Optimal Risk Transfer under Quantile-Based Risk Measures

Alexandru V. Asimit[1]
Cass Business School, City University, London EC1Y 8TZ, United Kingdom. E-mail: asimit@city.ac.uk

Alexandru M. Badescu
Department of Mathematics and Statistics, University of Calgary, Calgary, Alberta T2N 1N4 , Canada. E-mail: abadescu@math.ucalgary.ca

Tim Verdonck
Department of Mathematics, KU Leuven, 3001 Heverlee, Belgium
E-mail: tim.verdonck@wis.kuleuven.be

June 12, 2013

Abstract. The classical problem of identifying the optimal risk transfer from one insurance company to multiple reinsurance companies is examined under some quantile-based risk measure criteria. We develop a new methodology via a two-stage optimisation procedure which allows us to not only recover some existing results in the literature, but also makes possible the analysis of high dimensional problems in which the insurance company diversifies its risk with multiple reinsurance counter-parties, where the insurer risk position and the premium charged by the reinsurers are functions of the underlying risk quantile. Closed form solutions are elaborated for some particular settings, although numerical methods for the second part of our procedure represent viable alternatives for the ease of implementing it in more complex scenarios. Furthermore, we discuss some approaches to obtain more robust results.

Keywords and phrases: Expected Shortfall, Distorted Risk Measure, Premium Principle, Optimal Reinsurance, Truncated Tail-Value-at-Risk , Value-at-Risk.

1. Introduction

The optimality problem of the risk transfer contract between two insurance companies within a one-period setting appears in different forms in the literature. The first attempts are attributed to Borch (1960) and Arrow (1963) where maximising the expected utility defines the optimality criterion. Further extensions have been developed for various decision criteria that depend on the risk measure choice (for example, see Heerwaarden *et al.*, 1989, Young, 1999, Kaluska, 2001 and 2005, Verlaak and Beirlant, 2003, Kaluszka and Okolewski, 2008, Ludrovski and Young, 2009). Decisions based on two particular risk measures, *Value-at-risk (VaR)* and *Expected Shortfall (ES)*, are considered by Cai *et al.* (2008), Cheung (2010) and Chi and Tan (2011). All of the afore-mentioned papers deal with the one-period model. The classical risk model setting has been successfully studied in the literature by Centeno and Guerra (2010) and Guerra and Centeno (2008 and 2010), via maximisation of the adjustment coefficient.

There are two parties involved in a reinsurance contract: the *insurer* or *cedent* who has an interest in transferring part of its risk, and the *reinsurer*. Let $X \geq 0$ be the total loss amount incurred during the duration of the insurance contract, with distribution function denoted by $F(\cdot)$ and survival function $\bar{F}(\cdot) = 1 - F(\cdot)$. In addition, the right end-point $x_F := \inf\{z \in \Re : F(z) = 1\}$ of the loss distribution can be either finite or infinite. The reinsurance company agrees to pay, $R[X]$, the amount by which the entire loss exceeds the insurer amount, $I[X]$. Thus, $I[X] + R[X] = X$. There are many possible reinsurance arrangements, which depend on the particular choice of the

[1]Corresponding author. Phone: +44(0)2070405282. Fax: +44(0)2070408572.

insurer and reinsurer of sharing the premiums and underwritten risks. For example, the liabilities are shared in a fixed proportion under *proportional* reinsurance and therefore $I[X] = cX$, where $0 < c < 1$ is a constant. Another common arrangement is the *stop-loss* reinsurance contracts, for which the cedent loss is limited to a fixed amount, M, known as *retention limit*. The net amount paid by the insurer is therefore given by $\min\{X, M\} := X \wedge M$.

