



Scandinavian Actuarial Journal

ISSN: 0346-1238 (Print) 1651-2030 (Online) Journal homepage: https://www.tandfonline.com/loi/sact20

Optimal insurance in the presence of reinsurance

Sheng Chao Zhuang, Tim J. Boonen, Ken Seng Tan & Zuo Quan Xu

To cite this article: Sheng Chao Zhuang, Tim J. Boonen, Ken Seng Tan & Zuo Quan Xu (2017) Optimal insurance in the presence of reinsurance, Scandinavian Actuarial Journal, 2017:6, 535-554, DOI: 10.1080/03461238.2016.1184710

To link to this article: https://doi.org/10.1080/03461238.2016.1184710

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



0

Published online: 26 May 2016.

Submit your article to this journal 🗹

Article views: 1953



View related articles



View Crossmark data 🗹

Citing articles: 12 View citing articles 🕑



OPEN ACCESS

Check for updates

Optimal insurance in the presence of reinsurance

Sheng Chao Zhuang^a, Tim J. Boonen^b, Ken Seng Tan^a and Zuo Quan Xu^c

^aDepartment of Statistics and Actuarial Science, University of Waterloo, Waterloo, Canada; ^bAmsterdam School of Economics, University of Amsterdam, Amsterdam, The Netherlands; ^cDepartment of Applied Mathematics, Hong Kong Polytechnic University, Kowloon, Hong Kong

ABSTRACT

This paper studies an optimal insurance and reinsurance design problem among three agents: policyholder, insurer, and reinsurer. We assume that the preferences of the parties are given by distortion risk measures, which are equivalent to dual utilities. By maximizing the dual utility of the insurer and jointly solving the optimal insurance and reinsurance contracts, it is found that a layering insurance is optimal, with every layer being borne by one of the three agents. We also show that reinsurance encourages more insurance, and is welfare improving for the economy. Furthermore, it is optimal for the insurer to charge the maximum acceptable insurance premium to the policyholder. This paper also considers three other variants of the optimal insurance/reinsurance models. The first two variants impose a limit on the reinsurance premium so as to prevent insurer to reinsure all its risk. An optimal solution is still layering insurance, though the insurer will have to retain higher risk. Finally, we study the effect of competition by permitting the policyholder to insure its risk with an insurer, a reinsurer, or both. The competition from the reinsurer dampens the price at which an insurer could charge to the policyholder, although the optimal indemnities remain the same as the baseline model. The reinsurer will however not trade with the policyholder in this optimal solution.

ARTICLE HISTORY

Received 18 January 2016 Accepted 26 April 2016

KEYWORDS

Contract design; distortion risk measures; premium constraint; competition constraint; insurance; reinsurance

1. Introduction

Many papers study the bilateral contract design problem to determine optimal reinsurance contracts for a given risk of the insurer. This problem is first formally analyzed by Borch (1960) and Arrow (1963), who both show that if the reinsurance premium is calculated by the expected value principle, the stop-loss reinsurance treaty is the optimal strategy. The objective of Borch (1960) is to minimize the variance of the retained loss of the insurer, and the objective of Arrow (1963) is to maximize the expected utility of the terminal wealth of a risk-averse insurer. These pioneering results are later extended to situation where there is a more sophisticated objective function and/or more realistic premium principles (see, e.g. Young 1999, Gajek & Zagrodny 2000, 2004, Kaluszka 2001, 2005, Cai & Tan 2007, Balbás *et al.* 2009, 2015, Chi 2012, Asimit *et al.* 2013, 2015, Cai *et al.* 2013, Forthcoming, Chi & Tan 2013, Cui *et al.* 2013, Cheung *et al.* 2014, 2015, Bernard *et al.* 2015, Cheung *et al.* 2015, Boonen *et al.* 2016, Weng & Zhuang Forthcoming). In the above-mentioned papers, the risk of the insurer is typically given and the objective boils down to determining an optimal strategy of transferring part of its risk to a reinsurer. In this paper, we relax this assumption by assuming the risk of the insurer is

CONTACT Tim J. Boonen 🖾 t.j.boonen@uva.nl

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http:// creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way. endogenously determined from bargaining with a policyholder. This is a more challenging problem in that in addition to determining an optimal reinsurance, the insurer also needs to determine how much risk to underwrite from the policyholder. The problem is therefore formulated as a three-party problem involving policyholder, insurer, and reinsurer, and with the objective of jointly analyzing the optimal insurance and optimal reinsurance from the point of view of the insurer. To the best of our knowledge, this is the first paper that studies insurance and reinsurance contract design simultaneously.

In our baseline model, we assume that the insurer bargains with a policyholder to determine the coverage and premium of the risk of the policyholder. The bargaining setting is similar to Raviv (1979) but with some notable differences. We assume that both parties use dual utility as introduced by Yaari (1987) and that the insurer has access to a reinsurance market. The reinsurer can take over a part of the risk that the insurer bears, but cannot trade directly with the policyholder. In this baseline model, we formally show that optimal insurance and reinsurance contracts exist. We find that an optimal insurance contract is given by layering of the risk. Any layer is allocated to the one specific party for which the corresponding distortion function is minimal at a given quantile. This observation is in line with Pareto optimal risk sharing (Jouini *et al.* 2008, Ludkovski & Young 2009) and bilateral reinsurance contract design with a given distortion premium principle (Cui *et al.* 2013). We also establish that the exclusive accessibility to a secondary reinsurance market encourages insurance and that the insurer might profit more from selling coverage to a policyholder, and reselling a part of the risk to a reinsurance. Analytic expression for the extra profit in the presence of reinsurance is also derived.

Using the baseline model as the benchmark, we additionally consider three other variations of the optimal insurance–reinsurance models. The first two variants impose an additional (linear) constraint on how much risk an insurer could cede its risk to a reinsurer. This can be controlled by ensuring that the reinsurance premium does not exceed a certain pre-set dollar amount or a certain fractional amount of the insurance premium received. The motivation for imposing such constraint can be internal or external. In general, it is costly for an insurer to reinsure its risk and hence internally controlling the reinsurance budget relates to controlling the insurer's profitability. Externally, the regulator may impose a limit on reinsurance budget so as to prevent a situation where the insurer merely acts as an intermediary by transferring all the risk from the policyholder to reinsurer and reaping the profit margin. As we will discuss in greater details in Section 4, the presence of the reinsurance premium budget can have a non-trivial effect on the optimal risk sharing among all three stakeholders, depending on how we formulate the reinsurance premium constraint. Consistent with our intuition, the presence of the budget constraint forces insurer to retain higher risk.

A key assumption in our baseline model is that only the insurer has the exclusive access to the reinsurance market. Our third variant of the model is to relax this assumption so that the policyholder has the option of insuring its risk to an insurer, to a reinsurer, or to both. In this setting, the increase competition implies that the insurer can no longer charge the highest attainable insurance premium to the policyholder. In fact, the insurance premium charged by the insurer needs to be competitive in order to discourage policyholder from channeling its risk to the reinsurer. Our analytic results confirm this observation (see Section 5). It should be emphasized that the observation that competition reduces prices is well known in economics (see, e.g. Bertrand 1883), but we are the first one to show this analytically in the context of insurance–reinsurance design. We also formally establish that competition does not affect the optimal indemnity contracts (compared to unconstrained case), although the prices are lower and hence benefit policyholders.

The remaining of the paper is organized as follows. Section 2 introduces the model setup. Section 3 presents our proposed baseline model and derives its solutions. Section 4 discusses our first two variants of the baseline model and Section 5 analyzes the solution where policyholder has equal access to the secondary reinsurance market. Section 6 concludes the paper and the Appendix 1 collects proofs to the lemmas.

2. Model setup

Let (Ω, F, \mathbb{P}) be a probability space. Since our model involves three agents, namely the policyholder, the insurer, and the reinsurer, it is convenient to use the notations P, I, R, to represent these agents, respectively. We assume that a policyholder with an initial wealth of $W_P \in \mathbb{R}$ faces a non-negative, bounded risk X, which has a support on [0, M] for some M > 0. We denote L^{∞} as the class of bounded random variables on (Ω, F, \mathbb{P}) . A risk measure is called a distortion risk measure with distortion function $g \in \mathcal{G}$, denoted by ρ^g , if and only if it admits the following representation

$$\rho^{g}(Z) = -E^{g}[-Z] := \int_{0}^{\infty} g(S_{Z}(z)) dz + \int_{-\infty}^{0} \left[g(S_{Z}(z)) - 1 \right] dz, \tag{1}$$

for random variables $Z \in L^{\infty}$,¹ where $S_Z(z) = 1 - F_Z(z), g \in \mathcal{G}$, and²

$$\mathcal{G} := \left\{ g : [0,1] \to [0,1] \middle| g \text{ is increasing and left continuous, } g(0) = 0 \text{ and } g(1) = 1 \right\}$$

Note that the second term in (1) vanishes for non-negative risk. Distortion risk measures are based on dual utilities (Yaari 1987), and are introduced by Wang *et al.* (1997) as a premium principle. If the distortion function g is concave, the distortion risk measure is coherent, as defined by Artzner *et al.* (1999).

Two risks $Y, Z \in L^{\infty}$ are comonotonic if there exist a common random variable Q and increasing functions f_1, f_2 such that $(Y, Z) \stackrel{d}{=} (f_1(Q), f_2(Q))$. Furthermore, distortion risk measures ρ^g satisfy the following properties:

- Comonotonic additivity: $\rho^g(Y+Z) = \rho^g(Y) + \rho^g(Z)$ for two comonotonic random variables $Y, Z \in L^{\infty}$;
- Translation invariance: $\rho^g(Z + c) = \rho^g(Z) + c$ for any constant $c \in \mathbb{R}$ and random variable $Z \in L^{\infty}$.

Comonotonic additivity means that if two random variables are 'perfectly' dependent, there is no benefit from pooling them. Translation invariance states that we can interpret the risk measure in terms of a monetary amount. These two properties of distortion risk measures are well known in the literature (Wang *et al.* 1997) and results to be presented shortly rely heavily on these properties.

Based on the distortion function, we define the general distortion premium principle for the reinsurer as

$$\pi_k(Y) := (1+\theta)\rho^{g_k}(Y) = \int_0^\infty h(S_Y(z))dz, \text{ for all non-negative } Y \in L^\infty,$$
(2)

where the constant $\theta \ge 0$, $g_k \in \mathcal{G}$ and $h(s) := (1+\theta)g_k(s)$ for $s \in [0, 1]$. When $g_k(s) = s$, the distortion premium principle recovers the expected value premium principle and θ can be interpreted as the safety loading of the reinsurer. Furthermore, when the distortion function is concave with $\theta = 0$, the distortion premium principle recovers Wang's premium principle. Note that *h* is not a distortion function if $\theta > 0$ as $h(1) = 1 + \theta > 1$. It is shown by Boonen *et al.* (2015) that this premium principle can be representative for multiple reinsurers in markets where there are multiple reinsurers that all use distortion premium principles.

We assume the policyholder wishes to insure a part of its risk X with an insurer who has an initial wealth of $W_I \in \mathbb{R}$. Mathematically, an insurance indemnity f_I partitions the risk X into $f_I(X)$ and

¹Literally, distortion risk measures are given by (1) whenever the integrals converge. In this paper, we ignore this issue, and focus only on risks in L^{∞} .

²Throughout this paper, by an 'increasing' function we mean a 'non-decreasing' function, namely: g is increasing if $g(x) \ge g(y)$ whenever x > y. We say g is 'strictly increasing' if g(x) > g(y) whenever x > y. Similar conventions are used for 'decreasing' and 'strictly decreasing' functions.

538 👄 S. C. ZHUANG ET AL.

