Insurance: Mathematics and Economics 49 (2011) 175–187

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Optimality of general reinsurance contracts under CTE risk measure

Ken Seng Tan [a,b], Chengguo Weng [a], Yi Zhang [c,*]

[a] Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, N2L 3G1, Canada

[b] China Institute for Actuarial Science, Central University of Finance and Economics, China

[c] Department of Mathematics, Zhejiang University, Hangzhou, 310027, China

ARTICLE INFO

Article history:
Received September 2009
Received in revised form
March 2011
Accepted 21 March 2011

Keywords:
Optimal reinsurance
Ceded loss function
Conditional tail expectation (CTE)
Expectation premium principle
Convex analysis
Lagrangian method
Directional derivative
Subdifferential

ABSTRACT

By formulating a constrained optimization model, we address the problem of optimal reinsurance design using the criterion of minimizing the conditional tail expectation (CTE) risk measure of the insurer's total risk. For completeness, we analyze the optimal reinsurance model under both binding and unbinding reinsurance premium constraints. By resorting to the Lagrangian approach based on the concept of directional derivative, explicit and analytical optimal solutions are obtained in each case under some mild conditions. We show that pure stop-loss ceded loss function is always optimal. More interestingly, we demonstrate that ceded loss functions, that are not always non-decreasing, could be optimal. We also show that, in some cases, it is optimal to exhaust the entire reinsurance premium budget to determine the optimal reinsurance, while in other cases, it is rational to spend less than the prescribed reinsurance premium budget.

1. Introduction

Since the seminal papers by Borch (1960) and Kahn (1961), the quest for optimal reinsurance has remained a fascinating area of research and it has drawn significant interest from both academicians and practitioners. Numerous creative models have been proposed with elegant mathematical tools, and sophisticated optimization theories have also been used in deriving the optimal solutions to the proposed models. The fascination with the optimality of reinsurance stems from its potential as an effective risk management tool for insurers. Indeed, by resorting to a meticulous choice of reinsurance treaty, it allows the insurer to control better and thereby manage its risk exposure. The use of reinsurance, on the other hand, incurs an additional cost to the insurer in the form of reinsurance premium. Naturally, the larger the expected risk that is transferred to a reinsurer, the higher the reinsurance premium. This implies that an insurer has to deal with the classical risk and reward tradeoff in balancing the amount of risk retained and risk transferred.

In this paper, we assume a single-period setting. The optimal reinsurance treaty is typically determined by solving an optimization problem, which could involve either maximization or minimization, depending on the chosen criterion. For example, one of the most classical results is based on the variance minimization model. It states that pure stop-loss reinsurance is the optimal treaty in the sense that it yields the least variance of the insurer's retained loss among all the treaties with the same pure premium; see, for example Kaas et al. (2001). Another classical result corresponds to the utility maximization model, which is attributed to Arrow (1974). It asserts that stop-loss reinsurance maximizes the expected utility of the insurer, provided that the insurer has a concave utility function.

In recent years, extensive research on optimal reinsurance has been conducted by Kaluszka (2001, 2004a,b, 2005), who derived explicit optimal reinsurance policies on a number of ingenious risk measure based reinsurance models. Other related contributions include Gajek and Zagrodny (2000, 2004), Promislow and Young (2005), Balbás et al. (2009) and the references therein. Recent relevant papers on the expected utility maximization models include Zhou and Wu (2008) and Zhou et al. (2010).[1]

More recently, two important risk measures known as the Value-at-Risk (VaR) and the conditional tail expectation (CTE) have

* Corresponding author.
E-mail addresses: kstan@uwaterloo.ca (K.S. Tan), c2weng@uwaterloo.ca (C. Weng), zhangyi63@zju.edu.cn (Y. Zhang).

[1] Both papers analyze the optimal insurance purchase, but their results could be applied to the optimal reinsurance purchase.

0167-6687/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2011.03.002

been applied to insurance and reinsurance for the determination of optimal policies. This area of research is inspired by the prevalent use of these two risk measures among banks and insurance companies for risk assessment and for determining regulatory capital requirement (see, for example, Wang et al. (2005), Huang (2006), Cai and Tan (2007), Cai et al. (2008), Bernard and Tian (2009), Balbás et al. (2009), and Tan and Weng (2010)). In particular, Bernard and Tian (2009) analyzed the optimal reinsurance contracts under two tail risk measures: a VaR-like risk measure (the probability for the underlying loss to exceed a given threshold) and a CTE-like risk measure (the expected loss over a given threshold); see Remark 2.2 for more detailed comments. Cai and Tan (2007), Cai et al. (2008) and Tan and Weng (2010) derived the optimal reinsurance treaties under the strict definition of VaR and CTE. While the optimal reinsurance obtained in these three papers are explicit, one critical limitation is the lack of generality in that the optimality of the reinsurance designs is confined to reinsurance treaties of specific structure. For example, Cai and Tan (2007) assumed that the feasible ceded loss function is of the form stop-loss, while Cai et al. (2008) and Tan and Weng (2010) restricted to the class of increasing convex functions. Balbás et al. (2009) characterized the optimal reinsurance treaties under a very general risk measure including CTE as one of the special cases.

The objective of the paper is to explicitly derive the optimal solutions over all possible reinsurance treaties using the criterion of minimizing CTE of the insurer's resulting risk. Because of the generality of the optimal reinsurance model, we will see shortly that this is a mathematically more complex problem. In fact, our formulation of the reinsurance model entails us in solving some convex optimization problem over a Hilbert space using the Lagrangian method. Because the objective function is only directionally differentiable but not Gâteaux differentiable, we utilize the concept of directional derivative in searching for the optimal solutions.

It is interesting to note that pure stop-loss reinsurance is always optimal under our CTE minimization model, a result which is consistent with the variance minimization and expected utility maximization reinsurance models. More interestingly, we also establish formally that ceded loss function of other structures (such as those that do not need to be always non-decreasing) could also be optimal. Moreover, it should be emphasized that our proposed reinsurance model is a constrained optimization model in that one of the constraints can be interpreted as either a reinsurance premium budget or an insurer's profitability guarantee. For completeness, we analyze the optimal solutions under both binding and unbinding cases depending on the optimal reinsurance premium expenditure relative to the reinsurance premium budget. Enforcing the reinsurance premium budget constraint to be binding, it facilitates us in establishing the optimal risk and reward profile and hence leads to the insurer's reinsurance efficient frontier. On the other hand, if the reinsurance premium budget constraint does not have to be binding, then there are cases where it is optimal to spend less than the prescribed budget.

The remaining paper is organized as follows. Section 2 gives some preliminaries and describes the setup of the proposed reinsurance models. Section 3 states the solutions to our proposed optimal reinsurance models with an unbinding constraint. Remarks and numerical examples to further elaborate these key results are also provided in the same section. Section 4 discusses the optimal solutions to the binding reinsurance model. Section 5 concludes the paper. Key mathematical background with respect to the optimization theory in Banach spaces, together with some relevant concepts related to the directional derivative are collected in Appendix. The proofs of all the propositions and theorems are also given in the same Appendix.

2. Preliminaries and reinsurance model

Let X denote the (aggregate) loss initially assumed by an insurer. Suppose X is a nonnegative random variable, and identify it by a probability measure Pr on the measurable space $(\Omega, \mathcal{F})$ with $\Omega = [0, \infty)$ and $\mathcal{F}$ being the Borel σ-field on Ω, such that the distribution function F_X of the underlying risk X is defined by $F_X(t) = \Pr\{[0, t]\}$ for $t \geq 0$. It is worth noting that the distribution of the loss random variable defined in such a way it is general enough for modeling a loss distribution. It can be any of a general distribution, not necessarily either continuous or discrete. Denote by $f(X)$ the part of loss transferred from the insurer to a reinsurer in the presence of the reinsurance. The function $f : [0, \infty) \mapsto [0, \infty)$, satisfying $0 \leq f(x) \leq x$ for all $x \geq 0$, is known as the ceded loss function or the indemnification function. Associated with the ceded loss function $f(X)$, we denote $I_f(X) := X - f(X)$ as the retained loss function of the insurer in the presence of reinsurance. Similarly, I_f can also be recognized as a function $I_f : [0, \infty) \mapsto [0, \infty)$. By transferring part of its loss to the reinsurer, the insurer is obligated to pay the reinsurance premium $\Pi(f(X))$ to the reinsurer, where Π is a principle adopted for calculating the reinsurance premium. Consequently, the total cost or the total risk for the insurer in the presence of reinsurance, denoted by $T_f(X)$, is the sum of the retained loss and the reinsurance premium,[2] i.e.,

$$T_f(X) = I_f(X) + \Pi(f(X)) = X - f(X) + \Pi(f(X)). \tag{2.1}$$

In situation where there is no ambiguity on the explicit dependence on the random variable X, we simplify the notation by writing $f(X)$, $I_f(X)$ and $T_f(X)$ as f, I_f and T_f, respectively.

Eq. (2.1) demonstrates clearly the intricate role of the reinsurance treaty f on the resulting total risk T_f. A more conservative insurer could reduce its risk exposure by transferring most of the risk to a reinsurer at the expense of higher reinsurance premium. On the other hand, a more aggressive insurer could reduce the cost of reinsurance by exposing to a greater expected risk. This illustrates the classical tradeoff between risk retained and risk transferred. In determining the optimal reinsurance treaties, one prudent strategy from the insurer's perspective is to minimize the resulting risk exposure $T_f(X)$ in terms of an appropriately chosen risk measure. In this paper, we focus on the risk measure CTE risk measure.

Before providing a formal definition of CTE, it is necessary to define a closely related risk measure known as the Value-at-Risk (VaR):

Definition 2.1. The VaR of a loss random variable Z at a confidence level $1 - \alpha$, $0 < \alpha < 1$, is formally defined as

$$\mathrm{VaR}_\alpha(Z) = \inf\{z \in \mathbb{R} : \Pr(Z \leq z) \geq 1 - \alpha\}. \tag{2.2}$$

In practice, the parameter α typically is a small value such as 5% or even 1%. Consequently, $\mathrm{VaR}_\alpha(Z)$ captures the underlying risk exposure by ensuring that with a high degree of confidence (such as $1 - \alpha$ probability) the loss will not exceed the VaR level. While VaR is intuitive and is widely accepted among financial institutions as a risk measure for market risk, it is often criticized for its inadequacy in capturing the tail behavior of the loss distribution, in addition to its violation of properties such as the subadditivity. To overcome these drawbacks, the risk measure CTE has been proposed. CTE is defined as the expected loss given that the loss falls in the worst α part of the loss distribution.

[2] Alternatively, we can choose to work with the net risk or net loss random variable, $\Gamma(f)$, defined as $\Gamma(f) = T_f - p_0$, where p_0 is the insurance premium collected by the insurer from the policyholders. $\Gamma(f)$ takes into account the insurance premium received by the insurer for underwriting risk X. Because p_0 is a constant, our proposed optimal CTE-based reinsurance models, whether defined via T_f or $\Gamma(f)$, are equivalent due to the translation invariance property of the CTE risk measure.

Definition 2.2. The CTE of a loss random variable Z at a confidence level $1-\alpha$, $0<\alpha<1$, is defined as the average of VaR:

$$\mathrm{CTE}_\alpha(Z)=\frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_q(Z)\mathrm{d}q.$$

Remark 2.1. (a) At this point we caution the readers that the literature on CTE can be quite confusing, as different authors have adopted different name even though they essentially mean the same risk measure. For example, the term "conditional tail expectation" is coined by Wirch and Hardy (1999) while others have used names such as the Tail Conditional Expectation (see Artzner et al., 1999), Conditional Value-at-Risk (CVaR) (see Rockafellar and Uryasev, 2002), Tail Value-at-Risk (TVaR) (see Dhaene et al., 2006) and Expected Shortfall (ES) (see Tasche, 2002 and McNeil et al., 2005).

(b) The formal definition of CTE is also another potential area of confusion. For instance, many authors (see, for example, Landsman and Valdez, 2003 and Dhaene et al., 2006) have defined CTE as $\mathrm{CTE}_\alpha(Z)=\mathrm{E}[Z|Z>\mathrm{VaR}_\alpha(Z)]$. This is, however, not quite correct as Wirch and Hardy (1999) make it clear that the equality $\mathrm{CTE}_\alpha(Z)=\mathrm{E}[Z|Z>\mathrm{VaR}_\alpha(Z)]$ is true only when the distribution of Z is continuous. In fact, $\mathrm{E}[Z|Z>\mathrm{VaR}_\alpha(Z)]$ is not well defined in the case that $Z>\mathrm{VaR}_\alpha(Z)$ with zero probability.

(c) Generally, the risk measure CTE can also be equivalently defined as either of the following ways.