The reinsurer premium, $\mathbf{P}\big(R[X]\big)$, is usually assumed to satisfy $\mathbf{P}\big(R[X]\big) \geq \mathbf{E}\big(R[X]\big)$, since otherwise the risk bearer would become insolvent almost surely. Obviously, the total insurer loss becomes

$$\mathbf{L}\big(R[X]\big) := I[X] + \mathbf{P}\big(R[X]\big).$$

The aim of this paper is to identify the optimal arrangement that lays the cedent in the best possible situation towards the risk. That is, we intend to minimise $\varphi_I\Big(\mathbf{L}\big(R[X]\big)\Big)$ over a set of feasible reinsurance contracts, where φ_I represents a measure of the risk taken by the insurer. Motivated by the standard regulatory requirements developed within the insurance industry, the risk exposures are measured via VaR and ES risk measures. The VaR of a generic loss variable Z at a confidence level α, $VaR_\alpha(Z)$, represents the minimum amount of capital that makes the insurance company to be solvent at least $\alpha\%$ of the time. The mathematical formulation is then given by

$$VaR_\alpha(Z) := \inf\{z \geq z_0 : \Pr(Z \leq z) \geq \alpha\},$$

where $z_0 := \sup\{z \in \Re : \Pr(Z \leq z) = 0\}$ represents the left end point of the distribution of Z. The ES at a confidence level α, $ES_\alpha(Z)$, evaluates the expected loss amount incurred among the worst $(1-\alpha)\%$ scenarios. Clearly, ES represents a more conservative risk measure than VaR, especially in the situation in which both risk quantifications are made at the same confidence level. The ES has multiple formulations in the literature, and the comprehensive papers on this topic of Acerbi and Tasche (2002) and Hürlimann (2003) may help in clarifying the differences between equivalent representations. We only refer to the next definition:

$$ES_\alpha(Z) := \frac{1}{1-\alpha}\int_\alpha^1 VaR_s(Z)\,ds = VaR_\alpha(Z) + \frac{1}{1-\alpha}\mathbf{E}\big(Z - VaR_\alpha(Z)\big)_+, \tag{1.1}$$

where $(z)_+ = \max\{z, 0\}$. Interestingly, this risk measure is a special case of the Haezendonck-Goovaerts class, which was introduced many years ago by Haezendonck and Goovaerts (1982). Further details can be found in Bellini and Rosazza Gianin (2012), Goovaerts *et al.* (2004 and 2012) and the references therein.

As a result of the translation invariance property of the three risk measures, the following holds

$$\varphi_I\Big(\mathbf{L}\big(R[X]\big)\Big) = \varphi_I\big(I[X]\big) + \mathbf{P}\big(R[X]\big).$$

In order to avoid potential moral hazard issues related to the reinsurance arrangement, the set of feasible contracts is given by

$$\mathcal{F} := \big\{R(\cdot) : I(x) = x - R(x) \; and \; R(x) \; are \; non-decreasing \; functions\big\}.$$

It is useful to note that $R \in \mathcal{F}$ implies that the functions I and R are Lipschitz functions with unit constants, i.e. $|I(y) - I(x)| \leq |y - x|$ and $|R(y) - R(x)| \leq |y - x|$ are true for all $x, y \geq 0$. Therefore, our optimisation problem is reduced to

$$\min_{R \in \mathcal{F}} \Big\{\varphi_I\big(X\big) - \varphi_I\big(R[X]\big) + \mathbf{P}\big(R[X]\big)\Big\}, \tag{1.2}$$

since $I[X]$ and $R[X]$ are co-monotone random variables (for details, see Dhaene *et al.*, 2002 a and b, Denuit *et al.*, 2005). In other words, the co-monotonicity property implies that

$$\varphi_I\big(X\big) = \varphi_I\big(I[X]\big) + \varphi_I\big(R[X]\big).$$

It should be noted that we give explicit derivations whenever the insurer's risk position is evaluated via the ES and VaR risk measures, but our procedure can be easily applied to any risk measure that is a function of the risk quantile. In Section 4, we will illustrate this for a more robust risk measure, namely the *Truncated Tail Value-at-Risk (TrTVaR)*. A large class of such risk measures is given by

$$M_{\Phi}(Z) := \int_0^1 VaR_s(Z)\Phi(s)\,ds,$$

where the function $\Phi(\cdot)$ has certain properties. This class is known as the *distorted* (see Wang and Young, 1998 and Jones and Zitikis, 2003) and *spectral* (see Acerbi, 2002) class of risk measures, respectively. Therefore, our procedure is widely applicable to situations in which the insurer risk position is evaluated via many well-known risk measures.