 $X - f_I(X)$, where $f_I(X)$ represents the portion of loss that is ceded to an insurer, and $X - f_I(X)$ is the residual loss retained by the policyholder. In this paper, we require that $f_I(\cdot)$ belongs to the following set $\mathcal{F}_{\hat{M}}$:

$$\mathcal{F}_{\hat{M}} := \{ f(\cdot) : f(0) = 0, \ 0 \le f(x) - f(y) \le x - y, \ \forall \ 0 \le y < x \le \hat{M} \}.$$
(3)

for $0 \le \hat{M} \le M$. That is, the loss functions $f_I(z)$ and $z - f_I(z)$ are increasing and any incremental compensation is always less than or equal to the incremental loss. These are desirable properties on the ceded loss function $f_I(z)$ as these conditions discourage moral hazard of the policyholder (Denuit and Vermandele 1998, Huberman et al. 1983). Under the insurance arrangement, the final wealth for the policyholder is $W_P - X + f_I(X) - \pi_I(f_I(X))$ and the final wealth for the insurer becomes $W_I - f_I(X) + \pi_I(f_I(X))$, where $\pi_I(f_I(X))$ represents the premium charged by the insurer for insuring the ceded risk $f_I(X)$.

As part of a prudent risk management strategy, we assume that at the time of insuring risk $f_I(X)$ from the policyholder, the insurer wishes to reinsure its risk with a reinsurer. This implies that with the reinsurance, the risk $f_I(X)$ is further partitioned into $f_R(f_I(X))$ and $f_I(X) - f_R(f_I(X))$, where $f_R(f_I(X))$ captures the stochastic loss that is ceded to a reinsurer and $f_I(X) - f_R(f(X))$ is the net residual loss retained by the insurer. The final wealth for the insurer becomes $W_I - f_I(X) + \pi_I(f_I(X)) + f_R(f_I(X)) - \pi_R(f_R(f_I(X)))$ where $\pi_R(f_R(f_I(X)))$ corresponds to the reinsurance premium defined in (2). In this paper, we require that

$$(f_I, f_R) \in \left\{ (\hat{f}_I, \hat{f}_R) : \hat{f}_I \in \mathcal{F}_M, \hat{f}_R \in \mathcal{F}_{\hat{f}_I(M)} \right\} := \mathcal{F}^2$$

We further assume that the policyholder and the insurer are endowed with *dual utility* (Yaari 1987). Maximizing dual utility of gains is mathematically equivalent to minimizing a distortion risk measure of losses. Therefore, the utilities of the policyholder and the insurer are given by

$$U_k(W) := -\rho^{g_k}(-W) := E^{g_k}[W], \qquad k \in \{P, I\}.$$

Note that we do not require concavity of the distortion function, which allows us to focus on preferences derived from value-at-risk as well.

3. The baseline insurance-reinsurance problem

In this section, we first present our proposed insurance–reinsurance model that jointly integrates policyholder's insurance decision, insurer's reinsurance decision, as well as the corresponding optimal insurance premium. More specifically, our proposed optimal insurance–reinsurance model is formally stated as follows:

$$\max_{f_{I},f_{R},\pi_{I}} E^{g_{I}} \Big[W_{I} - f_{I}(X) + \pi_{I}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \Big],$$

s.t. $(f_{I},f_{R}) \in \mathcal{F}^{2}, \pi_{I}(f_{I}(X)) \ge 0,$
 $E^{g_{P}} [W_{P} - X + f_{I}(X) - \pi_{I}(f_{I}(X))] \ge E^{g_{P}} [W_{P} - X].$ (4)

In the above model, the objective function ensures that the insurer's dual utility is maximized in the presence of both insurance and reinsurance. The left- and right-hand sides of the second constraint measure the welfare of the policyholder with and without insurance. The former is at least as great as the latter to provide an incentive for the policyholder to participate in insurance. For this reason, this constraint can be interpreted as the insurance participation constraint. The above formulation of the insurance–reinsurance model represents the baseline of our proposed model. In our subsequent analysis, we will consider three other models that are variants of the baseline model.

Proposition 3.1 below first asserts that the participation constraint leads to a closed form expression of the insurance premium while Theorem 3.1 gives the optimal $f_R^*(\cdot)$ and $f_I^*(\cdot)$ analytically.

Proposition 3.1: Every solution to (4) is such that $\pi_I(f_I(X)) = -E^{g_P}[-f_I(X)] = \rho^{g_P}(f_I(X)).$

Proof: If we fix $f_I(X)$, then the objective function in (4) is strictly increasing and continuous in $\pi_I(f_I(X))$ and the left-hand side of the last constraint is strictly decreasing and continuous in $\pi_I(f_I(X))$. This implies if a solution exists, it is such that the participation constraint is binding, i.e.

$$E^{g_P}[W_P - X + f_I(X) - \pi_I(f_I(X))] = E^{g_P}[W_P - X].$$

Since E^{g_p} satisfies the properties, translation invariance and comonotonic additivity, we have

$$\pi_I(f_I(X)) = -E^{g_P}[-f_I(X)] = \rho^{g_P}(f_I(X))$$

and this concludes the proof.

We denote $\pi_I^*(f_I(X)) = \rho^{g_P}(f_I(X))$ as the indifference pricing function, which is characterized in Proposition 3.1. We now draw some remarks on this proposition. Firstly, the insurance premium charged by the insurer is the highest acceptable premium that the policyholder is willing to pay for insurance coverage. At this level of insurance premium, the policyholder is indifferent between insurance and no insurance. Secondly, it follows from Proposition 3.1 that $\pi_I^*(f_I(X))$ is a special case of the general distortion premium principle defined in (2) with distortion function g_P and $\theta = 0$. Finally, the insurance premium charged by the insurer depends on the policyholder's distortion function. This should not be surprising since to have insurance, the policyholder needs to go through an underwriting process and hence any risk pertaining to the policyholder should be reflected in its distortion function, and hence its insurance premium.

From the definition of the distortion risk measure and for any $f_I \in \mathcal{F}_M$, we have

$$\pi_I^*(f_I(X)) = \int_0^M g_P(S_X(z)) \mathrm{d}f_I(z).$$

See Zhuang *et al.* (2016, Lemma 2.1 therein) for a detailed proof. Hence, we use a unique and nonnegative pricing kernel to price risk. Together with Proposition 3.1, this implies that our baseline model (4) can be reformulated as:

$$\max_{f_{I},f_{R}} E^{g_{I}} \Big[W_{I} - f_{I}(X) + \pi_{I}^{*}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \Big],$$

s.t. $(f_{I},f_{R}) \in \mathcal{F}^{2}.$ (5)

Note that apart from the admissibility condition on the shape of the indemnity contracts $(f_I, f_R) \in \mathcal{F}^2$, there is no other constraint. Theorem below provides the solution of our proposed baseline problem:

Theorem 3.1: The solution to (4) is given by $f_R^*(\cdot)$ and $f_I^*(\cdot)$ such that

$$(f_I^*)'(z) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } \min\{g_I(S_X(z)), h(S_X(z))\} < g_P(S_X(z)), \\ \eta(z) & \text{if } \min\{g_I(S_X(z)), h(S_X(z))\} = g_P(S_X(z)), \\ 0, & \text{if } \min\{g_I(S_X(z)), h(S_X(z))\} > g_P(S_X(z)), \end{cases}$$
(6)

540 😓 S. C. ZHUANG ET AL.

and

$$(f_{R}^{*})'(f_{I}^{*}(z)) \stackrel{a.e.}{=} \begin{cases} 1, & ifh(S_{X}(z)) < g_{I}(S_{X}(z)), \\ \kappa(f_{I}^{*}(z)), & ifh(S_{X}(z)) = g_{I}(S_{X}(z)), \\ 0, & ifh(S_{X}(z)) > g_{I}(S_{X}(z)), \end{cases}$$
(7)

where $\kappa(f_I^*(z))$ and $\eta(z)$ are any [0,1]-valued function on $\{z \in [0,M] : h(S_X(z)) = g_I(S_X(z))\}$ and $\{z \in [0,M] : \min\{g_I(S_X(z)), h(S_X(z))\} = g_P(S_X(z))\}$, respectively.

Proof: Since a distortion risk measure is translation invariant, this implies that W_I can be removed from (5). Moreover, it is easy to verify that the objective of (5) is equivalent to

$$\int_{0}^{M} g_{P}(S_{X}(z)) df_{I}(z) - \int_{0}^{M} g_{I}(S_{X}(z)) df_{I}(z) + \int_{0}^{M} g_{I}(S_{X}(z)) df_{R}(f_{I}(z)) - \int_{0}^{M} h(S_{X}(z)) df_{R}(f_{I}(z)),$$
(8)

which in turn leads to

$$\int_{0}^{M} m(z)df_{I}(z) + \int_{0}^{M} n(z)df_{R}(f_{I}(z)) = \int_{0}^{M} m(z)f_{I}'(z)dz + \int_{0}^{M} n(z)f_{R}'(f_{I}(z))f_{I}'(z)dz, \quad (9)$$

where $m(z) := g_P(S_X(z)) - g_I(S_X(z))$ and $n(z) := g_I(S_X(z)) - h(S_X(z))$. Assuming $f_I \in \mathcal{F}_M$ is given, then the optimal f_R^* is of the following form:

$$(f_R^*)'(f_I(z)) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } n(z)f_I'(z) > 0, \\ \kappa(f_I(z)), & \text{if } n(z)f_I'(z) = 0, \\ 0, & \text{if } n(z)f_I'(z) < 0. \end{cases}$$
(10)

where $\kappa(f_I(z))$ can be any [0,1]-valued function on $\{z \in [0, f_I(M)] : n(z)f'_I(z) = 0\}$. Note that f_R is defined on $[0, f_I(M)]$. By fixing f_R^* as in (10), it follows that $\int_0^M n(z)(f_R^*)'(f_I(z))f'_I(z)dz = \int_0^M n^+(z)f'_I(z)dz$, where $n^+(z) := \max\{0, n(z)\}$. Therefore, (9) reduces to

$$\int_0^M m(z) f_I'(z) dz + \int_0^M n^+(z) f_I'(z) dz = \int_0^M \left(m(z) + n^+(z) \right) f_I'(z) dz.$$

It is easy to derive that the optimal f_I^* is given by

$$(f_I^*)'(z) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } m(z) + n^+(z) > 0, \\ \eta(z), & \text{if } m(z) + n^+(z) = 0, \\ 0, & \text{if } m(z) + n^+(z) < 0, \end{cases}$$
(11)

where $\eta(z)$ can be any [0, 1]-valued function on $\{z \in [0, M] : m(z) + n^+(z) = 0\}$. Moreover, the optimal objective value of problem (5) becomes $\int_0^M (m(z) + n^+(z))^+ dz$.

Note that $\{z \in [0, M] : m(z) + n^+(z) > 0\} = \{z \in [0, M] : n(z) > 0, m(z) + n(z) > 0\} \cup \{z \in [0, M] : n(z) \le 0, m(z) > 0\} = \{z \in [0, M] : \min\{g_I(S_X(z)), h(S_X(z))\} < g_P(S_X(z))\}$. Therefore, we obtain the solutions in (6) and (7) from (10) and (11), and from substituting the definitions of the functions *m* and *n*.

Theorem 3.1 can be seen as an extension of Cui *et al.* (2013) to the case that there are three parties involved in sharing the policyholder's risk exposure *X*. With only two parties involved in risk sharing, Cui *et al.* (2013) similarly show that layering is an optimal reinsurance contract.