(i) Let $\beta=\inf\{u:\mathrm{VaR}_u(Z)=\mathrm{VaR}_\alpha(Z)\}$, or equivalently $\beta=\Pr\{Z>\mathrm{VaR}_\alpha(Z)\}$, then

$$\mathrm{CTE}_\alpha(Z)=\frac{1}{\alpha}\Big((\alpha-\beta)\mathrm{VaR}_\alpha(Z)+\beta\mathrm{E}\big[Z|Z>\mathrm{VaR}_\alpha(Z)\big]\Big), \tag{2.3}$$

provided that $\{Z>\mathrm{VaR}_\alpha(Z)\}$ has nonzero probability; otherwise, $\mathrm{CTE}_\alpha(Z)=\mathrm{VaR}_\alpha(Z)$.

(ii) Consider the α-upper-tail distribution $\Psi_\alpha(\xi)$ constructed from loss distribution of Z as below:

$$\Psi_\alpha(\xi)=\begin{cases}0, & \text{for } \xi<\mathrm{VaR}_\alpha(Z),\\ \dfrac{\Pr(Z\le\xi)-(1-\alpha)}{\alpha}, & \text{for } \xi\ge\mathrm{VaR}_\alpha(Z).\end{cases} \tag{2.4}$$

$\mathrm{CTE}_\alpha(Z)$ is then defined as the mean of a random variable with $\Psi_\alpha(\xi)$ as its distribution.

For the proof of the equivalence between the above definitions and Definition 2.2, see Weng (2009, p. 14), where additional properties associated with CTE are also discussed.

We remark that both VaR and CTE satisfy the property of translation invariance. A risk measure $\rho(\cdot)$ is said to satisfy the translation invariance property if $\rho(Z+m)=\rho(Z)+m$ for any scalar $m\in\mathbb{R}$ and a loss random variable Z.

We now proceed with our reinsurance model formulation. Suppose that the reinsurance premium uses the expectation principle with a safety loading $\theta>0$, i.e., $\Pi(f)=(1+\theta)\mathrm{E}[f]$. We suppose further that the insurer is seeking an optimal reinsurance that minimizes the CTE of the total risk T_f. Then a plausible reinsurance model can be formulated as follows:

$$\begin{cases}\min\limits_f & \mathrm{CTE}_\alpha(T_f)=\min\limits_f\ \mathrm{CTE}_\alpha\Big(X-f(X)+(1+\theta)\mathrm{E}[f]\Big)\\ \text{s.t.} & 0\le f(x)\le x \quad \text{for all } x\ge 0,\\ & \mathrm{E}[f(X)]\in\left[0,\dfrac{\pi}{1+\theta}\right],\end{cases} \tag{2.5}$$

where α,θ, and π are constants satisfying $0<\alpha<1,\theta>0$, and $0\le\pi\le(1+\theta)\mathrm{E}[X]$. The above optimal reinsurance model generalizes those in Cai and Tan (2007) and Cai et al. (2008) in that the minimization is taken with respect to all possible ceded loss functions. In contrast, the model in Cai and Tan (2007) is confined to the class of stop-loss function so that the reinsurance problem simplifies to a one-dimensional optimization problem while the feasible ceded loss function in Cai et al. (2008) is a class of increasing convex function.

Remark 2.2. The "CTE model" analyzed in Bernard and Tian (2009) is expressed through the insurer's initial wealth W_0 and final wealth W. Under a reinsurance contract with a ceded loss function $f(x)$, the final wealth $W=W_0-(1+\theta)\mathrm{E}[f(X)]-X+f(X)$. Thus, their model essentially can be formulated as follows:

$$\begin{cases}\min\limits_f & \mathrm{E}\left[(W_0-W)\mathbf{1}_{(W_0-W>v)}\right]\\ \text{s.t.} & 0\le f(x)\le x \quad \text{for all } x\ge 0,\\ & \mathrm{E}[f(X)]\in\left[0,\dfrac{\pi}{1+\theta}\right],\end{cases}$$

where v, not related to the $(1-\alpha)$ quantile of the loss, is exogenously specified.

For mathematical convenience, we suppose that X has finite first two moments so that we can restrict to the space $\mathcal{L}^2:=\mathcal{L}^2(\Omega,\mathcal{F},P)$ for the optimal ceded loss functions. Let $\mathcal{Q}=\mathcal{Q}_f\bigcap\mathcal{Q}_\pi$ where

$$\mathcal{Q}_f:=\{f\in\mathcal{L}^2:0\le f(x)\le x \text{ for } x\ge 0\}, \tag{2.6}$$

and

$$\mathcal{Q}_\pi:=\{f\in\mathcal{L}^2:0\le(1+\theta)\mathrm{E}[f]\le\pi\}, \tag{2.7}$$

respectively. Then the reinsurance model (2.5) can be succinctly reformulated as

$$\min_{f\in\mathcal{Q}}\mathrm{CTE}_\alpha(T_f)=\min_{f\in\mathcal{Q}}\mathrm{CTE}_\alpha\Big(X-f(X)+(1+\theta)\mathrm{E}[f]\Big). \tag{2.8}$$

The following section is devoted to analyzing the solution to the optimization problem (2.8) while in Section 4, we will focus on the binding case with the constraint in $\mathcal{Q}_\pi$ replaced by $(1+\theta)\mathrm{E}[f]=\pi$.

3. Optimal reinsurance treaties: the unbinding case

The mathematical challenges of solving the reinsurance model (2.8) directly arise from at least two aspects. First, the model is an infinite-dimensional problem which involves searching for an optimal function instead of the optimal values of a finite number of parameters. Thus, many of the prevalent numerical techniques cannot be applied. Second, there is no analytical expression for the objective function $\mathrm{CTE}_\alpha(T_f)$ for a general feasible ceded loss function f. Recognizing that solving (2.8) directly can be very challenging, we resolve this through an auxiliary model. The auxiliary model, which will be defined in the following subsection, is much more tractable than the original reinsurance model (2.8). More importantly, a key result in Rockafellar and Uryasev (2002) asserts that the solution to the auxiliary model regarding the ceded loss function f is also the solution to our reinsurance model (2.8).

3.1. Auxiliary model and the optimality conditions

To describe the auxiliary model, it requires us to introduce the mapping $G_\alpha(\xi,f):\mathbb{R}\times\mathcal{L}^2\mapsto\mathbb{R}$ such that

$$G_\alpha(\xi,f)=\xi+\frac{1}{\alpha}\mathrm{E}\left[\Big(X-f+(1+\theta)\mathrm{E}[f]-\xi\Big)_+\right], \tag{3.1}$$

where α is the same constant associated with the risk measure CTE in (2.8). The significance of introducing $G_\alpha(\xi,f)$ can be deduced from the following Lemma 3.1, which is a direct consequence of Rockafellar and Uryasev (2002, Theorem 14).

Lemma 3.1. *Minimizing* $\mathrm{CTE}_\alpha(T_f)$ *with respect to* $f \in \mathcal{Q}$ *is equivalent to minimizing* $G_\alpha(\xi, f)$ *over all* $(\xi, f) \in \mathbb{R} \times \mathcal{Q}$*, in the sense that*

$$\min_{f\in\mathcal{Q}} \mathrm{CTE}_\alpha(T_f) = \min_{(\xi,f)\in\mathbb{R}\times\mathcal{Q}} G_\alpha(\xi, f), \tag{3.2}$$

where moreover,

$$(\xi^*, f^*) \in \arg\min_{(\xi,f)\in\mathbb{R}\times\mathcal{Q}} G_\alpha(\xi, f) \tag{3.3}$$

if and only if

$$f^* \in \arg\min_{f\in\mathcal{Q}} \mathrm{CTE}_\alpha(T_f), \quad \text{and} \quad \xi^* \in \arg\min_{\xi\in\mathbb{R}} G_\alpha(\xi, f^*). \tag{3.4}$$

The above lemma states that, in order to find the minimizer of $\mathrm{CTE}_\alpha(T_f)$ over $\mathcal{Q}$, it is sufficient to focus on minimizing $G_\alpha(\xi, f)$ over the product space $\mathbb{R} \times \mathcal{Q}$. The latter optimization problem can be written in a more explicit form as follows:

$$\begin{cases} \min\limits_{(\xi,f)\in\mathbb{R}\times\mathcal{Q}_f} & G_\alpha(\xi, f) \\ & \equiv \xi + \dfrac{1}{\alpha}\mathrm{E}\left[\left(X - f + (1+\theta)\mathrm{E}[f] - \xi\right)_+\right] \\ \text{s.t.} & \mathrm{E}[f] \in [0, \pi/(1+\theta)]. \end{cases} \tag{3.5}$$

By Lemma 3.1, if (ξ^*, f^*) is one solution to problem (3.5), then f^* solves the reinsurance model (2.8), i.e., f^* is one optimal ceded loss function. Comparing to problem (2.8), an obvious advantage of (3.5) is that its goal function is more tractable.

However, obtaining solution to (3.5) is still mathematical challenging since it remains an infinite-dimensional problem. Furthermore, its objective function is not Gâteaux differentiable which implies that the widely used Karush–Kuhn–Tucker Theorem is not helpful to tackle this problem. Our strategy of solving (3.5) involves the following. We demonstrate in Appendix that (3.5) is a convex problem and its goal function $G_\alpha(\xi, f)$ is directionally differentiable with respect to (ξ, f) over its feasible set. This result motivates us in adopting the Lagrangian method based on the directional derivatives to solving (3.5). More specifically, by defining g^* and V, respectively, as

$$g^* = X - f^* + (1+\theta)\mathrm{E}[f^*] - \xi^* \tag{3.6}$$

and

$$V = (1+\theta)\mathrm{E}[f] - \xi - f. \tag{3.7}$$

One of the key results of the paper is to establish the following optimality conditions for the optimization problem (3.5). The proof of these results is relegated to Appendix A.2.

Proposition 3.1. *An element* $(\xi^*, f^*) \in \mathbb{R}\times\mathcal{Q}$ *solves problem* (3.5) *if and only if there exist a constant* $r \in \mathbb{R}$ *and a random variable* $\lambda \in \mathcal{L}^2$ *such that the following three conditions are satisfied:*

C1. $A(\xi, f) \equiv \alpha\Big[\xi + r(1+\theta)\mathrm{E}[f] + \mathrm{E}[\lambda f]\Big] + \mathrm{E}[V\mathbf{1}_{\{g^*>0\}}] + \mathrm{E}[V_+\mathbf{1}_{\{g^*=0\}}] \geq 0, \ \forall(\xi, f) \in \mathbb{R}\times\mathcal{L}^2$;
C2. $\mathrm{E}[\lambda(f - f^*)] \leq 0, \ f \in \mathcal{Q}_f$;
C3. $r\big(\mathrm{E}[f] - \mathrm{E}[f^*]\big) \leq 0$ *for every* $f \in \mathcal{Q}_\pi$.

Using the above proposition, an optimal solution to (3.5) can be deduced by first selecting some potential candidate. The candidate is then shown to be an optimal solution by verifying conditions C1, C2 and C3 as asserted in Proposition 3.1. We will elaborate this procedure in the following subsection.

3.2. *Optimal ceded loss functions*

Throughout this subsection, we assume $\alpha(1+\theta) \leq 1$. Let

$$\pi_\alpha = (1+\theta)\mathrm{E}\left[(X - d_\alpha)_+\right] \tag{3.8}$$

where

$$d_\alpha = \inf\{d : \Pr[X > d] \leq \alpha\}. \tag{3.9}$$

The notation π_θ and d_θ are defined analogously as

$$\pi_\theta = (1+\theta)\mathrm{E}\left[(X - d_\theta)_+\right] \tag{3.10}$$

and

$$d_\theta = \inf\left\{d : \Pr[X > d] \leq \frac{1}{1+\theta}\right\}. \tag{3.11}$$

We emphasize that the condition $\alpha(1+\theta) \leq 1$ is quite mild as in practice both α and θ are typically much smaller than one. Note that the same condition also implies $d_\alpha \geq d_\theta$ so that $\pi_\alpha \leq \pi_\theta$.

To discuss the solution to the reinsurance model (3.5), we proceed by splitting into three cases, depending on the level of the reinsurance premium budget; i.e.

Case (i): $\pi \in (0, \pi_\alpha)$;
Case (ii): $\pi \in [\pi_\alpha, \pi_\theta]$; and
Case (iii): $\pi \in [\pi_\theta, \infty)$.

The solutions to these cases are formally stated in the following three subsections as Theorems 3.1–3.3, respectively. See Appendix A.4 for their proofs.