A common premium principle used in practice is the *expected value* principle. That is, the reinsurer premium is loaded as follows:

$$\mathbf{P}\big(R[X]\big) = (1+\rho)\mathbf{E}\big(R[X]\big),$$

where $\rho > 0$ is known as the *security loading factor*. The main problem defined by equation 1.2 has been previously investigated in the literature when the insurer has set VaR or ES as the baseline risk measure. Cai *et al.* (2008) and Cheung (2010) found the optimal reinsurance contract over a class of convex functions $R(\cdot)$. Note that both papers assume $F(\cdot)$ to be a strictly increasing and continuous function, and therefore their *Conditional Tail Expectation* evaluation coincides with that of ES. Chi and Tan (2011) elaborate a two-stage optimisation procedure to solve the VaR and ES problem, where the solution of the first problem is given a priori.

Once again, our procedure can be applied to other premium principles that are functions of the underlying risk quantile. Section 3 contains some distorted premium principles that should help in supporting our statement. The expected value principle is chosen only for the sake of exposition, and also to recover some existing results from the literature.

So far, it has been implicitly assumed that the insurer may transfer the risk to only one reinsurance counter-party. More realistic situations involve multiple reinsurance risk transfers available on the market. It is likely that each reinsurer has its own pricing model, and the cedent may choose to transfer specific layers from the total risk to competitive companies. Therefore, the insurer may improve the diversification gain by sharing the loss with multiple reinsurance counter-parties. At this point, it is worth mentioning that all previously-mentioned papers assume the reinsurance market to consist of only one agent.

In this paper, we present a general two-stage based algorithm for identifying optimal arrangements when the insurer risk is diversified through multiple reinsurance agreements. More specifically, we investigate some optimisation problems within a trivariate risk transfer setting with different assumptions regarding the premium principles used by the two reinsurance counter-parties. According to our knowledge, the proposed method represents the only choice for an insurance company of taking advantage of different available reinsurance pricing schemes on the market. However, this situation tends to be cumbersome if more than two reinsurance companies are potential players in the risk transfer game. We derive closed form solutions for two particular scenarios, but numerical methods are required for solving the second stage problem in order to overcome this potential computational issue. In the one dimensional reinsurance market case, our approach can also be viewed as an alternative solution to existing methods, but we clearly derive the solution of the first optimisation problem. However, our proposed method has the advantage of solving more complex situations, whenever the insurer's risk position and the reinsurer premiums are functions of the risk quantile. Note that our approach can be extended to many more reinsurance companies than just two.

The rest of the paper is organized as follows. The next section gives a different perspective for the ideal ES and VaR-based reinsurance arrangement under the assumption that there is only one reinsurance company on the market, for which some of the results can already be found in the existing literature. The third section illustrates some VaR and ES-based optimal decisions for an insurer, whenever the cedent shares the risk with two other reinsurance agents. The fourth section deals with robust estimation of the optimal reinsurance arrangement. In this section we will also obtain the ideal TrTVaR-based reinsurance arrangement given that there is only one reinsurance company on the market. The final section comprises a numerical analysis of some of our main results.

2. Optimal Single Reinsurance Contract

The current section explains our method of finding the optimal risk transfer whenever there is only one reinsurance company. This lays the foundation for extending our methodology to more complex situations where multiple risk absorbers are available to the risk transfer initiator. This section also provides an alternative proof for some existing results in the literature, and additionally, we are able to find all optimal solutions via our constructive methodology. Expected value premium principle is set to be used by the reinsurer, while the insurer risk is measured by VaR and ES, as has been anticipated in the previous section. These settings are chosen to simply explain our constructive method, but one may apply it for any situation in which the insurer's criterion and reinsurance premium are functions of the underlying risk quantile. The method consists of a two-stage optimisation problem, where the first one is an infinite dimensional problem, while the second stage becomes a classical constrained optimisation problem. Therefore, the first stage is the key step in our procedure, and it can be solved as shown in Proposition 2.1.