To continue, it is useful to define the following sets $\{K_1, K_2, K_3\}$ and $\{Q_1, Q_2, Q_3, Q_4\}$. As we will see shortly, the risk borne by the three agents can be expressed succinctly in terms of these sets:

$$K_{1} := \left\{ z \in [0, M] : g_{I}(S_{X}(z)) < \min\{g_{P}(S_{X}(z)), h(S_{X}(z))\} \right\}$$

$$K_{2} := \left\{ z \in [0, M] : g_{P}(S_{X}(z)) < \min\{h(S_{X}(z)), g_{I}(S_{X}(z))\} \right\}$$

$$K_{3} := \left\{ z \in [0, M] : h(S_{X}(z)) < \min\{g_{P}(S_{X}(z)), g_{I}(S_{X}(z))\} \right\}$$

$$Q_{1} := \left\{ z \in [0, M] : g_{I}(S_{X}(z)) = g_{P}(S_{X}(z)) = h(S_{X}(z)) \right\}$$

$$Q_{2} := \left\{ z \in [0, M] : h(S_{X}(z)) = g_{I}(S_{X}(z)) < g_{P}(S_{X}(z)) \right\}$$

$$Q_{3} := \left\{ z \in [0, M] : g_{P}(S_{X}(z)) = g_{I}(S_{X}(z)) < h(S_{X}(z)) \right\}$$

$$Q_{4} := \left\{ z \in [0, M] : g_{P}(S_{X}(z)) = h(S_{X}(z)) < g_{I}(S_{X}(z)) \right\}$$

By construction, we have $\left(\bigcup_{i=1}^{i=3} K_i\right) \cup \left(\bigcup_{j=1}^{j=4} Q_i\right) = [0, M]$, and that the sets are all disjoint from each other. It follows from Theorem 3.1 that the posterior risk of the policyholder, reinsurer, and insurer can be expressed explicitly as

$$1 - (f_I^*)'(z) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in K_2, \\ u_1(z), & \text{if } z \in Q_1, \\ u_2(z), & \text{if } z \in Q_3, \\ u_3(z), & \text{if } z \in Q_3, \\ 0, & \text{otherwise}, \end{cases}$$
(12)
$$\left(f_R^*(f_I^*(z)) \right)' \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in K_3, \\ v_1(z), & \text{if } z \in Q_1, \\ v_2(z), & \text{if } z \in Q_2, \\ v_3(z), & \text{if } z \in Q_4, \\ 0, & \text{otherwise}, \end{cases}$$
(13)
$$\left((z) - f_R^*(f_I^*(z)) \right)' \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in K_1, \\ w_1(z), & \text{if } z \in Q_1, \\ w_2(z), & \text{if } z \in Q_3, \\ w_3(z), & \text{if } z \in Q_3, \\ w_3(z), & \text{if } z \in Q_2, \\ 0, & \text{otherwise}, \end{cases}$$
(14)

where $u_1(z) + v_1(z) + w_1(z) = 1$ for $z \in Q_1$, $v_2(z) + w_3(z) = 1$ for $z \in Q_2$, $u_3(z) + w_2(z) = 1$ for $z \in Q_3$ and $u_2(z) + v_3(z) = 1$ for $z \in Q_4$. Note that the risk retained by the insurer is given by $f_I^*(X) - f_R^*(f_I^*(X))$.

 (f_I^*)

It is of interest to point out that the presence of a reinsurer leads to more insurance coverage, as can be inferred from Theorem 3.1. The profit of the insurer, as attributed to the presence of insurance and reinsurance, can be defined as

$$E^{g_{I}}\left[W_{I} - f_{I}^{*}(X) + \pi_{I}^{*}(f_{I}^{*}(X)) + f_{R}^{*}(f_{I}^{*}(X)) - \pi_{R}(f_{R}^{*}(f_{I}^{*}(X)))\right] - E^{g_{I}}\left[W_{I}\right]$$

Not surprisingly, the insurer makes more profit if there is a reinsurer, as formally asserted in the following proposition:

Proposition 3.2: The insurer makes more profit if there is a reinsurer, and the extra profit is given $by - E^{\min\{g_P,g_I\}}[-f_R^*(f_I^*(X))] - \pi_R(f_R^*(f_I^*(X))) \ge 0.$

542 😉 S. C. ZHUANG ET AL.

Proof: Note that the profit does not depend on initial wealth W_I and hence without loss of generality, we let $W_I = 0$. The profit of the insurer, in the presence of reinsurer, is given by

$$\begin{split} E^{g_{I}}\Big[-f_{I}^{*}(X) + \pi_{I}^{*}(f_{I}^{*}(X)) + f_{R}^{*}(f_{I}^{*}(X)) - \pi_{R}(f_{R}^{*}(f_{I}^{*}(X)))\Big] \\ &= E^{g_{I}}\Big[-f_{I}^{*}(X) + f_{R}^{*}(f_{I}^{*}(X))\Big] - E^{g_{P}}(-f_{I}^{*}(X)) + (1+\theta)E^{g_{R}}(-f_{R}^{*}(f_{I}^{*}(X))) \\ &= E^{\min\{g_{P},g_{I},h\}}\Big[-f_{I}^{*}(X) + f_{R}^{*}(f_{I}^{*}(X))\Big] - E^{g_{P}}[-f_{I}^{*}(X)] + (1+\theta)E^{g_{R}}(f_{R}^{*}(f_{I}^{*}(X))) \\ &= E^{\min\{g_{P},g_{I},h\}}\Big[-X + f_{R}^{*}(f_{I}^{*}(X))\Big] - E^{g_{P}}[-X] - (1+\theta)E^{g_{R}}(f_{R}^{*}(f_{I}^{*}(X))) \\ &= E^{\min\{g_{P},g_{I},h\}}\Big[-X\Big] - E^{g_{P}}[-X], \end{split}$$

where the second equality follows from $g_I = \min\{g_P, g_I, h\}$ on the domain where $(f_I^*(z) - f_R^*(f_I^*(z)))' > 0$ (see (14)), and the third equality follows from adding $E^{g_P}[-X + f_I^*(X)]$ to both sides of the equation, and the fact that $g_P = \min\{g_P, g_I, h\}$ on the domain where $1 - (f_I^*)'(z) > 0$ (see (12)).

If the reinsurer does not participate, we deduce in the same manner as in Theorem 3.1 that the optimal insurance contract, denoted by $\hat{f}_I^*(X)$, has the following form:

$$(\hat{f}_{I}^{*})'(z) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } g_{I}(S_{X}(z)) < g_{P}(S_{X}(z)), \\ \eta(z) & \text{if } g_{I}(S_{X}(z)) = g_{P}(S_{X}(z)), \\ 0, & \text{if } g_{I}(S_{X}(z)) > g_{P}(S_{X}(z)), \end{cases}$$

for all $z \in [0, M]$, where $\eta(z)$ is any [0, 1]-valued function on $\{z \in [0, M] : g_I(S_X(z)) = g_P(S_X(z))\}$. Then, we obtain directly that

$$E^{g_{I}}\left[-\hat{f}_{I}^{*}(X) + \pi_{I}^{*}(\hat{f}_{I}^{*}(X))\right] = E^{g_{I}}\left[-\hat{f}_{I}^{*}(X)\right] - E^{g_{P}}(-\hat{f}_{I}^{*}(X))$$
$$= E^{\min\{g_{P},g_{I}\}}\left[-X\right] - E^{g_{P}}[-X].$$

Hence, the extra profit is given by

$$\begin{split} E^{\min\{g_{P},g_{I},h\}}\big[-X\big] &- E^{g_{P}}[-X] - \left(E^{\min\{g_{P},g_{I}\}}\big[-X\big] - E^{g_{P}}[-X]\right) \\ &= E^{\min\{g_{P},g_{I},h\}}\big[-X\big] - E^{\min\{g_{P},g_{I}\}}\big[-X\big] \\ &= E^{\min\{g_{P},g_{I}\}}\big[-X + f_{R}^{*}(f_{I}^{*}(X))\big] - \pi_{R}(f_{R}^{*}(f_{I}^{*}(X))) \\ &- E^{\min\{g_{P},g_{I}\}}\big[-X + f_{R}^{*}(f_{I}^{*}(X))\big] - E^{\min\{g_{P},g_{I}\}}\big[-f_{R}^{*}(f_{I}^{*}(X))\big] \\ &= -E^{\min\{g_{P},g_{I}\}}\big[-f_{R}^{*}(f_{I}^{*}(X))\big] - \pi_{R}(f_{R}^{*}(f_{I}^{*}(X))). \end{split}$$

where the second equality follows from the fact that we get from (13) that $h = \min\{g_P, g_I, h\}$ on the domain where $\left(f_R^*(f_I^*(z))\right)' > 0$. This concludes the proof.

Remark 1: Suppose we introduce a reservation utility $\underline{u}_P \ge E^{g_P}[W_P - X]$. This affects the pricing function in Proposition 3.1 by a constant, i.e. if $E^{g_P}[W_P - X + f_I(X) - \pi_I(f_I(X))] \ge \underline{u}_P$, then $\pi_I(f_I(X)) = (\underline{u}_P - E^{g_P}[W_P - X]) + E^{g_P}[f_I(X)]$ for all $f_I \in \mathcal{F}_M$. We note that the indemnities in Theorem 3.1 still hold, but insuring might no longer be rational for the insurer. Hence, we need to check that the profit for the insurer is non-negative of the solution from Theorem 3.1. If it is negative, the only optimal indemnity contract is $f_I^*(z) \equiv 0$; i.e. no insurance.

4. The effect of imposing reinsurance premium budget

The baseline model that we analyzed in the previous section does not have any restriction on how much an insurer could spend on reinsurance. In practice, it may be desirable to impose a reinsurance premium budget constraint, either due to internal or external reasons. In general, it is costly for an insurer to reinsure its risk and hence internally controlling the reinsurance budget relates to controlling the insurer's profitability (Cheung & Lo Forthcoming). Externally, the reinsurance premium budget can be imposed by the regulator to prevent a situation where the insurer merely acts as an intermediary by transferring all the risk from the policyholder to reinsurer and reaping the profit margin. Motivated by these arguments, this section studies two variants of our baseline model depending on how we incorporate the reinsurance premium budget constraint. Section 4.1 imposes the premium budget in terms of dollar amount while Section 4.2 expresses the reinsurance premium budget based on a certain percentage of the insurance premium collected. In addition to deriving the analytic solutions of these models, the effects of the imposed reinsurance premium budget are also discussed.

4.1. Reinsurance premium budget in terms of dollar amount

In this subsection, we study the optimal insurance–reinsurance model by assuming that the insurer can only spend a given amount $C \ge 0$ on reinsurance. Recall that in a bilateral contract design between insurer and reinsurer, this problem is studied by Cui *et al.* (2013), and Cheung & Lo (Forthcoming).

In our setting, the constraint on the reinsurance premium can easily be incorporated by explicitly introducing an additional constraint to our proposed baseline model, as shows below:

$$\max_{\substack{f_{I},f_{R},\pi_{I} \\ \text{s.t.}}} E^{g_{I}} \begin{bmatrix} W_{I} - f_{I}(X) + \pi_{I}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \end{bmatrix}, \\ (f_{I},f_{R}) \in \mathcal{F}^{2}, \pi_{I}(f_{I}(X)) \ge 0, \\ E^{g_{P}}[W_{P} - X + f_{I}(X) - \pi_{I}(f_{I}(X))] \ge E^{g_{P}}[W_{P} - X], \\ \pi_{R}(f_{R}(f_{I}(X))) \le C.$$
(15)

As in Proposition 3.1, the insurance premium is given by $\pi_I^*(f_I(X)) = \rho^{g_P}(f_I(X))$ for all $f_I \in \mathcal{F}_M$. If there exist f_I^* and f_R^* solving problem (4) such that $\pi_R(f_R^*(f_I^*(X))) \leq C$, then f_I^* and f_R^* are also the solutions to problem (15). Therefore, we define $\bar{C} := \inf\{\pi_R(f_R^*(f_I^*(X))) : (f_I^*, f_R^*) \text{ solves (4)}\}$. Based on (13) in Theorem 3.1, it follows that $\bar{C} = \int_{K_3} h(S_X(z)) dz$. Without loss of generality, we assume that $C \leq \bar{C}$. Similarly, problem (15) can be reduced to

$$\max_{\substack{f_{I},f_{R}\\ \text{s.t.}}} \int_{0}^{M} m(z)df_{I}(z) + \int_{0}^{M} n(z)df_{R}(f_{I}(z)),$$
s.t. $(f_{I},f_{R}) \in \mathcal{F}^{2},$

$$\int_{0}^{M} h(S_{X}(z))df_{R}(f_{I}(z)) \leq C,$$
(16)

where we recall from the proof of Theorem 3.1 that $m(z) = g_P(S_X(z)) - g_I(S_X(z))$ and $n(z) = g_I(S_X(z)) - h(S_X(z))$ for $z \in [0, M]$. Note that the objective and constraint are linear in (f_I, f_R) . By introducing a Lagrangian multiplier to the constraint, we obtain the following auxiliary problem:

$$\max_{\substack{f_{I},f_{R} \\ \text{s.t.}}} \int_{0}^{M} m(z) df_{I}(z) + \int_{0}^{M} n(z) df_{R}(f_{I}(z)) - \lambda \Big(\int_{0}^{M} h(S_{X}(z)) df_{R}(f_{I}(z)) - C \Big),$$
s.t. $(f_{I},f_{R}) \in \mathcal{F}^{2}.$
(17)