3.2.1. *Case (i):* $\pi \in (0, \pi_\alpha)$

Theorem 3.1. *Suppose* $\alpha(1+\theta) \leq 1$*. Then all the ceded loss functions* f^* *of the following form are the optimal solutions to the reinsurance model* (3.5)*:*

$$f^*(x) = \begin{cases} 0, & x < \hat{d}, \\ l(x), & x \geq \hat{d} \end{cases} \tag{3.12}$$

where the function $l(x)$ *satisfies*

$$0 \leq l(x) < x - d_\alpha, \quad \text{for } x \geq \hat{d}, \tag{3.13}$$

and the retention $\hat{d} > 0$ *is determined by*

$$\mathrm{E}[f^*] = \frac{\pi}{(1+\theta)}. \tag{3.14}$$

Remark 3.1. We reiterate that Theorem 3.1 only provides solution, if it exists, for $\pi \in (0, \pi_\alpha)$. This is an immediate consequence of conditions (3.13) and (3.14). More explicitly, suppose f^* is an optimal ceded loss function identified by Theorem 3.1, then we must have $\pi \in (0, \pi_\alpha)$ as can be justified below:

$$\begin{aligned} \pi &= (1+\theta)\mathrm{E}[f^*(X)] \\ &= (1+\theta)\mathrm{E}[l(X)\cdot\mathbf{1}_{\{X\geq\hat{d}\}}] \\ &< (1+\theta)\mathrm{E}[(X - d_\alpha)\mathbf{1}_{\{X\geq d_\alpha\}}] \\ &= (1+\theta)\mathrm{E}[(X - d_\alpha)_+] \\ &= \pi_\alpha. \end{aligned}$$

Remark 3.2. The constraint (3.13) states that for $x \geq \hat{d}$, the function $l(x) \geq 0$ is bounded from above by $x - d_\alpha$. Furthermore, when $x = \hat{d}$, we have $\hat{d} > l(\hat{d}) + d_\alpha \geq d_\alpha$. Consequently, the optimal function (3.12) satisfying (3.13) and (3.14) defines a class of ceded loss functions which is bounded from above by the curve $f(x) = (x - d_\alpha)_+$ with a retention larger than d_α and a resulting reinsurance premium equals to the preset budget π. In Fig. 1, the dashed lines depict three samples of such ceded loss functions that are also optimal. In Examples 3.1 and 3.2 (see Section 3.3), we illustrate numerically that in addition to the pure stop-loss function, there exist many other more interesting ceded loss functions that could also be optimal. In particular, these optimal ceded loss functions could have a variety of shapes as long as they

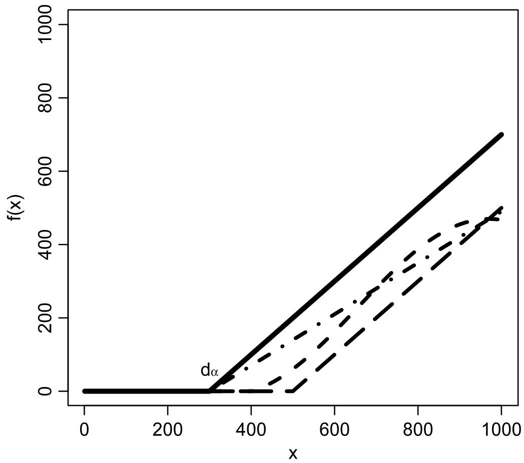

Fig. 1. Three typical optimal ceded loss functions.

are bounded from above by the curve $f(x) = (x - d_\alpha)_+$ and that they have the same reinsurance premium. In Example 3.1, we also present some other ceded loss functions which violate the above boundedness property and these ceded loss functions result in CTE values that are larger than the minimal one.

Remark 3.3. It is worth noting that pure stop-loss treaty is one of the optimal solutions for $\pi < \pi_\alpha$. The reason is as follows. Because $\mathrm{E}[(X-d)_+]$ is continuous in d, there must exist a constant d^* such that $(1+\theta)\mathrm{E}[(X-d^*)_+] = \pi$. Now that $\pi < \pi_\alpha$, d^* must be no less than d_α and hence the stop-loss treaty $f(x) = (x-d^*)_+$ is bounded from above by $f(x) = (x - d_\alpha)_+$. Applying the above Remark 3.2, we immediately conclude that the stop-loss treaty is optimal.

Remark 3.4. According to Remark 3.2, as long as there exists a ceded loss function that is bounded from above by the $f(x) = (x - d_\alpha)_+$ and that with a reinsurance premium equals to π, then it is an optimal solution to our CTE minimization model. Thus, an optimal ceded loss function $f(x)$ may not always be non-decreasing for all $x \geq 0$. We will present in Example 3.2 (see Section 3.3 and Fig. 2) some of these more involved ceded loss functions that are optimal. In some ceded loss functions, the indemnification $f(x)$ drops to zero when the ground loss x exceeds certain threshold, while in other cases, the indemnification starts to decrease after certain threshold. Nevertheless, all of them lead to the same minimal CTE level of the insurer's total risk.

It should be pointed out that issues related to moral hazard could surface when the ceded loss function is not non-decreasing; see Bernard and Tian (2009) for additional discussion on this aspect. Under the expectation premium principle, we could always choose the non-decreasing ceded loss functions in the reinsurance design, since we have identified a class of such optimal solutions in Theorem 3.1. However, the question remains interesting to discuss whether there always exists an optimal non-decreasing indemnification function to our CTE minimization model for other reinsurance premium principles. The promising techniques to tackle this research problem might be the rearrangement inequality and approximation method as developed by Carlier and Dana (2003) for a utility maximization model; see also Carlier and Dana (2005) and Dana and Scarsini (2007). We would leave this topic as one of the future research directions.

3.2.2. Case (ii): $\pi_\alpha \leq \pi \leq \pi_\theta$

Theorem 3.2. *For a given underlying loss random variable X, if there exists a positive constant d^* such that*

$$(1+\theta)\mathrm{E}[(X-d^*)_+] = \pi, \tag{3.15}$$

$$\Pr\{X \geq d^*\} \leq \frac{1}{1+\theta}, \tag{3.16}$$

$$\Pr\{X \geq d^*\} \geq \alpha, \tag{3.17}$$

then $f^* = (X - d^*)_+$ is an optimal ceded loss function to the reinsurance model (3.5).

Remark 3.5. Note that (3.16) and (3.17), respectively, imply $d_\theta \leq d^*$ and $d_\alpha \geq d^*$, where d_α and d_θ are defined in (3.9) and (3.11) respectively. The last condition, in turn, implies that Theorem 3.2 only provides solution, if it exists, for $\pi_\alpha \leq \pi \leq \pi_\theta$.

Remark 3.6. It is interesting to discuss whether there exist some optimal reinsurance treaties other than the stop-loss contract as identified in Theorem 3.2. Example 3.3 (see Section 3.3) analyzes a set of contracts that are either the combination of the quota-share and the stop-loss or the truncated stop-loss treaties with an upper limit. They all have the same reinsurance premium as the optimal stop-loss treaty. However, none of these ceded loss function is optimal.

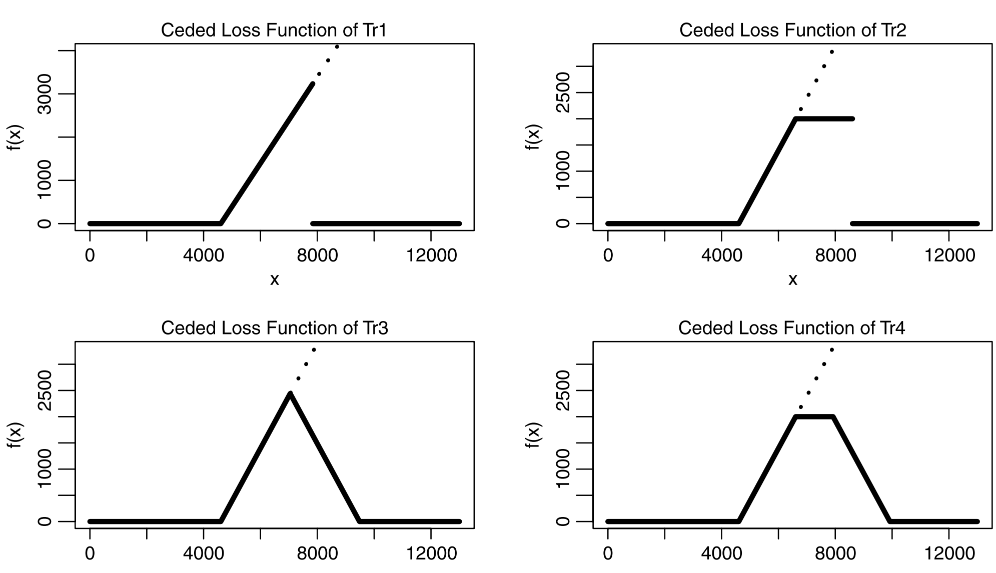

Fig. 2. Some optimal ceded loss functions in Example 3.2 that are not non-decreasing.

Table 1
CTE of some typical reinsurance treaties with $\pi = 10 < \pi_\alpha$.

	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
d_α	4605.17	2995.73	2302.59
π_α	12	60	120
Minimal CTE	*4781.84*	*3839.07*	*3229.25*
Tr1: $f(x) = c^*(x - d_\alpha)_+$	$c^* = \frac{10}{12}$	$c^* = \frac{2}{12}$	$c^* = \frac{1}{12}$
Tr2: $f(x) = (x - \hat{d})_+$	$\hat{d} = 4787.49$	$\hat{d} = 4787.49$	$\hat{d} = 4787.49$
Tr3: $f(x) = (x - d_\alpha)_+ \wedge l_\alpha$	$l_\alpha = 1791.76$	$l_\alpha = 182.32$	$l_\alpha = 87.01$
Tr4: $f(x) = cx$, $c = 1/120$	CTE = 5568.46	CTE = 3972.44	CTE = 3285.06
Tr5: $f(x) = c(x - 1000)_+$, $c = 0.0113$	CTE = 5568.67	CTE = 3977.47	CTE = 3292.17
Tr6: $f(x) = c(x - 1500)_+$, $c = 0.0149$	CTE = 5553.84	CTE = 3968.45	CTE = 3285.66
Tr7: $f(x) = c(x - 2000)_+$, $c = 0.0205$	CTE = 5530.91	CTE = 3954.51	CTE = 3275.59

3.2.3. Case (iii): $\pi \in [\pi_\theta, \infty)$

Theorem 3.3. *Suppose $\alpha(1+\theta) \le 1$ and $\pi \ge \pi_\theta$. Then $f^* = (X - d_\theta)_+$ is an optimal ceded loss function to the reinsurance model (3.5).*

Remark 3.7. According to Theorems 3.2 and 3.3 (see also Remarks 3.3 and 3.5), the pure stop-loss treaty $f^*(x) = (x - d^*)_+$ with the retention d^* satisfying $(1+\theta)\mathrm{E}[(X - d^*)_+] = \pi$ is an optimal ceded loss function for $\pi < \pi_\alpha$. Hence, combining this fact with Theorem 3.3, we see that a pure stop-loss treaty $f^*(X) = (X - d^*)_+$ is optimal for a general reinsurance premium budget π, where the retention d^* is determined by $(1+\theta)\mathrm{E}[(X - d^*)_+] = \min\{\pi, \pi_\theta\}$.

Remark 3.8. Suppose an insurer is willing to spend up to π with $\pi \ge \pi_\theta$ to transfer part of its risk to a reinsurer. Theorem 3.3 asserts that a rational insurer should only spend exactly π_θ. It is not possible to reduce its risk (in terms of smaller CTE) by spending more than π_θ. In fact, this fact is apparent if we restrict to the stop-loss treaties. Under a stop-loss treaty with a retention d, the total loss would be $T_f = X \wedge d + (1+\theta)\mathrm{E}[(X - d)_+]$, which implies

$$\mathrm{CTE}_\alpha(T_f) = d + (1+\theta)\int_d^\infty S_X(x)\mathrm{d}x.$$

Clearly, by the above expression, $\mathrm{CTE}_\alpha(T_f)$, as a function of d, is decreasing on $[0, d_\theta]$ while increasing on $[d_\theta, \infty)$. Thus, it is impossible for stop-loss treaties with reinsurance premium larger than π_θ to be optimal.

3.3. Some numerical examples

In this subsection, we will present some numerical examples to highlight the theoretical results developed in the preceding subsection. These examples will also complement the remarks as discussed in the last subsection. The basic setup of our examples below is based on the following information, which for ease of reference, we refer as the base case parameters set.