Proposition 2.1. *Let $f(\cdot)$ be a real valued function defined on $[s_1, s_2]$ with $0 \le s_1 \le s_2 \le 1$. Then,*

$$\min_{R\in\mathcal{F}} \int_{s_1}^{s_2} f(s) R\big(VaR_s(X)\big)\, ds \quad \text{subject to} \quad R\big(VaR_{s_1}(X)\big) = \xi_1, R\big(VaR_{s_2}(X)\big) = \xi_2,$$

is uniquely solved by

$$R^*[X;\xi_1,\xi_2] = \begin{cases} \big(X - VaR_{s_1}(X) + \xi_1\big) \wedge \xi_2, & \text{if } f(s) < 0 \text{ for all } s_1 \le s \le s_2 \\ \xi_1 + \big(X - VaR_{s_2}(X) + \xi_2 - \xi_1\big)_+, & \text{if } f(s) > 0 \text{ for all } s_1 \le s \le s_2 \end{cases},$$

where (ξ_1, ξ_2) are some constants such that $0 \le \xi_2 - \xi_1 \le VaR_{s_2}(X) - VaR_{s_1}(X)$.

Proof. Without loss of generality, we may only investigate the case in which $\xi_1 = 0$ and $\xi_2 = \xi$. Let $R \in \mathcal{F}$ be a generic risk transfer function such that $R\big(VaR_{s_1}(X)\big) = \xi_1$, $R\big(VaR_{s_2}(X)\big) = \xi_2$. We first assume that the function $f(\cdot)$ takes only negative values, and we show that $R(x) \le R^*(x; 0, \xi)$ holds for all $x \in [VaR_{s_1}(X), VaR_{s_2}(X)]$. If there exists $x_0 \in [VaR_{s_1}(X), VaR_{s_2}(X)]$ satisfying $R^*(x_0; 0, \xi) < R(x_0)$, then $x_0 < \xi$ due to the boundary condition. Thus,

$$R(x_0) - R\big(VaR_{s_1}(X)\big) = R(x_0) > R^*(x_0; 0, \xi) = x_0 - VaR_{s_1}(X),$$

which contradicts the Lipschitz continuity property. Therefore, the first scenario is concluded.

Similarly, if $f(\cdot)$ takes only positive values, then it is sufficient to show that $R(x) \ge R^*(x; 0, \xi)$ holds for all $x \in [VaR_{s_1}(X), VaR_{s_2}(X)]$. If there exists $x_0 \in [VaR_{s_1}(X), VaR_{s_2}(X)]$ satisfying $R(x_0) < R^*(x_0; 0, \xi)$, then $VaR_{s_2}(X) - \xi < x_0 \le VaR_{s_2}(X)$ due to the boundary condition. Thus,

$$R\big(VaR_{s_2}(X)\big) - R(x_0) = \xi - R(x_0) > \xi - R^*(x_0; 0, \xi) = VaR_{s_2}(X) - x_0,$$

which contradicts the Lipschitz continuity property. Thus, the first proof is now complete. □

We are now ready to exemplify our procedure, and we start out with the VaR criterion, which is given in Theorem 2.1.

Theorem 2.1. *Let $\rho^* = \rho/(1+\rho)$. The VaR-based optimal decision that minimises the insurer total loss from 1.2 is given by*

$$R^*[X] = \big(X - VaR_{\rho^*}(X)\big)_+ \wedge \big(VaR_\alpha(X) - VaR_{\rho^*}(X)\big)_+. \tag{2.1}$$

Therefore, the corresponding cedent risk becomes

$$VaR_\alpha\Big(\mathbf{L}\big(R^*[X]\big)\Big) = VaR_\alpha(X) \wedge VaR_{\rho^*}(X) + (1+\rho)\int_{VaR_\alpha(X)\wedge VaR_{\rho^*}(X)}^{VaR_\alpha(X)} \bar{F}(x)\, dx. \tag{2.2}$$