Now setting $n_{\lambda}(z) := n(z) - \lambda h(S_X(z))$, then (17) reduces to the following problem

$$\max_{f_{I},f_{R}} \int_{0}^{M} m(z) df_{I}(z) + \int_{0}^{M} n_{\lambda}(z) df_{R}(f_{I}(z)),$$
s.t. $(f_{I},f_{R}) \in \mathcal{F}^{2}.$
(18)

S. C. ZHUANG ET AL. 544

For a given λ , we can apply the proof of Theorem 3.1 by changing $h(S_X(z))$ to $(1 + \lambda)h(S_X(z))$ to get all solutions $(f_I^{\lambda}, f_R^{\lambda})$ that solve (18) explicitly. Or, equivalently, $(f_I^{\lambda}, f_R^{\lambda})$ satisfies Equations (12), (13), and (14) with $h(S_X(z))$ being replaced by $(1 + \lambda)h(S_X(z))$ in K_i for i = 1, 2, 3 and Q_j for j = 1, 2, 3, 4. Then, we need to find a $\lambda^* > 0$ such that $\int_0^M h(S_X(z))df_R^{\lambda^*}(f_I^{\lambda^*}(z)) = C$ for a certain pair $(f_I^{\lambda^*}, f_R^{\lambda^*})$ which solves (18) under λ^* . For any $C \leq \overline{C}$, let us define $A_C := \{\lambda \geq 0 : \text{there exists } (f_L^{\lambda}, f_R^{\lambda}) \text{ solving } \}$ (18) under λ with $\int_0^M h(S_X(z)) df_R^{\lambda}(f_I^{\lambda}(z)) = C$ and $\lambda_C := \sup \{\lambda \ge 0 : \lambda \in A_C\}.$

Lemma 4.1: When $C \leq \overline{C}$, there exists a $\lambda^* > 0$ such that $(f_I^{\lambda^*}, f_R^{\lambda^*})$ solving problem (18) with $\int_0^M h(S_X(z)) df_R^{\lambda^*}(f_I^{\lambda^*}(z)) = C.$ For all $C \leq \overline{C}$, it holds that λ_C is decreasing in C.

Proof: The techniques in the proof for the first part of this lemma are analogous to Xu et al. (2015). We relegate the proof to the Appendix 1 of this paper. \Box

It follows from Lemma 4.1 that for any given $C \leq \overline{C}$, $f_I^{\lambda C}$ and $f_R^{\lambda C}$ are the optimal solutions to problem (16), or problem (15). Here, λ^* is interpreted as shadow prices, i.e. if we increase C by 1, the profit of the insurer increased by λ^* .

Proposition 4.1: We get for all $(f_I^*, f_R^*) \in \mathcal{F}_2$ solving (4) that there exists a pair of solution $(f_I^{\lambda_C}, f_R^{\lambda_C})$ such that

- $\begin{array}{ll} (\mathrm{i}) & f_{I}^{\lambda_{C}}(X) \leq f_{I}^{*}(X); \\ (\mathrm{ii}) & f_{I}^{\lambda_{C}}(X) f_{R}^{\lambda_{C}}(f_{I}^{\lambda_{C}}(X)) \geq f_{I}^{*}(X) f_{I}^{*}(f_{R}^{*}(X)). \end{array}$

Moreover, $f_I^{\lambda_C}(X)$ is increasing and $f_I^{\lambda_C}(X) - f_R^{\lambda_C}(f_I^{\lambda_C}(X))$ is decreasing in C.

Proof: Let the pair $(f_I^*, f_R^*) \in \mathcal{F}^2$ solve (4). Let us consider problem (15) under C_1 and C_2 with $C_1 \leq C_2 \leq \overline{C}$. Lemma 4.1 implies that λ_{C_1} and λ_{C_2} exist, and are positive with $\lambda_{C_1} \geq \lambda_{C_2}$. Then, by denoting $h^{\lambda}(z) := (1 + \lambda)h(S_X(z))$, we have $h^{\lambda C_1}(S_X(z)) \ge h^{\lambda C_2}(S_X(z))$. Therefore, it follows that $\{z \in [0, M] : g_P(S_X(z)) > \min\{g_I(S_X(z)), h^{\lambda_{C_2}}(S_X(z))\}\} \supseteq \{z \in [0, M] : g_P(S_X(z)) > 0\}$ min{ $g_I(S_X(z)), h^{\lambda C_1}(S_X(z))$ }. When we apply Theorem 3.1 with $\eta(z)$ common for both problems, we get $(f_I^{\lambda C_1})'(z) \leq (f_I^{\lambda C_2})'(z)$ for all $z \in [0, M]$. This, together with $f_I^{\lambda C_2}(0) = f_I^{\lambda C_1}(0) = 0$, leads to $f_I^{\lambda_{C_1}}(X) \le f_I^{\lambda_{C_2}}(X).$

We now verify the second inequality, i.e. $f_I^{\lambda C_1}(X) - f_R^{\lambda C_1}(f_I^{\lambda C_1}(X)) \ge f_I^{\lambda C_2}(X) - f_I^{\lambda C_2}(f_R^{\lambda C_2}(X))$. We again show this via the derivatives, i.e. we show that $(f_I^{\lambda C_1})'(z) - (f_R^{\lambda C_1}(f_I^{\lambda C_1}(z)))' \ge (f_I^{\lambda C_2})'(z) - (f_R^{\lambda C_2}(f_I^{\lambda C_2}(z)))'$. We get from Theorem 3.1 that:

$$(f_I^{\lambda_C})'(z) - (f_R^{\lambda_C}(f_I^{\lambda_C}(z)))' = \begin{cases} 1 & \text{if } z \in \{z \in [0, M] : g_I(S_X(z)) < \min\{g_P(S_X(z)), h^{\lambda_C}(S_X(z))\}\}, \\ \zeta(z) & \text{if } z \in \{z \in [0, M] : g_I(S_X(z)) = \min\{g_P(S_X(z)), h^{\lambda_C}(S_X(z))\}\}, \\ 0 & \text{otherwise,} \end{cases}$$

where we note that $h^{\lambda_C}(S_X(z))$ decreases in *C* and sets $\zeta(z)$ to be the same for both problems.

When $C_2 = \tilde{C}$, the first part of this proposition follows easily as problem (15) reduces to the baseline problem (5). This concludes the proof. \square

The effects of imposing the reinsurance premium budget are clearly highlighted in Proposition 4.1. In particular, Part (i) of the proposition implies that the insurer is less aggressive in providing insurance coverage to the policyholder since there is a limit on how much it could spend on reinsurance. The situation is aggravated by the larger risk retained by both insurer and policyholder, as asserted in Part (ii) of the proposition. These adverse effects diminish as the reinsurance premium budget increases. Note also that since $\lambda^* > 0$, there exists a $\hat{M} > 0$ such that we get for all $(f_I, f_R) \in \mathcal{F}^2$ solving problem (18) that $f_R(f_I(z)) = 0$ for all $z \in [0, M]$. Therefore, the insurer does not reinsure the realizations of the risk X that are low.

Remark 2: Suppose there is a solvency requirement which requires the insurer to control its exposure to the net insured risk (i.e. after reinsurance). Such additional requirement can be reflected in our proposed insurance–reinsurance problem (15) by explicitly imposing a solvency constraint of the form $\hat{\rho}_I(f_I(X) - f_R(f_I(X))) \leq \hat{C}$, where $\hat{\rho}_I$ is a distortion risk measure with distortion function \hat{g}_I and \hat{C} is a constant. The Lagrangian method can similarly be used to solve the resulting problem. In this case, we have the Lagrangian multiplier $\lambda^* \geq 0$ and this leads to the transformation $g_I(S_Z(z)) \rightarrow g_I(S_Z(z)) + \lambda^* \hat{g}_I(S_Z(z))$, and the problem can be solved analogously to problem (15), but with the distortion function of the insurer being increased by a factor $(1 + \lambda^*)$. The presence of the solvency constraint yields a lower retained risk for the insurer, but at the expense of higher risk exposure for both the reinsurer and the policyholder.

4.2. Reinsurance premium budget as a function of insurance premium

In this subsection, we similarly study the baseline model except imposing the reinsurance premium budget in terms of the percentage of the insurance premium collected. We use the parameter $\alpha \in (0, 1]$ to capture the percentage of the insurance premium an insurer could spend on reinsurance. In this setting, the generalized baseline model becomes

$$\max_{\substack{f_{I}, f_{R}, \pi_{I} \\ \text{s.t.}}} E^{g_{I}} \left[W_{I} - f_{I}(X) + \pi_{I}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \right],$$

s.t. $(f_{I}, f_{R}) \in \mathcal{F}^{2}, \pi_{I}(f_{I}(X)) \geq 0,$
 $E^{g_{P}}[W_{P} - X + f_{I}(X) - \pi_{I}(f_{I}(X))] \geq E^{g_{P}}[W_{P} - X],$
 $\pi_{R}(f_{R}(f_{I}(X))) \leq \alpha \pi_{I}(f_{I}(X)).$ (19)

The approach to solving the above model parallels to that in the previous subsection. We similarly have the result that $\pi_I^*(f_I(X)) = \rho^{g_P}(f_I(X))$ for all $f_I \in \mathcal{F}_M$ (cf. Proposition 3.1) and that if there exist f_I^* and f_R^* solving problem (4) such that $\pi_R(f_R^*(f_I^*(X))) \leq \alpha \pi_I^*(f_I^*(X))$, then f_I^* and f_R^* are the solutions to problem (19). Also, we assume that $\pi_R(f_R^*(f_I^*(X))) > \alpha \pi_I^*(f_I^*(X))$ for all (f_I^*, f_R^*) which solve (4). We define $\bar{\alpha} := \inf \left\{ \frac{\pi_R(f_R^*(f_I^*(X)))}{\pi_I^*(f_I^*(X))} : (f_I^*, f_R^*) \text{ solves problem (4)} \right\}$. If $\pi_I^*(f_I^*(X)) = 0$ and $\pi_R(f_R^*(f_I^*(X))) = 0$, then we let $\frac{\pi_R(f_R^*(f_I^*(X)))}{\pi_I^*(f_I^*(X))} = 0$. If $\pi_I^*(f_I^*(X)) = 0$ and $\pi_R(f_R^*(f_I^*(X))) = \infty$. Without loss of generality, we assume that $\alpha \leq \bar{\alpha}$ to exclude the trivial case. Problem (19) can be reduced to

$$\max_{\substack{f_{I},f_{R} \\ \text{s.t.}}} \int_{0}^{M} m(z)df_{I}(z) + \int_{0}^{M} n(z)df_{R}(f_{I}(z)),
\text{s.t.} \quad (f_{I},f_{R}) \in \mathcal{F}^{2},
\int_{0}^{M} h(S_{X}(z))df_{R}(f_{I}(z)) \leq \alpha \int_{0}^{M} g_{P}(S_{X}(z))df_{I}(z).$$
(20)

Similarly, applying a Lagrangian multiplier γ to the constraint leads to the following auxiliary problem:

$$\max_{\substack{f_{I},f_{R} \\ \text{s.t.}}} \int_{0}^{M} m(z) df_{I}(z) + \int_{0}^{M} n(z) df_{R}(f_{I}(z)) - \gamma \left(\int_{0}^{M} h(S_{X}(z)) df_{R}(f_{I}(z)) - \alpha \int_{0}^{M} g_{P}(S_{X}(z)) df_{I}(z) \right),$$
s.t. $(f_{I},f_{R}) \in \mathcal{F}^{2}.$
(21)

Here, we assume $\gamma > 0$ since if $\gamma = 0$, then a solution of the unconstrained problem in Section 3 (i.e. our baseline model) appears to satisfy the constraint. By denoting $m_{\gamma}(z) := m(z) + \gamma \alpha g_P(S_X(z))$ and $n_{\gamma}(z) = n(z) - \gamma h(S_X(z))$, problem (21) reduces to the following problem

$$\max_{\substack{f_I,f_R\\ \text{s.t.}}} \int_0^M m_\gamma(z) df_I(z) + \int_0^M n_\gamma(z) df_R(f_I(z)),$$

s.t. $(f_I,f_R) \in \mathcal{F}^2.$ (22)

546 😓 S. C. ZHUANG ET AL.

