Shortfall. *Mathematical Finance*, **31**(4), 1190–1217.
- Föllmer, H. and Schied, A. (2002) Convex measures of risk and trading constraints. *Finance and Stochastics*, **6**(4), 429–447.
- Frittelli, M. and Rossaza Gianin, E. (2002) Putting order in risk measures. *Journal of Banking and Finance*, **26**, 1473–1486.
- Gavagan, J., Hu, L., Lee, G., Liu, H. and Weixel, A. (2022) Optimal reinsurance with model uncertainty. *Scandinavian Actuarial Journal*, **2022**(1), 29–48.
- Gneiting, T. (2011) Making and evaluating point forecast. *Journal of the American Statistical Association*, **106**(494), 746–762.
- Hu, X., Yang, H. and Zhang, L. (2015) Optimal retention for a stop-loss reinsurance with incomplete information. *Insurance: Mathematics and Economics*, **65**, 15–21.
- Jagannathan, R. (1977) Minimax procedure for a class of linear programs under uncertainty. *Operations Research*, **25**(1), 173–177.
- Kratz, M., Lok, Y.H. and McNeil, A. (2018) Multinomial VaR backtests: A simple implicit approach to backtesting expected shortfall. *Journal of Banking and Finance*, **88**, 393–407.
- Kuan, C., Yeh, J. and Hsu, Y. (2009) Assessing value at risk with CARE, the conditional autoregressive expectile models. *Journal of Econometrics*, **150**, 261–270.
- Liu, H. and Mao, T. (2022) Distributionally robust reinsurance with Value-at-Risk and Conditional Value-at-Risk. *Insurance: Mathematics and Economics*, **107**, 393–417.
- Li, J. (2018) Closed-form solutions for worst-case law invariant risk measures with application to robust portfolio optimization. *Operations Research*, **66**(6), 1533–1541.

- Natarajan, K., Sim, M. and Uichanco, J. (2010) Tractable robust expected utility and risk models for portfolio optimization. *Mathematical Finance*, **20**(4), 695–731.
- Newey, W. and Powell, J. (1987) Asymmetric least squares estimation and testing. *Econometrica*, **55**, 819–847.
- Pflug, G.Ch., Timonina-Farkas, A. and Hochrainer-Stigler, S. (2017) Incorporating model uncertainty into optimal insurance contract design. *Insurance: Mathematics and Economics*, **73**, 68–74.
- Schied, A., Föllmer, H. and Weber, S. (2009) Robust preferences and robust portfolio choice. In *Mathematical Modelling and Numerical Methods in Finance* (eds. P. Ciarlet, A. Bensoussan and Q. Zhang), Handbook of Numerical Analysis, vol. 15, pp. 29–88.
- Weber, S. (2006) Distribution-invariant risk measures, information, and dynamic consistency. *Mathematical Finance*, **16**(2), 419–441.
- Ziegel, J.F. (2016) Coherence and elicibility. *Mathematical Finance*, **26**(4), 901–918.