Proof. Recall that we need to minimise

$$\begin{aligned} VaR_\alpha\Big(\mathbf{L}\big(R[X]\big)\Big) &= VaR_\alpha(X) - VaR_\alpha\big(R[X]\big) + (1+\rho)\mathbf{E}\big(R[X]\big) \\ &= VaR_\alpha(X) - R\big(VaR_\alpha(X)\big) + (1+\rho)\int_0^1 R\big(VaR_s(X)\big)\, ds \end{aligned}$$

over $\mathcal{F}$. Note that the last relation is true since $R(\cdot)$ is a non-decreasing continuous function. The above optimisation problem is solved via a two-stage procedure, and the first step is given by

$$\min_{R\in\mathcal{F}} \int_0^1 R\big(VaR_s(X)\big)\, ds \quad \text{subject to} \quad R\big(VaR_\alpha(X)\big) = \xi, \tag{2.3}$$

where $0 \le \xi \le VaR_\alpha(X)$ is a constant. Proposition 2.1 allows us to conclude that 2.3 is solved by

$$R^*[X;\xi] := \big(X - VaR_\alpha(X) + \xi\big)_+ \wedge \xi. \tag{2.4}$$

Now,

$$\mathbf{E}\big(R^*[X;\xi]\big) = \int_0^\xi \Pr\big(R^*[X;\xi] > x\big)\, dx = \int_{VaR_\alpha(X)-\xi}^{VaR_\alpha(X)} \Pr(X > x)\, dx.$$

Thus, the second step is to minimise

$$H_1(\xi) := -\xi + (1+\rho)\int_{VaR_\alpha(X)-\xi}^{VaR_\alpha(X)} \bar{F}(x)\, dx \tag{2.5}$$

over the set $[0, VaR_\alpha(X)]$. Clearly, the derivative $H_1'(\xi) = -1 + (1+\rho)\bar{F}\big(VaR_\alpha(X) - \xi\big)$ takes non-positive values if and only if $VaR_\alpha(X) - VaR_{\rho^*}(X) \ge \xi$. Thus, if $VaR_\alpha(X) \ge VaR_{\rho^*}(X)$ then $H_1(\cdot)$ is minimised at $VaR_\alpha(X) - VaR_{\rho^*}(X)$, which replicates 2.1 in this case. Similarly, $H_1(\cdot)$ attains its global minimum at 0, whenever $VaR_\alpha(X) < VaR_{\rho^*}(X)$. Thus, the insurer should choose $R^*[X;0] = 0$. The proof is now complete. □

We are now able to find the optimal ES-based decision, which is further developed in Theorem 2.2.

Theorem 2.2. *The ES-based optimal reinsurance contract that solves 1.2 is given by*

$$R^*[X] = \begin{cases} \big(X - VaR_{\rho^*}(X)\big)_+, & \alpha > \rho^* \\ 0, & \alpha < \rho^* \\ h_1^*(X), & \alpha = \rho^* \end{cases},$$

where $h_1^*(\cdot)$ is a non-decreasing Lipschitz function with unit constant satisfying $h_1^*\big(VaR_\alpha(X)\big) = 0$. Therefore, the corresponding insurer risk becomes

$$ES_\alpha\Big(\mathbf{L}\big(R^*[X]\big)\Big) = \begin{cases} VaR_{\rho^*}(X) + (1+\rho)\int_{VaR_{\rho^*}(X)}^{x_F} \bar{F}(x)\, dx., & \alpha > \rho^* \\ ES_\alpha(X), & \alpha \le \rho^* \end{cases}.$$