Using Theorem 3.1, we can obtain f_I^{γ} and f_R^{γ} which solve (22) explicitly. Then, we need to find γ^* such that $\int_0^M h(S_X(z)) df_R^{\gamma^*}(f_I^{\gamma^*}(z)) = \alpha \int_0^M g_P(S_X(z)) df_I^{\gamma^*}(z)$. Analogous to Lemma 4.1, we have the following lemma and its proof is relegated to the Appendix 1.

Lemma 4.2: When $\alpha \leq \bar{\alpha}$, there exists a $\gamma^* > 0$ such that $(f_I^{\gamma^*}, f_R^{\gamma^*})$ solving (21) with

$$\int_0^M h(S_X(z)) df_R^{\gamma^*}(f_I^{\gamma^*}(z)) = \alpha \int_0^M g_P(S_X(z)) df_I^{\gamma^*}(z).$$

For $\alpha \leq \bar{\alpha}$, it follows from the above lemma that $f_I^{\gamma^*}$ and $f_R^{\gamma^*}$ are the optimal solutions to problem (20), or problem (19).

Proposition 4.2: In the constrained problem (19), we get for all $(f_I^*, f_R^*) \in \mathcal{F}_2$ solving (4) that there exists a pair of solution $(f_I^{\gamma^*}, f_R^{\gamma^*})$ solving (19) such that

(i) $f_R^{\gamma^*}(f_I^{\gamma^*}(X)) \leq f_R^*(f_I^*(X));$ (ii) $f_I^{\gamma^*}(X) - f_R^{\gamma^*}(f_I^{\gamma^*}(X)) \geq f_I^*(X) - f_R^*(f_I^*(X)).$

Proof: Let the pair $(f_I^*, f_R^*) \in \mathcal{F}_2$ solve (4). Lemma 4.2 asserts that γ^* exists and is positive. By denoting $h^{\gamma^*}(z) := (1 + \gamma^*)h(S_X(z))$ and $g_P^{\gamma^*,\alpha}(z) := (1 + \alpha\gamma^*)g_P(S_X(z))$, then it follows that $\{z \in [0, M] : h(S_X(z)) < \min\{g_I(S_X(z)), g_P(S_X(z))\}\} \supseteq \{z \in [0, M] : h^{\gamma^*}(z) < \min\{g_I(S_X(z)), g_P^{\gamma^*,\alpha}(z)\}\}$. When we apply Theorem 3.1 and using the same $\kappa(z)$ for both problems, we get $(f_R^{\gamma^*}(f_I^{\gamma^*}(z)))' \leq (f_R^*(f_I^*(z)))'$ for all $z \in [0, M]$. This, and together with $f_R^{\gamma^*}(f_I^{\gamma^*}(0)) = f_R^*(f_I^*(0)) = 0$, leads to $f_R^{\gamma^*}(f_I^{\gamma^*}(X)) \leq f_R^*(f_I^*(X))$. This holds for all $(f_I^*, f_R^*) \in \mathcal{F}_2$ solving (4) due to $\{z \in [0, M] : h(S_X(z)) = \min\{g_I(S_X(z)), g_P(S_X(z))\}\} \subseteq \{z \in [0, M] : h^{\gamma^*}(z) > \min\{g_I(S_X(z)), g_P^{\gamma^*,\alpha}(z)\}\}$.

Let us continue with the second result, where we again let the pair $(f_I^*, f_R^*) \in \mathcal{F}_2$ solve (4). We show this via the derivatives, i.e. we show that $(f_I^{\gamma^*}(z))' - (f_R^{\gamma^*}(f_I^{\gamma^*}(z)))' \ge (f_I^*)'(z) - (f_R^*(f_I^*(z)))'$. We get from Theorem 3.1 that:

$$\left(f_{I}^{\gamma^{*}}\right)'(z) - \left(f_{R}^{\gamma^{*}}\left(f_{I}^{\gamma^{*}}(z)\right)\right)' = \begin{cases} 1 & \text{if } z \in \{z \in [0, M] : g_{I}(S_{X}(z)) < \min\{h^{\gamma^{*}}(z), g_{P}^{\gamma^{*}, \alpha}(z)\}, \\ \zeta(z) & \text{if } z \in \{z \in [0, M] : g_{I}(S_{X}(z)) = \min\{h^{\gamma^{*}}(z), g_{P}^{\gamma^{*}, \alpha}(z)\}, \\ 0 & \text{otherwise.} \end{cases}$$

Thus, it is sufficient to show that $h^{\gamma^*}(z)$ and $g_P^{\gamma^*,\alpha}(z)$ increase in γ^* , which are true by the definitions of $h_{\gamma^*}(z)$ and $g_P^{\gamma^*,\alpha}(z)$. This holds for all $(f_I^*, f_R^*) \in \mathcal{F}_2$ solving (4) due to $\{z \in [0, M] : g_I(S_X(z)) = \min\{h(z), g_P(z)\}\} \subseteq \{z \in [0, M] : g_I(S_X(z)) < \min\{h^{\gamma^*}(z), g_P^{\gamma^*,\alpha}(z)\}\}$. This concludes the proof. \Box

Recall that the generalized baseline models (15) and (19) are similar in that both impose an additional constraint on the reinsurance premium budget. They differ from each other merely on how we explicitly construct the budget constraint. As a result, there are some similarities as well as differences on the effect of the constraint on the resulting models. In particular, the presence of the reinsurance premium budget constraint implies that the risk retained by the insurer is higher for problem (19), as in problem (15). However, as opposed to problem (15), the risk that is transferred to a reinsurer decreases and that the insurance coverage $f_I(X)$ is not necessarily increasing or decreasing for problem (19).

5. The presence of competition from the reinsurer

The optimal insurance–reinsurance models that we have considered so far assume that reinsurer can only trade with the insurer. What if we relax this assumption and permit direct bargaining between the policyholder and the reinsurer? In other words, the policyholder has the option of insuring its

risk exclusively with the insurer, or with the reinsurer, or with both. Under this setting, what are the implications on its insurance decision as well as its insurance premium?

To address these issues, we modify our baseline model by assuming that a part of the policyholder's risk could be traded directly to a reinsurer. Therefore, we will determine a pricing function $\pi_I : L^{\infty} \rightarrow \mathbb{R}$ instead of a single price of the optimal contract. Then, a plausible generalization of the baseline model can be formulated as follows:

$$\max_{f_{I},f_{R},\pi_{I}} E^{g_{I}} \Big[W_{I} - f_{I}(X) + \pi_{I}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \Big],$$
s.t.
$$(f_{I},f_{R}) \in \mathcal{F}^{2}, \pi_{I}(\hat{f}_{I}(X)) \ge 0, \forall \hat{f}_{I} \in \mathcal{F}_{M},$$

$$E^{g_{P}} [W_{P} - X + \hat{f}_{I}(X) - \pi_{I}(\hat{f}_{I}(X))] \ge E^{g_{P}} [W_{P} - X], \forall \hat{f}_{I} \in \mathcal{F}_{M},$$

$$\pi_{I}(\hat{f}_{I}(X)) \le \pi_{I}(\hat{f}_{I}(X) - \hat{f}_{R}(\hat{f}_{I}(X))) + \pi_{R}(\hat{f}_{R}(\hat{f}_{I}(X))), \forall (\hat{f}_{I}, \hat{f}_{R}) \in \mathcal{F}^{2}.$$
(23)

It is important to note that the individual rationality constraints are adjusted such that it holds for every insurance contract. The term $\pi_I(\hat{f}_I(X))$ on the left-hand side of the third constraint represents the cost to the policyholder for the insurer to provide the insurance coverage $\hat{f}_I(X)$. This assumes that the entire risk $\hat{f}_I(X)$ is ceded exclusively to the insurer. Suppose now the same amount of risk $\hat{f}_I(X)$ is partitioned into two portions, i.e. $\hat{f}_I(X) - \hat{f}_R(\hat{f}_I(X))$ and $\hat{f}_R(\hat{f}_I(X))$, and that these two parts of the risk are insured with the insurer and reinsurer, respectively. The right-hand side of the third constraint therefore denotes the total cost to the policyholder if it were to insure its risk from both insurer and reinsurer. The presence of this constraint ensures that it is economically cheaper for the policyholder to simply deal exclusively with the insurer, even though the policyholder could also insure directly with the reinsurer. For this reason, we refer this constraint as the competition constraint. Because of this constraint, Proposition 3.1 does not need to hold, i.e. competition from the reinsurer implies that the insurer may no longer able to charge the highest acceptable insurance premium and hence the participation constraint may be slack.

We assume that π_I is a linear pricing function. Then, the last constraint can be rewritten as $\pi_I(\hat{f}_R(\hat{f}_I(X))) \leq \pi_R(\hat{f}_R(\hat{f}_I(X))), \forall (\hat{f}_I, \hat{f}_R) \in \mathcal{F}^2$. Since $\{\hat{f}_R(\hat{f}_I(X)) : (\hat{f}_I, \hat{f}_R) \in \mathcal{F}^2\} = \{\tilde{f}_I(X) : \tilde{f}_I \in \mathcal{F}_M\}$, and that distortion risk measures are comonotonic additive and translation invariant, problem (23) can be reduced to

$$\max_{f_{I},f_{R},\pi_{I}} E^{g_{I}} \Big[W_{I} - f_{I}(X) + \pi_{I}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \Big],$$
s.t.
$$(f_{I},f_{R}) \in \mathcal{F}^{2}, \pi_{I}(f_{I}(X)) \geq 0, \qquad (24)$$

$$\pi_{I}(\hat{f}_{I}(X)) \leq -E^{g_{P}}[-\hat{f}_{I}(X)], \forall \hat{f}_{I} \in \mathcal{F}_{M}, \\
\pi_{I}(\hat{f}_{I}(X)) \leq \pi_{R}(\hat{f}_{I}(X)), \forall \hat{f}_{I} \in \mathcal{F}_{M}.$$

By defining the function $g_A(s) := \min\{g_P(s), h(s)\}$ for all $s \in [0, 1]$, which is a distortion function, we show that $\pi_I(\hat{f}_I(X)) = -E^{g_A}(-\hat{f}_I(X))$ for all $\hat{f}_I \in \mathcal{F}_M$, as asserted in the proposition below:

Proposition 5.1: For every solution to (24), it holds that $\pi_I(\hat{f}_I(X)) = -E^{g_A}[-\hat{f}_I(X)]$ for all $\hat{f}_I \in \mathcal{F}_M$. **Proof:** Note that the constraint in problem (24) can be written as $\pi_I(\hat{f}_I(X)) \leq \min\{-E^{g_P}(-\hat{f}_I(X)), \pi_R(\hat{f}_I(X))\}$ for all $\hat{f}_I \in \mathcal{F}_M$. Since the utility of the insurer is strictly increasing in $\pi_I(\hat{f}_I(X))$, we get that π_I is the largest linear function satisfying this constraint, i.e. there does not exist a linear pricing function $\tilde{\pi}_I$ satisfying the constraint such that there exist an $\tilde{f} \in \mathcal{F}_M$ with $\pi_I(\tilde{f}(X)) < \tilde{\pi}_I(\tilde{f}(X))$.

From $g_A(S_X(z)) \leq g_P(S_X(z))$ and $g_A(S_X(z)) \leq h(S_X(z))$ for every $z \in [0, M]$, we get $-E^{g_A}[-\hat{f}_I(X)] \leq \min\{-E^{g_P}(-\hat{f}_I(X)), \pi_R(\hat{f}_I(X))\}$ for all $\hat{f}_I \in \mathcal{F}_M$. Moreover, it follows from Boonen *et al.* (2015, Theorem 3.2) that $-E^{g_A}[-\hat{f}_I(X)] = \min\{-E^{g_P}[-\hat{f}_I(X) + \tilde{f}(\hat{f}_I(X))] + \pi_R(\tilde{f}(\hat{f}_I(X))) : \tilde{f} \in \mathcal{F}_{\hat{f}_I(M)}\}$. Suppose that there exists an $\hat{f}_I \in \mathcal{F}_M$ such that $\pi_I(\hat{f}_I(X)) > -E^{g_A}[-\hat{f}_I(X)]$. Then, there exists an $\tilde{f} \in \mathcal{F}_{\hat{f}_I(M)}$ such that $\pi_I(\hat{f}_I(X)) > -E^{g_P}[-\hat{f}_I(X) + \tilde{f}(\hat{f}_I(X))] + \pi_R(\tilde{f}(\hat{f}_I(X)))$. Since π_I is a linear function, we have $\pi_I(\hat{f}_I(X)) = \pi_I(\hat{f}_I(X) - \tilde{f}(\hat{f}_I(X))) + \pi_I(\tilde{f}(\hat{f}_I(X)))$. So, we must have

548 👄 S. C. ZHUANG ET AL.