(a) *Assume the safety loading factor $\theta = 0.2$ for the reinsurance premium calculation.*
(b) *Suppose the loss random variable X is exponentially distributed with mean $\mu = 1000$ and with survival function S_X and probability density function f_X*

$$S_X(x) = \mathrm{e}^{-\frac{x}{\mu}}, \qquad f_X(x) = \frac{1}{\mu}\mathrm{e}^{-\frac{x}{\mu}}, \quad \text{for } x \ge 0.$$

Thus, it follows from the definition (3.9) that $d_\alpha = S_X^{-1}(\alpha) = -\mu \ln \alpha$ for $0 < \alpha \le 1$. Consequently, it is easy to verify that $\pi_\alpha = (1+\theta)\mathrm{E}[(X - d_\alpha)_+] = (1+\theta)\mu\alpha$. □

Example 3.1. This example aims to illustrate Theorem 3.1 and to support the arguments as presented in Remark 3.2. Using the above base case parameter set and together with $\alpha \in \{1\%, 5\%, 10\%\}$, we have

$$\left.\begin{array}{lll} \pi_\alpha = 12 \text{ and } d_\alpha = 4605.17, & \text{for } \alpha = 1\% \\ \pi_\alpha = 60 \text{ and } d_\alpha = 2995.73, & \text{for } \alpha = 5\% \\ \pi_\alpha = 120 \text{ and } d_\alpha = 2302.56, & \text{for } \alpha = 10\%. \end{array}\right\} \quad (3.18)$$

In order for Theorem 3.1 to be applicable, we assume the reinsurance premium budget $\pi = 10$ so that the condition $\pi \le \pi_\alpha$ is satisfied for all three levels α. We consider seven reinsurance treaties. The first three treaties are

$$\mathrm{Tr1}: f(x) = c^*(x - d_\alpha)_+,$$
$$\mathrm{Tr2}: f(x) = (x - \hat{d})_+,$$
$$\mathrm{Tr3}: f(x) = (x - d_\alpha)_+ \wedge l_\alpha.$$

The remaining four treaties take the form: $f(x) = c(x - d)_+$, for $d = 0$, 1000, 1500 and 2000 and these treaties are labeled, respectively, as Tr4–Tr7. For these reinsurance treaties to be well defined, we have yet to specify their parameter values (such as c^* in Tr1, $\hat{d}$ in Tr2, ...). The required parameter is determined in such a way that the loaded reinsurance premium coincides exactly with the reinsurance premium budget; i.e.

$$(1+\theta)\mathrm{E}[f(X)] = \pi.$$

It is easy to see that while the graphs of the first three treaties are bounded from above by pure stop-loss function $f(x) = (x - d_\alpha)_+$, the other four samples (i.e. Tr4–Tr7) do not satisfy this property. By applying Theorem 3.1 (see also Remark 3.2), we see that treaties Tr1, Tr2 and Tr3 are all optimal reinsurance treaties, while Tr4–Tr7 may not be. Table 1 reports, for each reinsurance treaty, the CTE value of the resulting total loss. Clearly treaties Tr4–Tr7 are not optimal due to the higher CTE values.

Example 3.2. In this example we demonstrate that under our proposed CTE minimization model, the optimal ceded loss functions need not always be non-decreasing, as pointed out in Remark 3.2. We use the same base case parameters set with $\pi = 10$ and $\alpha = 1\%$. From (3.18), we have $d_\alpha = 4605.17$. We consider the following four reinsurance treaties:

- $f(x) = (x - d_\alpha)_+ \cdot \mathbf{1}_{(x \le d_l)}$, where $d_l = 7840.357$.
- $f(x) = \min\{(x - d_\alpha)_+, l\} \cdot \mathbf{1}_{(x \le d_l)}$, where $l = 2000$ and $d_l = 8761.452$.
- $f(x) = (x - d_\alpha)_+ \cdot \mathbf{1}_{(x \le d_l)} + (2d_l - d_\alpha - x)_+ \cdot \mathbf{1}_{(x > d_l)}$, where $d_l = 7045.535$.
- $f(x) = \min[(x - d_\alpha)_+, l] \cdot \mathbf{1}_{(x \le d_l)} + (d_l + l - x)_+ \cdot \mathbf{1}_{(x > d_l)}$, where $l = 2000$ and $d_l = 7922.892$.

Note that the parameters l and d_l in the above ceded loss functions are determined using the condition $(1+\theta)\mathrm{E}[f(X)] = \pi$. These four treaties are plotted in Fig. 2. As a benchmark, we also plot pure stop-loss function of the form $f(x) = (x - d_\alpha)_+$ and this is represented by the dashed curve. Note that by construction, the

Table 2
CTE of some typical reinsurance treaties with $\pi_\alpha \le \pi = 400 \le \pi_\theta$.

	$\alpha = 1\%$	$\alpha = 5\%$	$\alpha = 10\%$
d_α	4605.17	2995.73	2302.59
π_α	12	60	120
Tr1: $f(x) = (x - d^*)_+, d^* = 1098.61$	CTE = 1498.612	CTE = 1498.61	CTE = 1498.61
Tr2: $f(x) = cx, c = 0.3333$	CTE = 4136.780	CTE = 3063.82	CTE = 2601.72
Tr3: $f(x) = c(x - 200)_+, c = 0.4071$	CTE = 3804.54	CTE = 2850.36	CTE = 2439.42
Tr4: $f(x) = c(x - 400)_+, c = 0.4973$	CTE = 3416.77	CTE = 2607.67	CTE = 2259.20
Tr5: $f(x) = c(x - 600)_+, c = 0.6074$	CTE = 2965.17	CTE = 2333.26	CTE = 2061.11
Tr6: $f(x) = c(x - 800)_+, c = 0.7418$	CTE = 2440.47	CTE = 2024.99	CTE = 1846.05
Tr7: $f(x) = c(x - 1000)_+, c = 0.9061$	CTE = 1832.45	CTE = 1681.32	CTE = 1616.23
Tr8: $f(x) = (x - d)_+ \wedge l, d = 1000, l = 2365.46$	CTE = 3639.71	CTE = 2090.92	CTE = 3090.92

above four treaties are bounded from above by the dashed curve and hence according to Remark 3.2 these treaties are optimal. In fact, it is easy to verify that under each of these treaties, the CTE of the resulting total cost T_f is 4781.837, which corresponds to the minimum CTE as obtained in Example 3.1.

Example 3.3. The objective of this example is to illuminate Theorem 3.2 by demonstrating that pure stop-loss reinsurance treaty is optimal while many reinsurance treaties are not necessary optimal, even though these latter treaties are commonly discussed in the optimal reinsurance literature. These results supplement Remark 3.6.

We use the same basic setup as in Example 3.1 with $\alpha \in \{1\%, 5\%, 10\%\}$. These parameter values imply $\pi_\theta = (1+\theta)\mathrm{E}[(X - d_\theta)_+] = \mu = 1000$ and the values of d_α and π_α are given in (3.18) for the corresponding values of α. Furthermore, we set the reinsurance premium budget to $\pi = 400$ so as to ensure $\pi_\alpha < \pi < \pi_\theta$. Using Theorem 3.2, pure stop-loss reinsurance $f(x) = (x - d^*)_+$ satisfying $(1+\theta)\mathrm{E}[(X - d^*)_+] = \pi$ is one of the optimal treaties. The optimal retentions d^*, together with the resulting minimal CTE of the insurer's total loss, are reported in Table 2. In addition, we consider seven other reinsurance treaties and these are labeled as Tr2–Tr8 as shown in Table 2. Note that these treaties are either some combinations of the stop-loss and the quota-share contracts, or a truncated stop-loss treaty. The parameter values of these treaties are selected so that the resulting reinsurance premium coincides with 400. Clearly, none of these treaties are optimal as exemplified by the corresponding higher CTE.

4. Optimal reinsurance model: the binding case

To illustrate the best risk and reward profile for the insurer's optimal selection of reinsurance, this section focuses on the reinsurance model with binding reinsurance premium budget constraint as follows:

$$\begin{cases} \min\limits_{f \in \mathcal{Q}_f} & \mathrm{CTE}_\alpha(T_f), \\ \text{s.t.} & (1+\theta)\mathrm{E}[f] = \pi. \end{cases} \tag{4.1}$$

As previously defined, the notation $\mathcal{Q}_f$ denotes the set of feasible ceded loss functions, i.e. $\mathcal{Q}_f = \{f \in \mathcal{L}^2 : 0 \le f(x) \le x \text{ for } x \ge 0\}$, and π is an exogenous variable representing the reinsurance premium budget. The only difference between the reinsurance model (2.8) and the above (4.1) lies on how we interpret the constraint associated with the reinsurance premium budget π. In the former case, the insurer is willing to spend up to π while in the latter case, the constraint is binding in that the insurer is to spend exactly π on reinsuring its risk. Hence problem (4.1) is more restrictive and we are interested in its solution for $\pi \in (0, \pi_X]$ where $\pi_X = (1+\theta)\mathrm{E}[X]$.

One motivation for considering the above optimization problem (4.1) is that it allows us to address more explicitly the tradeoff between risk and reward. To see this, we will first focus on the objective function in model (4.1) and then on its constraint. Recall that in Footnote 2 on page 4, $\Gamma(f)$ was defined as the insurer's net risk; i.e., $\Gamma(f) = T_f - p_0 = X - f(X) + (1+\theta)\mathrm{E}[f] - p_0$, where p_0 is the insurance premium collected by the insurer from the policyholders. Because of the translation invariance property, we have $\mathrm{CTE}_\alpha(\Gamma(f)) = \mathrm{CTE}_\alpha(T_f) - p_0$. Since p_0 is a constant for a given X, this implies that if f^* is a minimizer of $\mathrm{CTE}_\alpha(T_f)$, it is also a minimizer of $\mathrm{CTE}_\alpha(\Gamma(f))$.

We now shift our attention to the constraint condition in (4.1). The term $b(f) \equiv -\mathrm{E}[\Gamma(f)] = p_0 - \mathrm{E}[X] - \theta\mathrm{E}[f]$ captures the insurer's expected net profit in the presence of reinsurance. Note that the insurer's expected net profit depends on the choice of the ceded loss function. Furthermore, the constraint $\mathrm{E}[f] = \frac{\pi}{1+\theta}$, where $\pi = \frac{1+\theta}{\theta}(p_0 - \mathrm{E}[X] - b(f))$, can be interpreted as the profitability requirement in that once the condition is attained, the resulting optimal ceded loss function f^* ensures a certain prescribed level of expected net profit $b(f^*)$. Consequently, f^* that solves model (4.1) represents the insurer's least risk exposure (as measured by the CTE) for a given level of expected profitability. Hence if model (4.1) is solved repeatedly for each $\pi \in (0, \pi_X]$, where $\pi_X = (1+\theta)\mathrm{E}[X]$, then we trace out pairs of $\big(\mathrm{CTE}_\alpha(\Gamma(f^*)), b(f^*)\big)$ that give the best possible risk and reward tradeoff. This is analogous to the efficient frontier of the Markowitz portfolio mean-variance analysis. For this reason, we refer the curve represented by $\big(\mathrm{CTE}_\alpha(\Gamma(f^*)), b(f^*)\big)$ as the insurer's reinsurance "efficient" frontier. Depending on the risk tolerance of an insurer, the reinsurance efficient frontier facilitates the insurer on its optimal selection of ceded loss function. Bernard and Tian (2009) also discussed the reinsurance efficient frontier under a VaR-like minimization model. For more detailed remarks, please refer to Remark (v) in Example 4.1.

The mathematical technique used to solve the reinsurance model (2.8) can similarly be used to derive the optimal solution to (4.1). This entails reformulating (4.1) as

$$\begin{cases} \min\limits_{(\xi,f)\in\mathbb{R}\times\mathcal{Q}_f} & G_\alpha(\xi, f) \\ \quad \equiv \xi + \dfrac{1}{\alpha}\mathrm{E}\left[\Big(X - f + (1+\theta)\mathrm{E}[f] - \xi\Big)_+\right] \\ \text{s.t.} \quad (1+\theta)\mathrm{E}[f] = \pi. \end{cases} \tag{4.2}$$

If (ξ^*, f^*) are the optimal solutions to (4.2), then f^* is also the optimal solution to (4.1) (see Lemma 3.1). Moreover, the problem (4.2) is convex and thus a ceded loss function f^* is a solution to (4.2) if and only if there exist constants ξ^* and r, and the random variable $\lambda \in \mathcal{L}^2$ such that the three optimality conditions C1, C2 and C3 in Proposition 3.1 are satisfied except with the binding condition $(1+\theta)\mathrm{E}[f] = \pi$ in defining the set $\mathcal{Q}_\pi$. To avoid any confusion, we will define $\mathcal{Q}'_\pi$ as $\mathcal{Q}'_\pi = \{f \in \mathcal{L}^2 : (1+\theta)\mathrm{E}[f] = \pi\}$ to distinguish from the previously introduced notation $\mathcal{Q}_\pi$.