Proof. Recall that we aim to minimise

$$\begin{aligned} ES_\alpha\Big(\mathbf{L}\big(R[X]\big)\Big) &= ES_\alpha(X) - ES_\alpha\big(R[X]\big) + (1+\rho)\mathbf{E}\big(R[X]\big) \\ &= ES_\alpha(X) - \frac{1}{1-\alpha}\int_\alpha^1 R\big(VaR_s(X)\big)\, ds + (1+\rho)\int_0^1 R\big(VaR_s(X)\big)\, ds \end{aligned}$$

over $\mathcal{F}$, where the second identity is due to equation 1.1 and the fact that $R(\cdot)$ is a non-decreasing continuous function. Once again, the latter is solved via a two-stage optimisation procedure, and the first step is given by

$$\begin{cases} \min_{R\in\mathcal{F}}(1+\rho)\displaystyle\int_0^\alpha R\big(VaR_s(X)\big)\,ds+\Big(1+\rho-\frac{1}{1-\alpha}\Big)\int_\alpha^1 R\big(VaR_s(X)\big)\,ds \\ \quad\text{subject to}\quad R\big(VaR_\alpha(X)\big)=\xi, \end{cases} \tag{2.6}$$

where the parameter ξ satisfies $0\le\xi\le VaR_\alpha(X)$.

Let us first assume that $\alpha>\rho^*$, which implies that $1+\rho-\frac{1}{1-\alpha}<0$. Thus, Proposition 2.1 shows that 2.6 is solved by $R^*[X;\xi]:=\big(X-VaR_\alpha(X)+\xi\big)_+$. Next,

$$\begin{aligned} H_2(\xi) &:= -ES_\alpha\big(R^*[X;\xi]\big)+(1+\rho)\mathbf{E}\big(R^*[X;\xi]\big)\\ &= -ES_\alpha(X)+VaR_\alpha(X)-\xi+(1+\rho)\int_{VaR_\alpha(X)-\xi}^{x_F}\bar{F}(x)\,dx, \end{aligned}$$

needs to be minimised over the set $[0,VaR_\alpha(X)]$. Note that the above differs from $H_1(\cdot)$ defined in 2.5 by just a constant, and therefore has the same behaviour. Similar reasoning to the one used in the proof of Theorem 2.1 shows that $H_2(\cdot)$ is minimised at $VaR_\alpha(X)-VaR_{\rho^*}(X)$, which completes the proof in this case.

The mirror setting $\alpha<\rho^*$ is considered. Now, the optimal reinsurance contract of 2.6 is given by 2.4, due to Proposition 2.1 and the fact that $1+\rho-\frac{1}{1-\alpha}>0$. The second stage in our optimisation can be solved as before.

Finally, the case in which $\alpha=\rho^*$ is discussed in greater detail. We have no available information on the behaviour of the optimal solution on $[VaR_\alpha(X),x_F]$, and the solution of 2.6 is no longer unique. The set of possible solutions is given by

$$R^*[X;\xi]:=\begin{cases}\big(X-VaR_\alpha(X)+\xi\big)_+, & X\le VaR_\alpha(X)\\ h_1^*(X;\xi), & X>VaR_\alpha(X)\end{cases},$$

where $h_1^*(\cdot;\xi)$ is a Lipschitz function with unit constant such that $h_1^*\big(VaR_\alpha(X);\xi\big)=\xi$. These are consequences of Proposition 2.1 and the fact that $R(\cdot)$ is a Lipschitz function with unit constant. The second stage optimisation problem is reduced to minimising

$$\begin{aligned} H_3(\xi) &:= \int_0^\alpha\big(VaR_s(X)-VaR_\alpha(X)+\xi\big)_+\,ds\\ &= \int_{F\left(VaR_\alpha(X)-\xi\right)}^\alpha\big(VaR_s(X)-VaR_\alpha(X)+\xi\big)\,ds \end{aligned}$$

over $[0,VaR_\alpha(X)]$. The above is increasing in ξ, since

$$H_3(\xi_1)\le\int_{F\left(VaR_\alpha(X)-\xi_1\right)}^\alpha\big(VaR_s(X)-VaR_\alpha(X)+\xi_2\big)\,ds\le H_3(\xi_2)$$

for all $0\le\xi_1<\xi_2\le VaR_\alpha(X)$. The second inequality is true as a result of

$$VaR_s(X)-VaR_\alpha(X)+\xi_2>0,\ \text{ for all } F\big(VaR_\alpha(X)-\xi_2\big)<s\le F\big(VaR_\alpha(X)-\xi_1\big).$$