 $\pi_{I}(\hat{f}_{I}(X) - \tilde{f}(\hat{f}_{I}(X))) > -E^{g_{P}}[-\hat{f}_{I}(X) + \tilde{f}(\hat{f}_{I}(X))] \text{ or } \pi_{I}(\tilde{f}(\hat{f}_{I}(X))) > \pi_{R}(\tilde{f}(\hat{f}_{I}(X))). \text{ This is a violation}$ of the feasibility constraints. Hence, the feasible linear pricing function $\pi_{I}(\hat{f}_{I}(X)) = -E^{g_{A}}[-\hat{f}_{I}(X)]$ is the largest. This concludes the proof. \Box

By denoting $\hat{\pi}_I(\hat{f}_I(X)) := -E^{g_A}[-\hat{f}_I(X)]$ and together with Proposition 5.1, problem (24) simplifies to

$$\max_{\substack{f_{I}, f_{R}, \pi_{I} \\ \text{s.t.}}} E^{g_{I}} \Big[W_{I} - f_{I}(X) + \hat{\pi}_{I}(f_{I}(X)) + f_{R}(f_{I}(X)) - \pi_{R}(f_{R}(f_{I}(X))) \Big],$$

$$(25)$$

It is easy to show that the objective of (25) is equivalent to

$$\int_{0}^{M} g_{A}(S_{X}(z)) df_{I}(z) - \int_{0}^{M} g_{I}(S_{X}(z)) df_{I}(z) + \int_{0}^{M} g_{I}(S_{X}(z)) df_{R}(f_{I}(z)) - \int_{0}^{M} h(S_{X}(z)) df_{R}(f_{I}(z)).$$
(26)

Theorem 5.1: Every solution pair $(f_I^*, f_R^*) \in \mathcal{F}^2$ solving (4) also solves problem (23). Within the set of solutions to (23), every indemnity profile $(f_I^*, f_R^*) \in \mathcal{F}^2$ solving (4) yields the highest utility of the policyholder.

Proof: We define $m_1(z) := g_A(S_X(z)) - g_I(S_X(z))$, and recall that $n(z) = g_I(S_X(z)) - h(S_X(z))$. It follows that $m_1(z) + n(z) = g_A(S_X(z)) - h(S_X(z)) \le 0$. Using Theorem 3.1 and defining $A_1 := \{z \in [0, M] : m_1(z) = n(z) = 0\}$, $B_1 := \{z \in [0, M] : n(z) > 0, m_1(z) + n(z) = 0\}$, and $C_1 := \{z \in [0, M] : n(z) < 0, m_1(z) = 0\}$, we have:

$$1 - (f_I^*)'(z) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in \{z \in [0, M] : m_1(z) < 0, m_1(z) + n(z) < 0\}, \\ u_{11}(z), & \text{if } z \in A_1, \\ u_{21}(z), & \text{if } z \in B_1, \\ u_{31}(z), & \text{if } z \in C_1, \\ 0, & \text{otherwise}, \end{cases}$$

$$\left(f_R^*(f_I^*(z))\right)' \stackrel{a.e.}{=} \begin{cases} v_{11}(z), & \text{if } z \in A_1, \\ v_{31}(z), & \text{if } z \in B_1, \\ 0, & \text{otherwise}, \end{cases}$$

$$\left(f_I^*(z) - f_R^*(f_I^*(z))\right)' \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in \{z \in [0, M] : m_1(z) > 0, n(z) < 0\}, \\ w_{11}(z), & \text{if } z \in A_1, \\ v_{21}(z), & \text{if } z \in A_1, \\ 0, & \text{otherwise}, \end{cases}$$

$$\left(f_I^*(z) - f_R^*(f_I^*(z))\right)' \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in A_1, \\ w_{11}(z), & \text{if } z \in A_1, \\ w_{21}(z), & \text{if } z \in C_1, \\ 0, & \text{otherwise}. \end{cases}$$

$$(29)$$

where $u_{11}(z) + v_{11}(z) + w_{11}(z) = 1$ for $z \in A_1, u_{21}(z) + v_{31}(z) = 1$ for $z \in B_1$, and $u_{31}(z) + w_{21}(z) = 1$ for $z \in C_1$.

Then, according to the same analysis after Theorem 3.1, we obtain the following expressions of the risks as borne by the three agents after the transactions in terms of K_i for i = 1, 2, 3 and Q_j for j = 1, 2, 3, 4 as

$$1 - (f_I^*)'(z) \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in K_2, \\ u_{11}(z), & \text{if } z \in Q_1 \cup Q_2, \\ u_{21}(z), & \text{if } z \in Q_4 \cup K_3, \\ u_{31}(z), & \text{if } z \in Q_3, \\ 0, & \text{otherwise}, \end{cases}$$
(30)
$$\left(f_R^*(f_I^*(z)) \right)' \stackrel{a.e.}{=} \begin{cases} v_{11}(z), & \text{if } z \in Q_1 \cup Q_2, \\ v_{31}(z), & \text{if } z \in Q_4 \cup K_3, \\ 0, & \text{otherwise}, \end{cases}$$
(31)
$$\left(f_I^*(z) - f_R^*(f_I^*(z)) \right)' \stackrel{a.e.}{=} \begin{cases} 1, & \text{if } z \in K_1, \\ w_{11}(z), & \text{if } z \in Q_1 \cup Q_2, \\ w_{21}(z), & \text{if } z \in Q_3, \\ 0, & \text{otherwise}, \end{cases}$$
(32)

where $u_{11}(z) + v_{11}(z) + w_{11}(z) = 1$ for $z \in Q_1 \cup Q_2$, $u_{21}(z) + v_{31}(z) = 1$ for $z \in Q_4 \cup K_3$, and $u_{31}(z) + w_{21}(z) = 1$ for $z \in Q_3$. The above contracts recover solutions solving the baseline model by setting

$$\begin{split} u_{11}(z) &\stackrel{a.e.}{=} \begin{cases} u_1(z), & \text{ if } z \in Q_1, \\ 0, & \text{ if } z \in Q_2, \end{cases} \\ u_{21}(z) &\stackrel{a.e.}{=} \begin{cases} u_2(z), & \text{ if } z \in Q_4, \\ 0, & \text{ if } z \in K_3, \end{cases} \\ u_{31}(z) \stackrel{a.e.}{=} \begin{cases} u_3(z), & \text{ if } z \in Q_3, \\ v_{11}(z) \stackrel{a.e.}{=} \end{cases} \begin{cases} v_1(z), & \text{ if } z \in Q_1, \\ v_2(z), & \text{ if } z \in Q_2, \end{cases} \\ v_{31}(z) \stackrel{a.e.}{=} \begin{cases} v_3(z), & \text{ if } z \in Q_4, \\ 1, & \text{ if } z \in K_3, \end{cases} \\ u_{11}(z) \stackrel{a.e.}{=} \begin{cases} w_1(z), & \text{ if } z \in Q_1, \\ w_3(z), & \text{ if } z \in Q_1, \\ w_3(z), & \text{ if } z \in Q_1, \end{cases} \\ w_{21}(z) \stackrel{a.e.}{=} \begin{cases} w_2(z), & \text{ if } z \in Q_2, \end{cases} \end{split}$$

Hence, all contracts (f_I^*, f_R^*) solving the baseline problem (4) are contracts that solve the problem (23) as well.

Next, we show the second result. This result states that (f_I^*, f_R^*) solving (4) also solves the following optimization problem:

$$\max_{\substack{f_{I},f_{R} \\ \text{s.t.}}} E^{g_{P}} \Big[W_{P} - X + f_{I}(X) - \hat{\pi}_{I}(f_{I}(X)) \Big],$$
s.t. $(f_{I},f_{R}) \text{ solves (23).}$
(33)

550 🔄 S. C. ZHUANG ET AL.

where $\hat{\pi}_I(f_I(X)) = -E^{g_A}[-f_I(X)]$. Since E^g is comonotonic additive and translation invariant, the objective of problem (33) is, modulo a constant, given by

$$\begin{split} &\int_{0}^{M} \left(g_{P}(S_{X}(z)) - g_{A}(S_{X}(z)) \right) df_{I}(z) \\ &= \int_{g_{P}(S_{X}(z)) \leq h(S_{X}(z))} \left(g_{P}(S_{X}(z)) - g_{P}(S_{X}(z)) \right) df_{I}(z) \\ &+ \int_{g_{P}(S_{X}(z)) > h(S_{X}(z))} \left(g_{P}(S_{X}(z)) - h(S_{X}(z)) \right) df_{I}(z) \\ &= \int_{g_{P}(S_{X}(z)) > h(S_{X}(z))} \left(g_{P}(S_{X}(z)) - h(S_{X}(z)) \right) df_{I}(z). \end{split}$$

Therefore, it yields that for the solution of problem (23) in (30), it is optimal for the policyholder to set $u_{11}(z) = 0$ for $z \in Q_2$, and $u_{21}(z) = 0$ for $z \in K_3$. This implies that every solution $(f_I^*, f_R^*) \in \mathcal{F}^2$ solving (4) also solves (33). The proof is thus complete.

From Theorem 5.1, we get that from the set of solutions to (23), the unconstrained indemnity profile is the only one that is not Pareto dominated by another solution of (23). The competition constraint does not affect the indemnity functions, which remain the ones that are improving total welfare. It only lowers the insurance price, and therefore it affects the way the profits are shared between the policyholder and the insurer. The policyholder pays a smaller price for the same coverage.

To conclude, we can straightforwardly see that if we introduce linear constraint on the reinsurance premium to problem (23), an optimal pair of indemnity functions is the same as in Section 4, but the insurance premium principle is as in Proposition 5.1. Therefore, introducing competition affects prices, whereas linear constraints on the reinsurance premium affect the indemnity functions.

6. Conclusion

This paper is the first one to study insurance and reinsurance contract design in a single model. We assume that the preferences of the parties are given by distortion risk measures, which are equivalent to dual utilities. We show the existence of optimal insurance and reinsurance contracts simultaneously. We find that layering of the insurance risk is optimal, so are under various possible constrained models. We give explicit expressions of the optimal risk sharing contracts and the insurance pricing function. We show that classical risk sharing models do apply to the setting with insurance and reinsurance jointly as well.

Moreover, we show the effect of a possible competition by permitting the reinsurer to join the insurance market. In this setting, the policyholder can trade directly with the reinsurer. We show that this competition between the insurer and reinsurer reduces prices, but that optimal indemnity contracts remain the same as for the case without competition. This result shows the relevance of competition in the insurance market. Hence, adding linear constraints on the reinsurance premium affects the indemnity functions and not the insurance premium function. Moreover, introducing a competition constraint does influence the insurance premium function, but not the indemnity functions.

Acknowledgements

The authors are grateful for the constructive comments from the anonymous reviewer. All four authors acknowledge the financial support from the Society of Actuaries Centers of Actuarial Excellence Research Grant. In addition, Zhuang acknowledges financial support from the Department of Statistics and Actuarial Science, University of Waterloo. Tan thanks financial support from the Natural Sciences and Engineering Research Council of Canada (PIN: 220010). Xu acknowledges financial supports from the Hong Kong Early Career Scheme (No.533112), the Hong Kong General Research Fund (No.529711), NNSF of China (No.11471276), and the Hong Kong Polytechnic University.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was financially supported by the Society of Actuaries Centers of Actuarial Excellence Research Grant; Department of Statistics and Actuarial Science, University of Waterloo; Natural Sciences and Engineering Research Council of Canada [PIN: 220010]; Hong Kong Early Career Scheme [number 533112]; the Hong Kong General Research Fund [number 529711]; NNSF of China [number 11471276]; and the Hong Kong Polytechnic University.