Remark 4.1. Theorems 3.1 and 3.2 indicate that for any given reinsurance premium budget $\pi \in (0, \pi_\theta]$, pure stop-loss treaty $f^*(x) = (X - d^*)_+$, where $(1+\theta)\mathrm{E}[f^*] = \pi$, is an optimal reinsurance solution to the reinsurance model (2.8). Note that the optimal retention d^* is determined such that the resulting reinsurance premium coincides with the reinsurance premium budget π. In other words, the optimal ceded loss function is attained at the reinsurance premium budget. Theorem 3.3, on the other hand, suggests that even if an insurer is willing to spend $\pi \geq \pi_\theta$, the stop-loss treaty is still a possible optimal reinsurance treaty, except that the solution is no longer binding. The optimal retention d^* is always d_θ so that the reinsurance premium is capped at π_θ. Hence it is never rational for an insurer to spend more than π_θ to reinsure its risk. Nevertheless, it is of theoretical interest to examine the solution to our optimal reinsurance model under the binding reinsurance premium budget constraint, as we establish in the following theorem for $\pi \in (0, \pi_X]$.

Theorem 4.1. *Assume $\alpha(1+\theta) \leq 1$ and there exists a constant d^* such that $(1+\theta)\mathrm{E}[(X-d^*)_+] = \pi$ for each $\pi \in (0, \pi_X]$. Then the stop-loss treaty $f^*(x) = (x-d^*)_+$ is an optimal solution to the reinsurance model (4.1).*

Proof. See Appendix A.4. □

Remark 4.2. The above Theorem 4.1 can also be used to explain why the deduction level d_α and the constant d_θ play such an important role in determining solutions to our model (2.5) in the unbinding case. Under a stop-loss treaty with deductible d, the insurer's total loss is

$$\begin{aligned} T_f &\equiv X - f(X) + (1+\theta)\mathrm{E}[f(X)] \\ &= \min\{X, d\} + (1+\theta)\mathrm{E}[(X-d)_+], \end{aligned}$$

and its resulting CTE level is

$$\mathrm{CTE}_\alpha(T_f; d) = \begin{cases} d + (1+\theta)\displaystyle\int_d^\infty S_X(x)\mathrm{d}x, \\ \quad \text{for } d \leq d_\alpha, \\ d_\alpha + \dfrac{1}{\alpha}\displaystyle\int_{d_\alpha}^d S(x)\mathrm{d}x + (1+\theta)\int_d^\infty S(x)\mathrm{d}x, \\ \quad \text{for } d \geq d_\alpha. \end{cases}$$

According to Theorem 4.1, this is the minimal CTE level the insurer can attain by spending a reinsurance premium of $(1+\theta)\mathrm{E}[(X-d)_+]$. Clearly, as a function of d, $\mathrm{CTE}_\alpha(T_f; d)$ is increasing on interval $[0, d_\theta]$, and decreasing on interval $[d_\theta, \infty]$. For d decreases from ∞ to d_θ, the resulting premium of the stop-loss treaty increases from 0 to π_θ (see (3.8)) and $\mathrm{CTE}_\alpha(T_f; d)$ keeps decreasing. Thus, to be optimal under our model (2.5) with an unbinding premium constraint $(1+\theta)\mathrm{E}[f(X)] \leq \pi$ for $\pi \in [0, \pi_\theta]$, a ceded loss function f must exhaust the premium budget π, i.e., $(1+\theta)\mathrm{E}[f(X)] = \pi$. On the other hand, for d decreases from d_θ to 0, the resulting reinsurance premium of the stop-loss treaty increases from π_θ to $(1+\theta)\mathrm{E}[X]$, while $\mathrm{CTE}_\alpha(T_f; d)$ keeps increasing. This implies that, to be optimal under the unbinding case with reinsurance premium budget constraint $(1+\theta)\mathrm{E}[f(X)] \leq \pi$ for $\pi \geq \pi_\theta$, a ceded loss function f must cost a reinsurance premium of π_θ.

Remark 4.3. From the above Theorem 4.1, the ceded loss function $f^*(x) = (x-d^*)_+$ with $(1+\theta)\mathrm{E}[f^*] = \pi$ solves model (4.1) if $\alpha(1+\theta) \leq 1$. Thus the reinsurance efficient frontier is given by

$$\Big\{\big(\mathrm{CTE}_\alpha(\Gamma(f^*)), b(f^*)\big) : f^* = (X-d^*)_+,\ (1+\theta)\mathrm{E}[f^*] = \pi, \\ \text{and } \pi \in (0, \pi_X]\Big\},$$

where $\Gamma(f^*) = (X \wedge d^*) + \pi - p_0$, and $b(f^*) = p_0 - \mathrm{E}[X] - \frac{\theta}{1+\theta}\pi$.

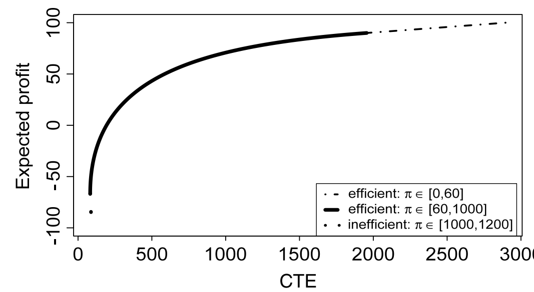

Fig. 3. Risk reward under optimal reinsurance arrangement.

Example 4.1. Using our base case parameter set and together with $\eta = 0.1$ and $\alpha = 5\%$, it is easy to show that $d_\alpha = 2995.73$, $d_\theta = 182.32$, $\pi_\alpha = 60$, $\pi_\theta = 1000$ and $\pi_X \equiv (1+\theta)\mathrm{E}[X] = 1200$. With these parameter values, Theorem 4.1 asserts that to obtain an optimal ceded loss function, we merely need to determine the retention level d^* that satisfies $(1+\theta)\mathrm{E}[(X-d^*)_+] = \pi$ for each $\pi \in (0, \pi_X]$. Under the exponential distribution with mean μ, it is easy to show that

$$d^* = \mu \ln\left(\frac{\mu(1+\theta)}{\pi}\right).$$

Furthermore, it is clear that $\mathrm{CTE}_\alpha(X \wedge d) = d$ for $d \leq d_\alpha \equiv -\mu\ln\alpha$, or equivalently $\pi \geq \pi_\alpha$. For $d \geq d_\alpha$, i.e., $\pi \leq \pi_\alpha$,

$$\begin{aligned} \mathrm{CTE}_\alpha(X \wedge d) &= d_\alpha + \frac{1}{\alpha}\int_{d_\alpha}^\infty \Pr\{X \wedge d > x\}\mathrm{d}x \\ &= d_\alpha + \frac{\mu}{\alpha}\left[\mathrm{e}^{-d_\alpha/\mu} - \mathrm{e}^{-d/\mu}\right] \\ &= \mu(1-\ln\alpha) - \frac{\pi}{\alpha(1+\theta)}. \end{aligned}$$

Thus the reinsurance efficient frontier, $\big(\mathrm{CTE}_\alpha(\Gamma(f^*)), b(f^*)\big)$, is given by

$$\begin{aligned} \mathrm{CTE}_\alpha(\Gamma(f^*)) &= \mathrm{CTE}_\alpha(X \wedge d^*) + \pi - p_0 \\ &= \begin{cases} \mu(1-\ln\alpha) - p_0 + \pi\left[1 - \dfrac{1}{\alpha(1+\theta)}\right], \\ \quad \pi \leq \pi_\alpha, \\ \mu\ln\left(\dfrac{\mu(1+\theta)}{\pi}\right) + \pi - p_0, \\ \quad \pi \geq \pi_\alpha, \end{cases} \\ &= \begin{cases} -\dfrac{47}{3}\pi + 2895.73, & \pi \leq 60, \\ 1000\ln\left(\dfrac{1200}{\pi}\right) + \pi - 1100, & \pi \geq 60, \end{cases} \end{aligned} \tag{4.3}$$

and

$$b(f^*) = p_0 - \mathrm{E}[X] - \frac{\theta}{1+\theta}\pi = 100 - \frac{1}{6}\pi. \tag{4.4}$$

Fig. 3 plots the resulting reinsurance efficient frontier for $\pi \in [0, \pi_X]$. We now conclude the example with the following remarks:

(i) It is striking to note that the reinsurance efficient frontier has a tremendous resemblance to the classical Markowitz mean-variance efficient frontier even though the risk in our reinsurance model is captured by the CTE.

(ii) Without reinsurance, the insurer retains the entire amount of the insurance premium and hence its expected profit margin is 100.[3] This is not surprising since we have assumed that the insurer's loading factor is $\eta = 10\%$. Moreover, the insurer's risk exposure in term of CTE reaches its peak at 2895.72. These values can be obtained by setting $\pi = 0$ in (4.3) and (4.4). However, as the insurer becomes more risk averse and is willing to spend more on reinsurance, its expected profit declines with decreasing CTE risk exposure. This is the classical risk and reward tradeoff. More precisely, as the reinsurance premium budget π increases from 0 to 60, both the expected profit and the CTE decline linearly at the rate of $\frac{1}{6}$ and $15\frac{2}{3}$ from 100 and 2895.72, respectively. The dot-dashed line in Fig. 3 depicts the tradeoff for $\pi \in [0, 60]$.

(iii) When an insurer is willing to spend more than 60 on the reinsurance premium budget, the insurer's expected profit continues to drop linearly. The CTE, on the other hand, continues to decrease but it reaches its minimum at $\pi = \pi_\theta = 1000$. This is depicted as the solid line in Fig. 3. When $\pi > 1000$, the CTE actually increases even though the expected profit still declines. Consequently, it is never rational to spend more than 1000 in reinsuring its risk as already noted in Theorem 3.3. To distinguish these two parts of the frontier, we denote the portion with $\pi \le 1000$ as the efficient frontier while the portion with $\pi > 1000$ as the inefficient frontier, in analogous to the Markowitz model. The efficient and inefficient reinsurance frontiers are depicted in Fig. 3.

(iv) We point out that while $\pi \le 1000$ yields a reinsurance frontier that is "efficient", we also note that for $\pi > 600$, the expected profit of the insurer is negative (see (4.4)). Hence under ordinary circumstances, the insurer will not be spending more than 600 on reinsurance, otherwise it would be prudent of not insuring the risk at all.

(v) A reinsurance efficient frontier with a similar shape has been obtained in Bernard and Tian (2009) (see Fig. 1 in their paper). While their efficient frontier depicts the efficient risk-return profile between the guaranteed insurer's expected final wealth and the minimum probability that the underlying loss X exceeds an exogenously specified threshold, ours clearly describes the risk-return tradeoff between the insurer's expected profit and the minimum risk measure CTE of the insurer's net cost. The insurer's expected final wealth can be recovered by adding a constant to the insurer's expected profit defined in the present paper. Thus, the only difference between these two frontiers lies in the risk measure that is used.

5. Conclusion

In this paper, we generalized the reinsurance models of Cai and Tan (2007) and Cai et al. (2008) by seeking optimal reinsurance design over all possible ceded loss function under the CTE minimization criterion. Under some very mild conditions, analytic optimal solutions were derived for both binding and unbinding reinsurance premium budget constraint. In particular, we found that stop-loss reinsurance treaty is one of the optimal solutions. In some other cases, we also found that more elaborate ceded loss functions (such as those depicted in Fig. 2) could be optimal. For the unbinding reinsurance model, we showed that in some cases the entire reinsurance premium budget should be optimally used in determining the reinsurance design while in other cases, it was rational to spend less than the allowable budget. For the binding reinsurance model, we derived explicitly the classical tradeoff between risk (in term of CTE) and reward (in term of expected profit) and we presented the insurer's reinsurance "efficient" and "inefficient" frontiers.

[3] In practice, the profit margin will be less than 100 since this amount also includes expenses, administration charges, in addition to profits. In our analysis, we ignore these charges for simplicity.

Acknowledgments

Tan acknowledges the research support from the Natural Sciences and Engineering Research Council of Canada. Weng thanks the funding support from the University of Waterloo Start-up Grant and the Waterloo Research institute in Insurance, Securities and Quantitative finance. Zhang acknowledges the research support by the grant NSFC (70871103).

Appendix

The main objective of this Appendix is to establish formally the optimality conditions of the reinsurance model (3.5) as asserted in Proposition 3.1. As the objective functional in model (3.5) is not Gâteaux differentiable, the analysis is quite mathematically involved. We adopt the Lagrangian method and develop the optimality conditions by exploiting the concept of directional derivative. Appendix A.1 provides a brief summary of the mathematical concepts of relevance to subsequent discussions. Detailed descriptions of these topics can be found in Bonnans and Shapiro (2000). Appendix A.2 presents the proof of Proposition 3.1, and Appendix A.3 devotes to establishing the directional derivative of the Lagrangian function of problem (3.5), which is essential in the proof of Proposition 3.1. Finally, Appendix A.4 collects the proofs of Theorems 3.1–3.3 and 4.1.