Thus, $H_3(\cdot)$ attains its minimum at 0. It may be worth mentioning that the objective function that we start with has its minimal value

$$ES_\alpha(X)+(1+\rho)\int_0^\alpha\big(VaR_s(X)-VaR_\alpha(X)\big)_+\,ds=ES_\alpha(X).$$

The last case is now elucidated, which completes the proof. □

Note that the VaR-based insurer optimal contract includes extremely high layers as opposed to the ES-based ideal solution. This is due to the fact that VaR is blind to more extreme situations than its level, while ES incorporates the entire tail risk. It should also be noted that our findings from Theorem 2.1 replicate some existing results, such as Theorem 3.2 of Chi and Tan (2011). Moreover, most of our findings from Theorem 2.2 fully overlap with previous results from Theorem 4.1 of Chi and Tan (2011). In most of the cases, the optimal solution is unique, as observed in the aforementioned paper. Our first stage optimisation problem does not have a single solution whenever $\alpha = \rho^*$, which makes the optimal solution of the initial problem to not be unique anymore. The previous approaches are able to capture only one optimal solution, which is due to the fact that their solution for the first optimisation problem is given a priori, and is sensitive to the problem to be solved, while the proposed approach is an instructive method of finding the optimal solution(s). Our method is able to identify the entire set of optimal solutions, which obviously lays the insurer at the same level of risk, but it has the advantage of allowing the decision-maker to choose the best arrangement among these optimal solutions according to a different criterion. For example, if the insurer wishes to find the ES-based optimal risk transfer for $\alpha_1 = \rho^*$, then Theorem 2.2 says that the whole risk prior to $VaR_{\rho*}(X)$ should be retained by the reinsurer, while the remaining layer can be shared with the reinsurer in any possible way. In addition, if the insurer decides to find the ideal way of sharing layers higher than $VaR_{\rho*}(X)$, the optimal VaR_{α_2}-based decision, with $\alpha_2 > \rho^*$, is as given by Theorem 2.1. Thus, the multiple criteria decision (the optimal $VaR_{\alpha_2}(X)$ contract among the set of ES_{α_1}-based optimal risk transfer, where $\alpha_1 = \rho^* < \alpha_2$), is given by

$$I^*[X] = X \wedge VaR_{\rho^*}(X) + \big(X - VaR_{\alpha_2}(X)\big).$$

In the next section, one may find why previous methods, based on knowledge of the solution a priori, fail to be applicable since the optimal solution becomes quite cumbersome. In contrast, our constructive method is able to accommodate even higher dimensional problems where there are more than one reinsurers that may absorb the insurer's risk.

3. Optimal Multiple Reinsurance Contract

It is natural to consider situations in which an insurance company may be able to share the risk with multiple reinsurers that use different premium principles. Clever allocation of the risk layers among the players, would produce a better position for the insurer towards the risk, as a result of risk reduction. In this section, we derive optimal solutions for the VaR-based problem when the reinsurance market consists of two agents with different pricing schemes. The two-stage optimisation procedure allows one to tackle complex situations with more than two agents, although analysis becomes cumbersome in high dimension. We choose two settings in order to exemplify our procedure in such cases. The key point is to solve the first stage problem, while the second one can be solved numerically. Generally speaking, high dimensional problems are numerically tractable, but one ought to carefully design the variable restrictions associated with the second stage problem. A general assumption that is required by our proposed method is that each reinsurer could price any risk layer, although we do not exclude the possibility of discouraging the insurer to place certain layers by some reinsurers.

The bivariate reinsurance problem has the following mathematical formulation

$$\min_{(R_1,R_2)\in\mathcal{G}} VaR_\alpha\Big(\mathbf{L}\big(R_1[X], R_2[X]\big)\Big) = VaR_\alpha\big(I[X]\big) + \mathbf{P_1}\big(R_1[X]\big) + \mathbf{P_2}\big(R_2[X]\big), \tag{3.1}$$

where $\mathcal{G} := \big\{(R_1, R_2) : I(x) = x - R_1(x) - R_2(x), R_1(x) \; and \; R_2(x) \; are \; non-decreasing \; functions\big\}$.