References

- Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review* **53**, 941–973. Artzner, P., Delbaen, F., Eber, J. & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance* **9**, 203–228.
- Assa, H. (2015). On optimal reinsurance policy with distortion risk measures and premiums. *Insurance: Mathematics and Economics* **61**, 70–75.
- Asimit, A. V., Badescu, A. M. & Cheung, K. C. (2013). Optimal reinsurance in the presence of counterparty default risk. *Insurance: Mathematics and Economics* **53**, 690–697.
- Asimit, A. V., Chi, Y. & Hu, J. (2015). Optimal non-life reinsurance under Solvency II Regime. Insurance: Mathematics and Economics 65, 227–237.
- Balbás, A., Balbás, B. & Heras, A. (2009). Optimal reinsurance with general risk measures. *Insurance: Mathematics and Economics* 44, 374–384.
- Balbás, A., Balbás, B., Balbás, R. & Heras, A. (2015). Optimal reinsurance under risk and uncertainty. *Insurance: Mathematics and Economics* **60**, 61–74.
- Bernard, C., He, X., Yan, J.-A. & Zhou, X. Y. (2015). Optimal insurance design under rank-dependent expected utility. *Mathematical Finance* 25, 154–186.
- Bertrand, J. (1883). Review of 'Théorie mathématique de la richesse sociale' and 'Recherches sur les principles mathématiques de la théorie des richesses'. *Journal des Savants* 48, 499–508.
- Boonen, T. J., Tan, K. S. & Zhuang, S. C. (2015). Optimal reinsurance with one insurer and multiple reinsurers. Working Paper. Available at SSRN: http://ssrn.com/abstract=2628950
- Boonen, T. J., Tan, K. S. & Zhuang, S. C., (2016). Pricing in reinsurance bargaining with comonotonic additive utility functions. *Astin Bulletin* **46**, 507–530.
- Borch, K., (1960). An attempt to determine the optimum amount of stop loss reinsurance. *Transactions of the 16th International Congress of Actuaries* 1, 597–610.
- Cai, J., Fang, Y., Li, Z. & Willmot, G. E. (2013). Optimal reciprocal reinsurance treaties under the joint survival probability and the joint profitable probability. *Journal of Risk and Insurance* **80**(1), 145–168.
- Cai, J., Lemieux, C. & Liu, F., (Forthcoming). Optimal reinsurance from the perspectives of both an insurer and a reinsurer. *ASTIN Bulletin*.
- Cai, J. & Tan, K. S. (2007). Optimal retention for a stop-loss reinsurance under the VaR and CTE risk measures. *ASTIN Bulletin* **37**, 93–112.
- Cheung, K. C., Chong, W. F. & Yam, S. C. P. (2015). The optimal insurance under disappointment theories. *Insurance: Mathematics and Economics* **64**, 77–90.
- Cheung, K. C. & Lo, A., (Forthcoming). Characterizations of optimal reinsurance treaties: a cost-benefit approach. *Scandinavian Actuarial Journal.*
- Cheung, K. C., Sung, K. C. J., Yam, S. C. P. & Yung, S. P. (2014). Optimal reinsurance under general law-invariant risk measures. *Scandinavian Actuarial Journal* 1, 72–91.
- Chi, Y. (2012). Reinsurance arrangements minimizing the risk-adjusted value of an insurer's liability. *ASTIN Bulletin* **42**, 529–557.
- Chi, Y. & Tan, K. S. (2013). Optimal reinsurance with general premium principles. *Insurance: Mathematics and Economics* **52**, 180–189.
- Cui, W., Yang, J. & Wu, L. (2013). Optimal reinsurance minimizing the distortion risk measure under general reinsurance premium principles. *Insurance: Mathematics and Economics* **53**, 74–85.
- Denuit, M. & Vermandele, C. (1998). Optimal reinsurance and stop-loss order. *Insurance: Mathematics and Economics* **22**, 229–233.
- Gajek, L. & Zagrodny, D. (2000). Insurer's optimal reinsurance strategies. *Insurance: Mathematics and Economics* 27, 105–112.
- Huberman, G., Mayers, D. & Smith Jr, C. W. (1983). Optimal insurance policy indemnity schedules. *The Bell Journal of Economics* 14, 415–426.
- Gajek, L. & Zagrodny, D. (2004). Reinsurance arrangements maximizing insurer's survival probability. *Journal of Risk and Insurance* 71, 421–435.

552 👄 S. C. ZHUANG ET AL.

- Jouini, E., Schachermayer, W. & Touzi, N. (2008). Optimal risk sharing for law invariant monetary utility functions. *Mathematical Finance* 18, 269–292.
- Kaluszka, M. (2001). Optimal reinsurance under mean-variance premium principles. *Insurance: Mathematics and Economics* 28, 61–67.
- Kaluszka, M. (2005). Optimal reinsurance under convex principles of premium calculation. *Insurance: Mathematics and Economics* **36**, 375–398.
- Komiya, H. (1988). Elementary proof for Sion's minimax theorem. Kodai mathematical journal 11, 5-7.
- Ludkovski, M. & Young, V. R. (2009). Optimal risk sharing under distorted probabilities. *Mathematics and Financial Economics* 2, 87–105.
- Raviv, A. (1979). The design of an optimal insurance policy. American Economic Review 69, 84–96.
- Wang, S. S., Young, V. R. & Panjer, H. H. (1997). Axiomatic characterization of insurance prices. Insurance: Mathematics and Economics 21, 173–183.
- Weng, C. & Zhuang, S. C., (Forthcoming). CDF formulation for solving an optimal reinsurance problem. *Scandinavian Actuarial Journal*.
- Xu, Z. Q., Zhou, X. Y. & Zhuang, S. C. (2015). Optimal insurance with rank-dependent utility and increasing indemnities, Working Paper. Available at SSRN: http://ssrn.com/abstract=2660113

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

- Young, V. R. (1999). Optimal insurance under Wang's premium principle. *Insurance: Mathematics and Economics* 25, 109–122.
- Zhuang, S. C., Weng, C., Tan, K. S. & Assa, H. (2016). Marginal Indemnification Function formulation for optimal reinsurance. *Insurance: Mathematics and Economics* 67, 65–76.

Appendix 1

This Appendix 1 provides the proof of Lemma 4.1 and Lemma 4.2. In order to prove these lemmas, we first define the sets $\mathcal{Q} := \{(f_1(z), f_R(f_1(z))) : (f_1, f_R) \in \mathcal{F}^2, \text{ for } z \in [0, M]\}, \mathcal{H} := \{(k(z), j(z)) : k(0) = j(0) = 0, 0 \le j'(z) \le k'(z) \le 1, \text{ for } z \in [0, M] \text{ a.e.}\}$, as well as the inverse function of k as $k^{-1}(t) := \inf\{z \in [0, M] : k(z) \ge t\}$ for $t \in [0, k(M)]$. The following lemmas are essential in proving Lemma 4.1 and Lemma 4.2.

Lemma A.1: $k(k^{-1}(x)) = x$ for $x \in [0, k(M)]$ and $f_R(f_I(x)) = j(x)$ for $x \in [0, M]$.

Proof: Let $y := k^{-1}(x)$ for a given $x \in [0, k(M)]$. If k(y) > x, then there exists q < y such that k(q) > x using the property that k is continuous. Therefore, the definition of $k^{-1}(x)$ implies $q \ge k^{-1}(x) = y$, which contradicts q < y. If k(y) < x, then there exists p > y such that k(p) < x. Then, it follows that $p \le k^{-1}(x) = y$ which contradicts p > y. Consequently, we have $k(k^{-1}(x)) = x$ for $x \in [0, k(M)]$.

By the definition of k^{-1} , for any $x \in [0, M]$, we have $k^{-1}(k(x)) \le x$. Therefore, it follows that $j(k^{-1}(k(x))) \le j(x)$. If $j(k^{-1}(k(x))) < j(x)$, we denote $y := k^{-1}(k(x)) < x$ which results in $k(y) = k(k^{-1}(k(x))) = k(x)$. Hence, j(x) - j(y) > k(x) - k(y) = 0 which contradicts the fact that $j'(z) \le k'(z)$ for $z \in [0, M]$ *a.e.* The proof is thus complete.

Using the above lemma, we obtain the following lemma which establishes the equivalence between Q and H.

Lemma A.2: The set Q is equivalent to the set H.

Proof: It is easy to show that $\mathcal{Q} \subseteq \mathcal{H}$ by the chain rule. Next, we prove that the opposite direction. For any (\hat{k}, \hat{j}) , define $f_I(\cdot) := k(\cdot)$ and $f_R(\cdot) := j(k^{-1}(\cdot))$. Then, by applying Lemma A.1, we have $f_R(f_I(x)) = j(k^{-1}(k(x))) = j(x)$ and $f_R(s) - f_R(t) = j(k^{-1}(s)) - j(k^{-1}(t)) \le k(k^{-1}(s)) - k(k^{-1}(t)) = s - t$ for any $0 \le t < s \le k(M) = f_I(M)$. The proof is thus complete.

For any $(k(z), j(z)) \in \mathcal{H}$, let us define

$$\ell(z) := \begin{cases} k(z) & \text{if } z \in [0, M], \\ k(M) + j(z) & \text{if } z \in [M, 2M]. \end{cases}$$
(A.1)

Moreover, we define $\overline{\mathcal{H}} := \{\ell(z) : \ell(z) \text{ is defined by Equation (A.1), } (k(z), j(z)) \in \mathcal{H}\}$. By Arzela–Ascoli's theorem, $\overline{\mathcal{H}}$ is a compact set, i.e. for any sequence $\ell_i \in \overline{\mathcal{H}}$, there exists a subsequence ℓ_{i_k} that converges uniformly to l^* which belongs to $\overline{\mathcal{H}}$. Therefore, \mathcal{H} is also a compact set under the norm $\rho((k_1, j_1), (k_2, j_2)) := \max_{z \in [0,M]} |k_1(z) - k_2(z)| + |j_1(z) - j_2(z)|$.

Lemma A.2 enables us to write problems (5), (16), and (17), respectively, as

$$\max_{\substack{k,j\\ s,t}} \quad U_C(k,j) := \int_0^M m(z) \mathrm{d}k(z) + \int_0^M n(z) \mathrm{d}j(z),$$
s.t. $k, j \in \mathcal{H},$
(A.2)

$$\max_{k,j} \quad U_C(k,j) := \int_0^M m(z) dk(z) + \int_0^M n(z) dj(z),$$

s.t. $k,j \in \mathcal{H},$
 $\int_0^M h(S_X(z)) dj(z) \le C,$ (A.3)

and

$$\max_{k,j} \quad U_C(\lambda,k,j) := \int_0^M m(z) \mathrm{d}k(z) + \int_0^M n(z) \mathrm{d}j(z) - \lambda \Big(\int_0^M h(S_X(z)) \mathrm{d}j(z) - C \Big),$$
s.t. $k, j \in \mathcal{H}.$

$$(A.4)$$

Let $(\widetilde{k}, \widetilde{j})$ be an optimal solution to problem (A.2). If $\int_0^M h(S_X(z)) d\widetilde{j}(z) \leq C$, then $(\widetilde{k}, \widetilde{j})$ solves (A.3). Therefore, we assume that $\int_0^M h(S_X(z)) d\widetilde{j}(z) > C$ for all $(\widetilde{k}, \widetilde{j})$ that solves (A.2). Then, we have the following lemma.

Lemma A.3: If (k^*, j^*) solves problem (A.3), then $\int_0^M h(S_X(z)) dj^*(z) = C$.