A.1. Mathematical background for optimization over Banach spaces

Throughout this subsection, let $\mathcal{E}$ and $\mathcal{F}$ be two Banach spaces with dual spaces $\mathcal{E}^*$ and $\mathcal{F}^*$ respectively. Note that the dual space consists of all linear and continuous operator which map the Banach space onto the real line. By convention, for any operator $L \in \mathcal{E}^*$ (or $\mathcal{F}^*$), we use $\langle L, x\rangle$ to denote $L(x)$ for $x \in \mathcal{E}$ (or $\mathcal{F}$).

Definition A.1. A mapping $g : \mathcal{E} \mapsto \mathcal{F}$ is said to be directionally differentiable at a point $x \in \mathcal{E}$ in a direction $h \in \mathcal{E}$ if the limit

$$g'(x)[h] := \lim_{t\to 0^+} \frac{g(x+th) - g(x)}{t}$$

exists, and in this case, $g'(x)[h]$ is called the directional derivative of g at point x in direction h.[4] If g is directionally differentiable at x in every direction $h \in \mathcal{E}$, then g is said to be directionally differentiable at x.

Remark A.1. By the above definition, g is directionally differentiable at x in a direction h if and only if $g(x+th) = g(x)+tw+o^+(t)$ for $t \ge 0$, where $o^+(t)$ denotes a function such that $o^+(t)/t \to 0$ as $t \to 0^+$, and w is a vector in $\mathcal{F}$, which indeed is identified as the corresponding directional derivative.

Definition A.2. (1) A multifunction $\Psi : \mathcal{E} \to 2^{\mathcal{F}}$ is said to be convex, if its graph grh(Ψ) is a convex subset of $\mathcal{E} \times \mathcal{F}$, or equivalently $t\Psi(x_1) + (1-t)\Psi(x_2) \subset \Psi(tx_1 + (1-t)x_2)$ for any $x_1, x_2 \in \mathcal{E}$ and $t \in [0, 1]$.

(2) We say that a mapping $H_1 : \mathcal{E} \to \mathcal{F}$ is convex with respect to a convex closed set $C \subset \mathcal{F}$, if the corresponding multifunction $M_{H_1}(x) = H_1(x) + C$ is convex, where $H_1(x) + C$ denotes $\{H_1(x) + y : y \in C\}$.

Remark A.2. By the above definition, if $H_0(x) : \mathcal{E} \mapsto \mathcal{F}$ is linear then it is convex with respect to any convex subset of $\mathcal{F}$.

[4] The directional derivative can be defined in the same way for any mapping from a general normed vector space to another.

Definition A.3. The collection of all subgradients of H_0 at x, denoted by $\partial H_0(x)$, is called the subdifferential of H_0 at x, i.e.,

$$\partial H_0(x) = \{x^* \in \mathcal{E}^* : H_0(y) - H_0(x) \geq \langle x^*, y - x\rangle \text{ holds for all } y \in \mathcal{E}\}.$$

Definition A.4. The normal cone of the closed convex subset K of $\mathcal{F}$ at point y_0, denoted by $N_K(y_0)$, is defined as the set $\{\lambda \in \mathcal{F}^* : \langle \lambda, y - y_0\rangle \leq 0,\ \text{holds for all } y \in K\}$.

Now let Q and K denote, respectively, two nonempty subsets of $\mathcal{E}$ and $\mathcal{F}$, and consider the following program:

$$\min_{x\in Q} H_0(x) \quad \text{s.t. } H_1(x) \in K, \tag{A.1}$$

where H_0 and H_1 are two mappings such that $H_0 : \mathcal{E} \to \overline{\mathbb{R}}$ and $H_1 : \mathcal{E} \to \mathcal{F}$. Note that the feasible set of problem (A.1) is $\{x \in Q : H_1(x) \in K\} = Q \cap H_1^{-1}(K)$, where $H_1^{-1}(K) = \{x \in \mathcal{E} : H_1(x) \in K\}$.

Definition A.5. The program (A.1) is called convex, if it satisfies the following three conditions: (i) $H_0(x)$ is convex, (ii) $H_1(x)$ is convex with respect to the set $(-K)$, and (iii) both Q and K are convex and closed subsets.

Lemma A.1. *Assume that problem (A.1) is convex. Then one sufficient and necessary condition for a feasible point x_0 to solve problem (A.1) is as follows: there exists $\lambda \in \mathcal{F}^*$ such that*

$$0 \in \partial_x L(x_0, \lambda) + N_Q(x_0), \quad \text{and} \quad \lambda \in N_K(H_1(x_0)).$$

Here, $L(x, \lambda)$ denotes the Lagrangian function corresponding to problem (A.1), which is defined as

$$L(x, \lambda) = H_0(x) + \langle \lambda, H_1(x)\rangle, \quad (x, \lambda) \in \mathcal{E} \times \mathcal{F}^*.$$

Proof. See Bonnans and Shapiro (2000, p. 148). □

Lemma A.2. *Suppose $\mathcal{X}$ is a linear vector space, and let f be a convex functional from $\mathcal{X}$ to the extended real line $\overline{\mathbb{R}}$ taking a finite value at a point $x \in \mathcal{X}$, and let $\psi(\cdot)$ denote the directional derivative $f'(x)[\cdot]$ of f. Then $\partial f(x) = \partial\psi(0)$.*

Proof. See Proposition 2.15 of Bonnans and Shapiro (2000, p. 86). □

A.2. Proof of Proposition 3.1 (optimality conditions)

In this subsection we will provide the formal proof of Proposition 3.1. Recall that this proposition establishes the optimality conditions corresponds to the reinsurance model (3.5). We will compare the problem (3.5) with the general optimization problem (A.1) and complete the proof by applying Lemmas A.1 and A.2.

The goal function $G_\alpha(\xi, f)$ in problem (3.5) is a functional defined on the product space $H := \mathbb{R} \times \mathcal{L}^2$. It is clear that H is a Hilbert space if we equip it with the inner product $\ll \cdot, \cdot \gg$ defined by $\ll u_1, u_2 \gg = \mathsf{E}[\xi_1\xi_2 + f_1 f_2] = \xi_1\xi_2 + \mathsf{E}[f_1 f_2]$ for $u_i = (\xi_i, f_i) \in H$ and $i = 1, 2$.[5] Thus, problem (3.5) can be discussed as an optimization problem over a Hilbert space. Since Hilbert spaces are special cases of Banach spaces, any results about the optimization problem (A.1) can be applied to (3.5).

In order to apply Lemma A.1, we need to show the convexity of the optimization problem (3.5). First note that the feasible set $\mathcal{Q} \equiv \mathcal{Q}_f \bigcap \mathcal{Q}_\pi$ of the problem is clearly a closed convex subset of H. Moreover, for any $u_1 = (\xi_1, f_1)$ and $u_2 = (\xi_2, f_2)$ from $\mathcal{Q}$, and any scalar $b \in [0, 1]$,

$$\begin{aligned}
&bG_\alpha(\xi_1, f_1) + (1-b)G_\alpha(\xi_2, f_2)\\
&= b\xi_1 + b\frac{1}{\alpha}\mathsf{E}\left[\Big(X - f_1 + (1+\theta)\mathsf{E}[f_1] - \xi_1\Big)_+\right]\\
&\quad + (1-b)\xi_2 + (1-b)\frac{1}{\alpha}\mathsf{E}\left[\Big(X - f_2 + (1+\theta)\mathsf{E}[f_2] - \xi_2\Big)_+\right]\\
&\geq [b\xi_1 + (1-b)\xi_2] + \frac{1}{\alpha}\mathsf{E}\Big\{\Big[b\Big(X - f_1 + (1+\theta)\mathsf{E}[f_1] - \xi_1\Big)\\
&\quad + (1-b)\Big(X - f_2 + (1+\theta)\mathsf{E}[f_2] - \xi_2\Big)\Big]_+\Big\}\\
&= G_\alpha\Big(b\xi_1 + (1-b)\xi_2, bf_1 + (1-b)f_2\Big),
\end{aligned}$$

which implies the convexity of the functional $G_\alpha(\xi, f)$. Finally, $\mathsf{E}(f)$ is linear as a functional mapping $\mathcal{L}^2$ into $\mathbb{R}$, and hence, in view of Remark A.2, it is clearly convex with respect to the interval $[0, \pi]$. Therefore, (3.5) is a convex problem.

To proceed, it is worth emphasizing a fact about the Hilbert space H resulted from the Riesz representation theorem: For any linear mapping $M \in H^*$,[6] there exists a unique pair of elements $(r, \lambda) \in H$ such that

$$\langle M, (\xi, f)\rangle = \ll (r, \lambda), (\xi, f) \gg \equiv r\xi + \mathsf{E}[\lambda f]$$

for all $(\xi, f) \in H$. Therefore, the Lagrangian function corresponds to (3.5) has the following form:

$$L(\xi, f; r) = G_\alpha(\xi, f) + r(1+\theta)\mathsf{E}[f], \tag{A.2}$$

for $\xi \in \mathbb{R}, f \in \mathcal{L}^2$ and $r \in \mathbb{R}$.

Furthermore by setting $\mathcal{K}_\pi = [0, \pi]$ and applying Lemma A.1, we establish the following two optimality conditions for $u^* \equiv (\xi^*, f^*)$ that solves (3.5): there exists a constant $r \in \mathbb{R}$ such that

$$r \in N_{\mathcal{K}_\pi}(\mathsf{E}[f^*]) \tag{A.3}$$

and that

$$0 \in \partial_{(\xi,f)} L(\xi^*, f^*; r) + N_{\mathbb{R}\times\mathcal{Q}_f}(\xi^*, f^*). \tag{A.4}$$

Here, $\partial_{(\xi,f)} L(\xi^*, f^*; r)$ denotes the subdifferential of $L(\xi, f; r)$ at point (ξ^*, f^*), $N_{\mathcal{K}_\pi}(\mathsf{E}[f^*])$ is the normal cone to the convex set $\mathcal{K}_\pi$ at $\mathsf{E}[f^*]$, and $N_{\mathbb{R}\times\mathcal{Q}_f}(\xi^*, f^*)$ is the normal cone to $\mathbb{R} \times \mathcal{Q}_f$ at point (ξ^*, f^*).

We now complete the proof of Proposition 3.1 by first demonstrating condition (A.3) is equivalent to the proposition's condition C3 and then verifying condition (A.4) is equivalent to both conditions C1 and C2. The first equivalence is easily shown by using the definition of normal cone that (A.3) is equivalent to $r(m - \mathsf{E}[f^*]) \leq 0$ for all $m \in \mathcal{K}_\pi$, or equivalently

$$r(\mathsf{E}[f] - \mathsf{E}[f^*]) \leq 0 \quad \text{for all } f \in \mathcal{Q}_f.$$

This is exactly the condition C3.

To verify the second equivalence, it is useful to first focus on $N_{\mathbb{R}\times\mathcal{Q}_f}(\xi^*, f^*)$ and let $(\zeta, \lambda) \in N_{\mathbb{R}\times\mathcal{Q}_f}(\xi^*, f^*)$. Then by the definition of $N_{\mathbb{R}\times\mathcal{Q}_f}(\xi^*, f^*)$, we have

$$\zeta(\xi - \xi^*) \leq 0, \quad \text{and} \quad \mathsf{E}[\lambda(f - f^*)] \leq 0 \quad \text{for all } (\xi, f) \in \mathbb{R} \times \mathcal{Q}_f;$$

thus $\zeta = 0$ and (A.4) is equivalent to the condition that there exists a random variable $\lambda \in \mathcal{L}^2$ such that

$$\mathsf{E}[\lambda(f - f^*)] \leq 0 \quad \text{for all } f \in \mathcal{Q}_f, \tag{A.5}$$

and

$$(0, -\lambda) \in \partial_{(\xi,f)} L(\xi^*, f^*; r). \tag{A.6}$$

[5] The Hilbert space is a special Banach space endowed with an appropriately defined inner product.

[6] Here, H^* denotes the dual space of the Hilbert space H. The dual space consists of all the bounded linear functionals defined on the Hilbert space H.

Note that (A.5) corresponds to condition C2. The final part of the proof is to establish the equivalence between condition (A.6) and condition C1. To achieve this, we first derive the following directional derivative of the Lagrangian function $L(\cdot,\cdot;r)$:

$$\begin{aligned}\Psi(\xi,f) &:= L'(\xi^*,f^*;r)[\xi,f]\\ &= \xi+\frac{1}{\alpha}\left\{\mathsf{E}[V\mathbf{1}_{\{g^*>0\}}]+\mathsf{E}[V_+\mathbf{1}_{\{g^*=0\}}]\right\}\\ &\quad +r(1+\theta)\mathsf{E}[f],\end{aligned} \tag{A.7}$$

where g^* and V are defined, respectively, in (3.6) and (3.7). As establishing (A.7) is mathematically quite involved, we relegate the details to Appendix A.3. The above directional derivative $\Psi(\xi,f)$, together with Lemma A.2, imply that condition (A.6) is equivalent to $(0,-\lambda)\in\partial\Psi(0,0)$, which in turn is equivalent to condition C1 since $\Psi(0,0)=0$. This completes the proof of Proposition 3.1. □

A.3. Directional derivative of the Lagrangian function

Recall that we have used the notation $L(\xi,f;r):\mathbb{R}\times\mathcal{L}^2\times\mathbb{R}\mapsto\mathbb{R}$ to denote the Lagrangian function of the optimization reinsurance model (3.5). Moreover, an explicit form of $L(\xi,f;r)$ in term of $G_\alpha(\xi,f)$ is given in (A.2). The objective of this subsection is to establish the directional derivative of the functional $L(\xi,f;r)$ at (ξ^*,f^*) in the direction (ξ,f) is $\Psi(\xi,f)=L'(\xi^*,f^*;r)[\xi,f]$, as asserted in (A.7). Before presenting the proof, it is useful to first introduce the following notation

$$g(\xi,f)=X-f+(1+\theta)\mathsf{E}[f]-\xi, \tag{A.8}$$

$$h(Y)=Y_+, \tag{A.9}$$

$$e(Y)=\mathsf{E}[h(Y)], \tag{A.10}$$

so that (3.1) can equivalently be expressed as

$$\begin{aligned}G_\alpha(\xi,f) &= \xi+\frac{1}{\alpha}\mathsf{E}\left[\Big(X-f+(1+\theta)\mathsf{E}[f]-\xi\Big)_+\right]\\ &= \xi+\frac{1}{\alpha}(e\circ g)(\xi,f),\end{aligned} \tag{A.11}$$

where "$e\circ g$" denotes the composition of functions e and g.

By examining (A.2), (A.10) and (A.11), we would first focus on $e(Y)$ by deriving its directional derivative $e'(Y)[Z]$ at $Y\in\mathcal{L}^2$ in a direction $Z\in\mathcal{L}^2$. The derivation of $e'(Y)[Z]$ is outlined in Steps 1–5 below. The final Step 6 is then devoted to deriving the directional derivative $G'(\xi^*,f^*)[\xi,f]$ of $G_\alpha(\xi,f)$ and hence the required $L'(\xi^*,f^*;r)[\xi,f]$.

Step 1: Let Z be an indicator random variable such that $Z(x)=a\mathbf{1}_{[\delta_l,\delta_r)}(x)$ for all $x\in\Omega$, where a, δ_l and δ_r are nonnegative constants and $\delta_l<\delta_r$. Using Definition A.1, we have

$$\begin{aligned}h'(Y)[Z](x) &= \lim_{t\to0^+}\frac{1}{t}\left[(Y+tZ)_+-Y_+\right](x)\\ &= \begin{cases}a\cdot\mathbf{1}_{[\delta_l,\delta_r)}(x), & Y(x)\ge0,\\ 0, & \text{otherwise.}\end{cases}\\ &= Z(x)\cdot\mathbf{1}_{\{Y\ge0\}}(x).\end{aligned} \tag{A.12}$$

Step 2: Suppose Z is a nonnegative simple random variable such that $Z(x)=\sum_{i=1}^n a_iZ_i(x)$ for all $x\in\Omega$, where n is some positive integer, $\{Z_i, i=1,2,\ldots,n\}$ are indicator random variables of the form $\mathbf{1}_{[\delta_l,\delta_r)}$ with disjoint domains $[\delta_l,\delta_r)$, and $\{a_i\}_{i=1}^n$ is a sequence of positive real numbers. Then,

$$\begin{aligned}h'(Y)[Z] &= \lim_{t\to0^+}\frac{1}{t}\{(Y+tZ)_+-Y_+\}\\ &= \sum_{i=1}^n\lim_{t\to0^+}\frac{1}{t}\left\{\left[(Y+ta_iZ_i)_+-Y_+\right]\right\}\\ &= \sum_{i=1}^n a_iZ_i\cdot\mathbf{1}_{\{Y\ge0\}}\\ &= Z\cdot\mathbf{1}_{\{Y\ge0\}},\end{aligned} \tag{A.13}$$

where the second equality follows from the assumption that Z_i, $i=1,2,\ldots,n$ have disjoint domains, and the third equality is due to Step 1.

Step 3: Assume Z is a general *nonnegative* random variable from $\mathcal{L}^2$. From the definition of directional derivative, we have

$$e'(Y)[Z]=\lim_{t\to0^+}\left\{\frac{1}{t}\left[\mathsf{E}(Y+tZ)_+-\mathsf{E}(Y_+)\right]\right\}.$$

Clearly, on $\{x:Y(x)<0\}$ we have

$$\left|\frac{1}{t}\left[(Y+tZ)_+-Y_+\right]\right|=\left|\left(\frac{Y}{t}+Z\right)_+-\left(\frac{Y}{t}\right)_+\right|\le Z,$$

while on $\{x:Y(x)\ge0\}$

$$\left|\frac{1}{t}\left[(Y+tZ)_+-Y_+\right]\right|=\left|\left(\frac{Y}{t}+Z\right)-\left(\frac{Y}{t}\right)\right|=Z.$$

Combining the above results with the dominated convergence theorem results in

$$e'(Y)[Z]=\mathsf{E}\left\{\lim_{t\to0^+}\frac{1}{t}\left[(Y+tZ)_+-Y_+\right]\right\}.$$

Furthermore, it is well known that there exists a non-decreasing sequence of nonnegative simple random variable $\{Z_n, n\ge1\}$ such that $Z_n\to Z$ almost surely; thus

$$e'(Y)[Z]=\mathsf{E}\left\{\lim_{t\to0^+}\lim_{n\to\infty}\frac{1}{t}\left[(Y+tZ_n)_+-Y_+\right]\right\}.$$

To proceed, we denote $M(t,n)=\frac{1}{t}\left[(Y+tZ_n)_+-Y_+\right]$ as a function of the variables t and n. Clearly, $M(t,n)$ is non-decreasing in n for any fixed $t>0$. We now fix n and consider the monotonicity of $M(t,n)$ as a function of t. Recall that Z_n is a nonnegative random variable. Thus, on $\{x:Y(x)\ge0\}$, $M(t,n)$ is uniformly equal to Z_n for all $t\ge0$. On $\{x:Y(x)<0\}$, for $0<t\le -Y/Z_n$, $M(t,n)=0-(Y/t)_+=0$ and for $t\ge -Y/Z_n$, $M(t,n)=Y/t+Z_n$, which is monotonically decreasing to 0 as t decreases to $-Y/Z_n$. Therefore, for any sample point in Ω and any fixed n, $M(t,n)$ is decreasing as t decreases to 0. This implies that the two limits in the above expression of $e'(Y)[Z]$ are exchangeable and hence

$$\begin{aligned}e'(Y)[Z] &= \mathsf{E}\left\{\lim_{n\to\infty}\lim_{t\to0^+}\frac{1}{t}\left[(Y+tZ_n)_+-Y_+\right]\right\}\\ &= \mathsf{E}\left[\lim_{n\to\infty}Z_n\cdot\mathbf{1}_{Y\ge0}\right]=\mathsf{E}\left[Z\cdot\mathbf{1}_{Y\ge0}\right]\end{aligned} \tag{A.14}$$

where the second equality follows from Step 2.

Step 4: Now assume Z is a general *negative* random variable from $\mathcal{L}^2$. Then the procedure described in Steps 1–3 can similarly be used to establish the following directional derivative

$$e'(Y)[Z]=\mathsf{E}\left[Z\cdot\mathbf{1}_{Y>0}\right]. \tag{A.15}$$

Note the slight (but important) difference between (A.14) and (A.15).

Step 5: We now consider the directional derivative $e'(Y)[Z]$ in the direction of a general random variable $Z\in\mathcal{L}^2$. Let $N=\{x:Z(x)<0\}$, $\overline{N}=\{x:Z(x)\ge0\}$ and $Z_-=\max\{0,-Z\}$, then we have

$$\begin{aligned}e'(Y)[Z] &= \lim_{t\to0^+}\frac{1}{t}\left[\mathsf{E}\,(Y+tZ)_+-\mathsf{E}(Y_+)\right]\\ &= \lim_{t\to0^+}\frac{1}{t}\mathsf{E}\left\{\mathbf{1}_{\overline{N}}\left[(Y+tZ_+)_+-Y_+\right]\right\}\\ &\quad+\lim_{t\to0^+}\frac{1}{t}\mathsf{E}\left\{\mathbf{1}_N\left[(Y-tZ_-)_+-Y_+\right]\right\}\\ &= e'(Y)[Z_+]+e'(Y)[-Z_-]\\ &= \mathsf{E}[\mathbf{1}_{\{Y\ge0\}}Z_+]-\mathsf{E}[\mathbf{1}_{\{Y>0\}}Z_-],\end{aligned} \tag{A.16}$$

where the last equality is due to (A.14) and (A.15).

Step 6: Because of (A.2), we first focus on obtaining the derivative $G'(\xi^*, f^*)[\xi, f]$ so that $L'(\xi^*, f^*; r)[\xi, f]$ follows trivially. For brevity, we let $u^* := (\xi^*, f^*) \in \mathbb{R} \times \mathcal{L}^2$ and $u := (\xi, f) \in \mathbb{R} \times \mathcal{L}^2$. Furthermore, in what follows, all the equalities can be understood as $t \to 0^+$, if necessary.

From (A.8) and (A.10), we have

$$g(u^* + tu) = X - (f^* + tf) + (1+\theta)\mathsf{E}[f^* + tf] - (\xi^* + t\xi) = g(\xi^*, f^*) + t\Big[(1+\theta)\mathsf{E}[f] - f - \xi\Big],$$

and

$$(e \circ g)(u^* + tu) = e\Big(g^* + tV\Big),$$

where $g^* = g(\xi^*, f^*)$ and $V = (1+\theta)\mathsf{E}[f] - f - \xi$. Thus, combining (A.16) results in

$$\begin{aligned}(e \circ g)(u^* + tu) &= e(g^*) + te'(g^*)[V] + o^+(t)\\ &= e(g^*) + t\mathsf{E}[V_+\mathbf{1}_{\{g^*\geq 0\}} - V_-\mathbf{1}_{\{g_0>0\}}] + o^+(t)\\ &= e(g^*) + t\mathsf{E}[V\mathbf{1}_{\{g^*>0\}} + V_+\mathbf{1}_{\{g^*=0\}}] + o^+(t).\end{aligned}$$

It follows immediately from Definition A.1 and Remark A.1 that

$$\Big(e \circ g\Big)'(\xi^*, f^*)[(\xi, f)] = \mathsf{E}[V\mathbf{1}_{\{g^*>0\}} + V_+\mathbf{1}_{\{g^*=0\}}],$$

which, together with (A.11), leads to

$$G'_\alpha(\xi^*, f^*)[\xi, f] = \xi + \frac{1}{\alpha}\mathsf{E}\left[V\mathbf{1}_{\{g^*>0\}} + V_+\mathbf{1}_{\{g^*=0\}}\right].$$

Hence, by (A.2) we obtain the directional derivative $L'(\xi^*, f^*; r)[\xi, f]$ of the Lagrangian function L as asserted in (A.7).

A.4. Proofs of Theorems 3.1–3.3 and 4.1

The key to establishing the ceded loss function f^* is indeed an optimal solution to the reinsurance model (3.5) is to exploit Proposition 3.1. This entails two steps. The first step is to select appropriate constants $\xi^*, r \in \mathbb{R}$, and random variable $\lambda \in \mathcal{L}^2$. Based on the chosen ξ^*, r and λ, step two is to verify the three sufficient conditions C1–C3 in Proposition 3.1. If these conditions are satisfied, then f^* is an optimal solution and the proof is complete. This is the strategy we will adopt for the proofs of Theorems 3.1–3.3 and 4.1. We will provide a detailed proof of Theorem 3.1. Because of the similarity in the proofs, we will simply state the choices of ξ^*, r and λ that are essential to the proofs of the remaining theorems.