Proof: For any $\theta \in [0, 1]$, we define $j_{\theta} := \theta \tilde{j} + (1 - \theta) j^*$. If $\int_0^M h(S_X(z)) dj^*(z) < C$, then there exists θ^* such that $\int_0^M h(S_X(z)) dj_{\theta}(z) = C$. Moreover, it follows that

$$U_{C}(\theta^{*}\widetilde{k} + (1 - \theta^{*})k^{*}, \theta^{*}\widetilde{j} + (1 - \theta^{*})j^{*}) = \theta^{*}U_{C}(\widetilde{k}, \widetilde{j}) + (1 - \theta^{*})U_{C}(k^{*}, j^{*}) > U_{C}(k^{*}, j^{*}),$$

where we use the fact that (k^*, j^*) is not the optimal solution to problem (A.2). This is a contradiction to the fact that (k^*, j^*) solves problem (A.3). Thus the proof is complete.

For the following lemma, its proof can be found in Komiya (1988).

Lemma A.4: (Sion's Minimax Theorem) Let X be a compact convex subset of a linear topological space and Y a convex subset of a linear topological space. If f is a real-valued function on $X \times Y$ such that $f(x, \cdot)$ is continuous and concave on Y for all $x \in X$, and $f(\cdot, y)$ is continuous and convex on X for all $y \in Y$, then, $\min_{x \in X} \max_{y \in Y} f(x, y) = \max_{y \in Y} \min_{x \in X} f(x, y)$.

We are now ready to present the proof of Lemma 4.1.

Proof of Lemma 4.1: Let C > 0. Moreover, we define v(C) and $v(\lambda, C)$ as the optimal values of (A.3) and (A.4), respectively. We first prove that $v(\lambda, C)$ is a convex function in λ for a given C. Noting that $U_C(\lambda, k, j)$ is linear in λ for any given (k, j), we have

$$\begin{aligned} v(\alpha\lambda_1 + (1-\alpha)\lambda_2, C) &= \max_{k,j} U_C(\alpha\lambda_1 + (1-\alpha)\lambda_2, k, j) \\ &= \max_{k,j} \left\{ \alpha U_C(\lambda_1, k, j) + (1-\alpha)U_C(\lambda_2, k, j) \right\} \\ &\leq \max_{k,j} \left\{ \alpha U_C(\lambda_1, k, j) + \max_{k,j} \left\{ (1-\alpha)U_C(\lambda_2, k, j) \right\} \\ &= \alpha \max_{k,j} \left\{ U_C(\lambda_1, k, j) + (1-\alpha) \max_{k,j} \left\{ U_C(\lambda_2, k, j) \right\} \\ &= \alpha v(\lambda_1, C) + (1-\alpha)v(\lambda_2, C). \end{aligned}$$

Moreover, by Sion's Minimax Theorem, the following equality holds: $\max_{0 \le \lambda} \min_{k,j \in \mathcal{H}} - U_C(\lambda, k, j) = \min_{k,j \in \mathcal{H}} \max_{0 \le \lambda} - U_C(\lambda, k, j)$; hence $\min_{0 \le \lambda} \max_{k,j \in \mathcal{H}} U_C(\lambda, k, j) = \max_{k,j \in \mathcal{H}} \min_{0 \le \lambda} U_C(\lambda, k, j)$. Finally, we have $\nu(C) = \inf_{0 \le \lambda} \nu(\lambda, C)$ (i.e. $U_C(k^*, j^*) = \min_{0 \le \lambda} \max_{k,j \in \mathcal{H}} U_C(\lambda, k, j)$).

By defining $\overline{\lambda} := \frac{\nu(C)+1}{C} = \frac{U_C(k^*j^*)+1}{C} > 0$, then for any $\lambda \ge \overline{\lambda}$ we have

$$\nu(\lambda, C) = \max_{k, j \in \mathcal{H}} U_C(\lambda, k, j) \ge U_C(\lambda, 0, 0) = \lambda C \ge \nu(C) + 1$$

which yields $\nu(C) = \inf_{0 \le \lambda} \nu(\lambda, C) = \inf_{0 < \lambda < \overline{\lambda}} \nu(\lambda, C).$

Therefore, using the convexity of $\nu(\lambda, C)$, there exists an $\lambda^* \in [0, \overline{\lambda}]$ that minimizes the right part, and is such that $\nu(C) = \nu(\lambda^*, C)$. Moreover,

$$\begin{aligned} v(\lambda^*, C) &\geq U_C(\lambda^*, k^*, j^*) = \int_0^M m(z) dk^*(z) + \int_0^M n(z) dj^*(z) - \lambda \Big(\int_0^M h(S_X(z)) dj^*(z) - C \Big) \\ &= \int_0^M m(z) dk^*(z) + \int_0^M n(z) dj^*(z) = U_C(k^*, j^*) = v(C). \end{aligned}$$

The second equality follows from the fact that (k^*, j^*) is the optimal solution to (A.3) for a given *C*. Hence, it follows from Lemma (A.3) that $\int_0^M h(S_X(z))dj^*(z) = C$. By $v(C) = v(\lambda^*, C)$ and $v(\lambda^*, C) \ge U_C(\lambda^*, k^*, j^*) = v(C)$, we have that (k^*, j^*) is optimal solution to (A.4) under a given λ^* . Hence, we know that (k^*, j^*) is the optimal solution to (A.4)

S. C. ZHUANG ET AL. 554

under a given λ^* and satisfying $\int_0^M h(S_X(z)) dj^*(x) = C$. If $\lambda^* = 0$, then $\nu(C) = \nu(0, C)$. This is a contradiction with the assumption that $\int_0^M h(S_X(z)) d\tilde{j}(x) > C$ for all (\tilde{k}, \tilde{j}) which are the optimal solutions to problem (A.2). So, $\lambda^* > 0$. If λ_C is strictly increasing in C, then there exists $C_1 < C_2$ such that $\lambda_{C_1} < \lambda_{C_2}$. It follows from the definitions of A_C and λ_C that there exist λ_1 and λ_2 such that $\lambda_1 \in A_{C_1}$ and $\lambda_2 \in A_{C_2}$ with $\lambda_1 < \lambda_2$. Lemma 4.1 implies that $\int_0^M h(S_X(z)) d\tilde{j}_{\lambda_1}(x) = C_1$ and $\int_0^M h(S_X(z)) d\tilde{j}_{\lambda_2}(x) = C_2$. Moreover, Theorem 3.1 suggests that \tilde{j}_{λ} can be represented as follows:

$$\widetilde{(j_{\lambda})}'(z) \stackrel{a.e.}{=} \begin{cases}
1, & \text{if } (1+\lambda)h(S_{X}(z)) < \min\{g_{I}(S_{X}(z)), g_{P}(S_{X}(z))\}, \\
\kappa(z) & \text{if } (1+\lambda)h(S_{X}(z)) = \min\{g_{I}(S_{X}(z)), g_{P}(S_{X}(z))\}, \\
0, & \text{if } (1+\lambda)h(S_{X}(z)) > \min\{g_{I}(S_{X}(z)), g_{P}(S_{X}(z))\},
\end{cases} \tag{A.5}$$

where $\kappa(z)$ is any [0,1]-valued function on $\{z \in [0,M] : (1+\lambda)h(S_X(z)) = \min\{g_I(S_X(z)), g_P(S_X(z))\}\}$. Obviously, by (A.5), we have $\tilde{j}'_{\lambda_2}(z) \leq \tilde{j}'_{\lambda_1}(z)$ on $\{z \in [0, M] : h(S_X(z)) \neq 0\}$, which results in that $\int_0^M h(S_X(z))d\tilde{j}_{\lambda_2}(z) \leq \int_0^M h(S_X(z))d\tilde{j}_{\lambda_1}(z)$. This is a contradiction to the fact that $\int_0^M h(S_X(z))d\tilde{j}_{\lambda_1}(z) = C_1 < C_2 = \int_0^M h(S_X(z))d\tilde{j}_{\lambda_2}(z)$. This concludes the proof of Lemma 4.1.

Similarly, by Lemma A.2, problems (20) and (21) can be written, respectively, as

$$\max_{\substack{f_{l}f_{R}\\ s.t.}} \quad U_{\alpha}(k,j) := \int_{0}^{M} m(z) dk(z) + \int_{0}^{M} n(z) dj(z),$$
s.t.
$$k, j \in \mathcal{H},$$

$$\int_{0}^{M} h(S_{X}(z)) dj(z) \leq \alpha \int_{0}^{M} g_{P}(S_{X}(z)) dk(z),$$
(A.6)

and

$$\max_{\substack{f_1,f_R\\\text{s.t.}}} \quad U_{\alpha}(\gamma,k,j) := \int_0^M m(z) \mathrm{d}k(z) + \int_0^M n(z) \mathrm{d}j(z) - \gamma \left(\int_0^M h(S_X(z)) \mathrm{d}j(z) - \alpha \int_0^M g_P(S_X(z)) \mathrm{d}k(z)\right),$$
(A.7)
s.t. $k, j \in \mathcal{H}.$

Let $(\widetilde{k}, \widetilde{j})$ denote the optimal solution to problem (A.2). If $\int_0^M h(S_X(z))d\widetilde{j}(z) \le \alpha \int_0^M g_P(S_X(z))d\widetilde{k}(z)$, then $(\widetilde{k}, \widetilde{j})$ solves problem (A.3). Therefore, we assume that $\int_0^M h(S_X(z))d\widetilde{j}(z) > \alpha \int_0^M g_P(S_X(z))d\widetilde{k}(z)$ for all $(\widetilde{k}, \widetilde{j})$. Before proving Lemma 4.2, it is useful to have the following lemma.

Lemma A.5: If (k^*, j^*) solves problem (A.6), then $\int_0^M h(S_X(z)) dj^*(z) = \alpha \int_0^M g_P(S_X(z)) dk^*(z)$.

Proof: We assume that $\int_0^M h(S_X(z))dj^*(z) < \alpha \int_0^M g_P(S_X(z))dk^*(z)$. For any $\theta \in [0, 1]$, we define $k_{\theta} := \theta \widetilde{k} + (1 - \theta)k^*$, $j_{\theta} := \theta \widetilde{j} + (1 - \theta)j^*$ and $p(\theta) := \int_0^M h(S_X(z))dj_{\theta}(z) - \alpha \int_0^M g_P(S_X(z))dk_{\theta}(z)$. Since p is continuous, p(0) < 0 and p(1) > 0, there exists a θ^* such that $p(\theta^*) = 0$ due to the intermediate value theorem. Therefore, it follows that

$$U_{\alpha}(\theta^{*}\tilde{k} + (1 - \theta^{*})k^{*}, \theta^{*}\tilde{j} + (1 - \theta^{*})j^{*}) = \theta^{*}U_{\alpha}(\tilde{k}, \tilde{j}) + (1 - \theta^{*})U_{\alpha}(k^{*}, j^{*}) > U_{\alpha}(k^{*}, j^{*}),$$

where we use the fact that (k^*, j^*) is not an optimal solution to problem (A.2) as this would be a contradiction to the fact that (k^*, j^*) solves problem (A.6). This concludes the proof.

Proof of Lemma 4.2: As the proof to this lemma is very similar to that of Lemma 4.1, we only highlight their difference. We assume that $\alpha \in (0, 1]$ is given and let $\hat{\nu}(\alpha)$ and $\hat{\nu}(\gamma, \alpha)$ be the optimal values of (20) and (21), respectively. By Lemma A.2, problem (21) reduces to

$$\max_{f_{1},f_{R}} \quad \hat{U}(\gamma,j,k) := \int_{0}^{M} m(z) dk(z) + \int_{0}^{M} n(z) dj(z) - \gamma \left(\int_{0}^{M} h(S_{X}(z)) dj(z) - \alpha \int_{0}^{M} g_{P}(S_{X}(z)) dk(z) \right),$$
s.t. $(f_{1},f_{R}) \in \mathcal{F}^{2}.$
(A.8)

Thus, we need to set $\overline{\gamma} := \frac{\hat{v}(\alpha) + 1 - \int_0^M m(z) dz}{\alpha \int_0^M g_P(S_X(z)) dz} > 0$. For any $\gamma \ge \overline{\gamma}$, we have

$$\hat{v}(\gamma,\alpha) = \max_{k,j\in\mathcal{H}} \hat{U}(\gamma,k,j) \ge \hat{U}(\gamma,z,0) = \int_0^M m(z) dz + \gamma \alpha \int_0^M g_P(S_X(z)) dz \ge \hat{v}(\alpha) + 1.$$

The remaining proof is analogous to that of the first part of the proof of Lemma 4.1.