Proof of Theorem 3.1. We begin the proof by first focusing on condition C3. It is easy to see that condition C3 holds immediately by setting

$$r = \frac{1}{\alpha(1+\theta)} - 1. \tag{A.17}$$

Note that $r \geq 0$ since $\alpha(1+\theta) \leq 1$.

We now proceed to verifying condition C2. Recall that as pointed out in Remark 3.2, we have $d_\alpha < \hat{d}$. Then by choosing $\xi^* = \pi + d_\alpha$, (3.6) becomes

$$g^*(x) = \begin{cases} x - d_\alpha < 0, & x < d_\alpha,\\ x - d_\alpha \geq 0, & d_\alpha \leq x < \hat{d},\\ x - l(x) - d_\alpha > 0, & x \geq \hat{d}.\end{cases}$$

Clearly, we have $\mathbf{1}_{\{g^*<0\}} = \mathbf{1}_{\{X<d_\alpha\}}$ and $\mathbf{1}_{\{g^*=0\}} = \mathbf{1}_{\{X=d_\alpha\}}$. By defining

$$\beta_\alpha = \begin{cases} \dfrac{\alpha - \Pr\{X > d_\alpha\}}{\Pr\{X = d_\alpha\}}, & \text{if } \Pr\{X = d_\alpha\} \neq 0;\\ 0, & \text{if } \Pr\{X = d_\alpha\} = 0,\end{cases} \tag{A.18}$$

we note that $0 \leq \beta_\alpha \leq 1$, since from the definition of d_α in (3.9), we have

$$\Pr\{X > d_\alpha\} \leq \alpha, \quad \text{and} \quad \Pr\{X \geq d_\alpha\} \geq \alpha.$$

Moreover $\Pr\{X = d_\alpha\} = 0$ provided that $\Pr\{X > d_\alpha\} = \alpha$. By setting

$$\lambda = -\frac{1}{\alpha}\Big(\mathbf{1}_{\{X<d_\alpha\}} + (1-\beta_\alpha)\mathbf{1}_{\{X=d_\alpha\}}\Big). \tag{A.19}$$

Then, for any $f \in \mathcal{Q}_f$,

$$\begin{aligned}\alpha\mathsf{E}[\lambda(f - f^*)] &= -\mathsf{E}\left[\Big(\mathbf{1}_{\{X<d_\alpha\}} + (1-\beta_\alpha)\mathbf{1}_{\{X=d_\alpha\}}\Big)(f - f^*)\right]\\ &= -\mathsf{E}\left[f \cdot \Big(\mathbf{1}_{\{X<d_\alpha\}} + (1-\beta_\alpha)\mathbf{1}_{\{X=d_\alpha\}}\Big)\right]\\ &\leq 0,\end{aligned}$$

where the second equality follows from the definition that $f^*(x) = 0$ for $x \leq d_\alpha$. Hence, condition C2 is satisfied.

To demonstrate condition C1, we first combine (A.18) with (A.19) to obtain

$$\mathsf{E}[\lambda] = -\frac{1}{\alpha}\Big[\Pr\{X < d_\alpha\} + (1-\beta_\alpha)\Pr\{X = d_\alpha\}\Big] = -\frac{1-\alpha}{\alpha}.$$

This result in turn leads to

$$\mathsf{E}[1 + \alpha\lambda] = \alpha. \tag{A.20}$$

Moreover, we establish the following relation:

$$\begin{aligned}\mathsf{E}\Big[V\mathbf{1}_{\{g^*>0\}} + V_+\mathbf{1}_{\{g^*=0\}}\Big] &= \mathsf{E}\Big[V\mathbf{1}_{\{X>d_\alpha\}} + V_+\mathbf{1}_{\{X=d_\alpha\}}\Big]\\ &\geq \mathsf{E}\Big[V\mathbf{1}_{\{X>d_\alpha\}} + \beta_\alpha V_+\mathbf{1}_{\{X=d_\alpha\}}\Big]\\ &\geq \mathsf{E}\Big[V\Big(\mathbf{1}_{\{X>d_\alpha\}} + \beta_\alpha\mathbf{1}_{\{X=d_\alpha\}}\Big)\Big]\\ &= \mathsf{E}[V(1+\alpha\lambda)],\end{aligned} \tag{A.21}$$

where the last equality follows from (A.19). The above result, together with (A.17), (A.19) and (A.20), assert condition C1 as shown below:

$$\begin{aligned}A(\xi, f) &\geq \alpha\Big[\xi + r(1+\theta)\mathsf{E}[f] + \mathsf{E}(\lambda f)\Big] + \mathsf{E}[V(1+\alpha\lambda)]\\ &= \alpha\Big[\xi + r(1+\theta)\mathsf{E}[f] + \mathsf{E}(\lambda f)\Big]\\ &\quad + \mathsf{E}\Big[\Big((1+\theta)\mathsf{E}[f] - f - \xi\Big)(1+\alpha\lambda)\Big]\\ &= \xi\Big(\alpha - \mathsf{E}[1+\alpha\lambda]\Big) + \mathsf{E}[f]\Big(\alpha(1+\theta)(1+r) - 1\Big)\\ &= 0.\end{aligned}$$

Since conditions C1, C2 and C3 hold with constants $\xi^* = \pi + d_\alpha$, r as defined in (A.17), and the random variable $\lambda \in \mathcal{L}^2$ as defined in (A.19), f^* defined in (3.12) is indeed an optimal ceded loss function and hence the proof is complete. □

Proof of Theorem 3.2. We set

$$\xi^* = d^* + (1+\theta)\mathsf{E}[f^*]$$

$$\lambda = -\frac{\delta}{\alpha}\mathbf{1}_{\{X<d^*\}}$$

$$r = \frac{\delta}{\alpha}\left[\frac{1}{1+\theta} - \Pr\{X \geq d^*\}\right],$$

where $\delta = \frac{\alpha}{\Pr\{X \geq d^*\}}$ so that $0 < \delta \leq 1$. Similar to the proof of Theorem 3.1, we can verify that the ceded loss function of the form $f^* = (X - d^*)_+$ satisfies the three sufficient conditions C1, C2 and C3 in Proposition 3.1 with the above selected ξ^*, λ and r, and hence complete the proof. □

Proof of Theorem 3.3. Using the same line of arguments as in the proof of Theorem 3.1, we can show that the ceded loss function of the form $f^* = (X - d_\theta)_+$ satisfies the required conditions C1, C2 and C3 if we were to set

$$\xi^* = d_\theta + (1+\theta)\mathrm{E}[f^*]$$
$$r = 0$$
$$\lambda = -(1+\theta)\left\{\mathbf{1}_{\{X<d_\theta\}} + (1-\beta_\theta)\mathbf{1}_{\{X=d_\theta\}}\right\},$$

where

$$\beta_\theta = \begin{cases} \dfrac{1/(1+\theta) - \Pr\{X > d_\theta\}}{\Pr\{X = d_\theta\}}, & \text{if } \Pr\{X = d_\theta\} \neq 0; \\ 0, & \text{if } \Pr\{X = d_\theta\} = 0. \end{cases} \quad \text{(A.22)}$$ □

Proof of Theorem 4.1. It follows from Remark 4.1 that we only need to address the optimal solution for the reinsurance premium that falls in the range $(\pi_\theta, \pi_X]$. Moreover, it suffices to demonstrate that $f^*(x) = (x - d^*)_+$ satisfies the three optimality conditions C1, C2 and C3 in Proposition 3.1 with $\mathcal{Q}_\pi$ replaced by the binding constraint $\mathcal{Q}'_\pi$. In fact, because of the binding condition $f \in \mathcal{Q}'_\pi$, condition C3 holds trivially for any constant $r \in \mathbb{R}$. This implies that we only need to verify conditions C1 and C2 which, in turn, can be demonstrated by setting

$$\xi^* = d^* + (1+\theta)\mathrm{E}[f^*]$$
$$r = \frac{1}{\delta(1+\theta)} - 1$$
$$\lambda = -\frac{1}{\delta}\mathbf{1}_{\{X<d^*\}},$$

where $\delta = \Pr\{X \geq d^*\}$. □

References

Arrow, K.J., 1974. Optimal insurance and generalized deductibles. Scandinavian Actuarial Journal 1, 1–42.
Artzner, P., Delbaen, F., Eber, J.M., Heath, D., 1999. Coherent measures of risks. Mathematical Finance 9, 203–228.
Balbás, A., Balbás, B., Heras, A., 2009. Optimal reinsurance with general risk measures. Insurance: Mathematics and Economics 44, 374–384.
Bernard, C., Tian, W., 2009. Optimal reinsurance arrangements under tail risk measures. The Journal of Risk and Insurance 76, 709–725.
Bonnans, J.F., Shapiro, A., 2000. Perturbation Analysis of Optimization Problems. Springer Verlag.
Borch, K., 1960. The safety loading of reinsurance premiums. Skandinavisk Aktuarietidskrift 43, 163–184.
Cai, J., Tan, K.S., 2007. Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measure. ASTIN Bulletin 37, 93–112.
Cai, J., Tan, K.S., Weng, C., Zhang, Y., 2008. Optimal reinsurance under VaR and CTE risk measures. Insurance: Mathematics and Economics 43, 185–196.
Carlier, G., Dana, R.-A., 2003. Pareto efficient insurance contracts when the insurer's cost function is discontinuous. Economic Theory 21 (4), 871–893.
Carlier, G., Dana, R.-A., 2005. Rearrangement inequalities in non-convex insurance models. Journal of Mathematical Economics 41, 483–503.
Dana, R.-A., Scarsini, M., 2007. Optimal risk sharing with background risk. Journal of Economic Theory 133 (1), 152–176.
Dhaene, J., Vanduffel, S., Goovaerts, M.J., Kaas, R., Tang, Q., Vyncke, D., 2006. Risk measures and comotonicity: a review. Stochastic Models 22, 573–606.
Gajek, L., Zagrodny, D., 2000. Insurer's optimal reinsurance strategies. Insurance: Mathematics and Economics 27, 105–112.
Gajek, L., Zagrodny, D., 2004. Optimal reinsurance under general risk measures. Insurance: Mathematics and Economics 34, 227–240.
Huang, H.H., 2006. Optimal insurance contract under a Value-at-Risk constraint. The GENEVA Risk and Insurance Review 31, 91–110.
Kaas, R., Goovaerts, M., Dhaene, J., Denuit, M., 2001. Modern Actuarial Risk Theory. Kluwer Academic Publishers, Boston.
Kahn, P.M., 1961. Some remarks on a recent paper by Borch. The ASTIN Bulletin 1, 265–272.
Kaluszka, M., 2001. Optimal reinsurance under mean-variance premium principles. Insurance: Mathematics and Economics 28, 61–67.
Kaluszka, M., 2004a. An extension of Arrow's result on optimality of a stop loss contract. Insurance: Mathematics and Economics 35, 527–536.
Kaluszka, M., 2004b. Mean-variance optimal reinsurance arrangements. Scandinavian Actuarial Journal 1, 28–41.
Kaluszka, M., 2005. Optimal reinsurance under convex principles of premium calculation. Insurance: Mathematics and Economics 36, 375–398.
Landsman, Z., Valdez, E.A., 2003. Tail conditional expectations for elliptical distributions. North American Actuarial Journal 7 (4), 55–71.
McNeil, A.J., Frey, R., Embrechts, P., 2005. Quantitative Risk Management: Concepts, Techniques, and Tools. Princeton University Press.
Promislow, S.D., Young, V.R., 2005. Unifying framework for optimal insurance. Insurance: Mathematics and Economics 36, 347–364.
Rockafellar, R.T., Uryasev, S., 2002. Conditional Value-at-Risk for general loss distributions. Journal of Banking & Finance 26, 1443–1471.
Tan, K.S., Weng, C., (2010). Enhancing insurer value using reinsurance and Value-at-Risk criterion. Manuscript.
Tasche, D., 2002. Expected shortfall and beyond. Journal of Banking and Finance 26, 1519–1533.
Wang, C.P., Shyu, D., Huang, H.H., 2005. Optimal insurance design under a Value-at-Risk framework. The GENEVA Risk and Insurance Review 30, 161–179.
Weng, C., (2009). Optimal reinsurance designs: from an insurer's perspective. Ph.D. Thesis. University of Waterloo.
Wirch, J.W., Hardy, M.R., 1999. A synthesis of risk measures for capital adequacy. Insurance: Mathematics and Economics 25, 337–347.
Zhou, C., Wu, C., 2008. Optimal insurance under the insurer's risk constraint. Insurance: Mathematics and Economics 42, 992–999.
Zhou, C., Wu, W., Wu, C., 2010. Optimal insurance in the presence of insurer's loss limit. Insurance: Mathematics and Economics 46, 300–307.