Next, an explanation is provided for the above requirements, though standard assumptions that impose the amounts paid by insurance and reinsurance players to be increasing functions of the total loss. The starting point lies within the fundamentals of the reinsurance market mechanism.

Therefore, whenever multiple insurers share a risk, the risk allocation process is made sequentially. Specifically, a feasible allocation for our setting from 3.1 should satisfy

(**C1**) I and R_1 are non-decreasing functions of $X - R_2$
(**C2**) I and R_2 are non-decreasing functions of $X - R_1$
(**C3**) R_1 and R_2 are non-decreasing functions of $X - I$

in order to avoid potential moral hazard issues with such allocation.

Proposition 3.1. *Under conditions (**C1**) to (**C3**), I, R_1 and R_2 are non-decreasing functions of X. In addition, I, R_1 and R_2 are Lipschitz functions with unit constants.*

Proof. Let ω_1 and ω_2 be two possible outcomes such that $X(\omega_1) \leq X(\omega_2)$. We may assume without loss of generality that $I(\omega_1) \leq I(\omega_2)$, and therefore it is sufficient to show that $R_1(\omega_1) \leq R_1(\omega_2)$ and $R_2(\omega_1) \leq R_2(\omega_2)$. Now, if $R_2(\omega_1) > R_2(\omega_2)$, then condition (**C2**) implies that

$$X(\omega_1) - R_1(\omega_1) > X(\omega_2) - R_1(\omega_2) \Rightarrow I(\omega_1) \geq I(\omega_2).$$

Thus, $I(\omega_1) = I(\omega_2)$, which in turn suggests that $X(\omega_1) > X(\omega_2)$, since condition (**C3**) implies that $X(\omega_1) - I(\omega_1) > X(\omega_2) - I(\omega_2)$. Consequently, our assumption is violated, and therefore I, R_1 and R_2 should be non-decreasing functions of the total risk X.

The Lipschitz property is proved only for the insurer retained loss. In other words, we need to show that $I(\omega_2) - I(\omega_1) \leq X(\omega_2) - X(\omega_1)$ is true as long as $I(\omega_1) \leq I(\omega_2)$. Clearly,

$$I(\omega_2) - I(\omega_1) = X(\omega_2) - X(\omega_1) - \big(R_1(\omega_2) - R_1(\omega_1)\big) - \big(R_2(\omega_2) - R_2(\omega_1)\big) \leq X(\omega_2) - X(\omega_1).$$

The proof is now complete. □

As it has been demonstrated that the infinite dimensional optimisation problem given in the first stage of our procedure is the key step in finding the optimal risk transfer. The next result extends Proposition 2.1 when there are two risk absorbers.

Proposition 3.2. *Let $f_1(\cdot)$ and $f_2(\cdot)$ be two real valued functions defined on $[s_1, s_2]$ with $0 \leq s_1 \leq s_2 \leq 1$. Then,*

$$\min_{R \in \mathcal{G}} \int_{s_1}^{s_2} f_1(s) R_1\big(VaR_s(X)\big)\, ds + \int_{s_1}^{s_2} f_2(s) R_2\big(VaR_s(X)\big)\, ds \quad \textit{subject to} \tag{3.2}$$
$$R_1\big(VaR_{s_1}(X)\big) = \xi_{11}, R_1\big(VaR_{s_2}(X)\big) = \xi_{12}, R_2\big(VaR_{s_1}(X)\big) = \xi_{21}, R_2\big(VaR_{s_2}(X)\big) = \xi_{22},$$

is uniquely solved by

i) *if $f_1(s) < f_2(s) < 0$ for all $s_1 \leq s \leq s_2$, then*
$$\begin{cases} R_1^*[X;\underline{\xi}] = \big(X - VaR_{s_1}(X) + \xi_{11}\big)_+ \wedge \xi_{12} \\ R_2^*[X;\underline{\xi}] = \xi_{21} + \big(X - VaR_{s_1}(X) - (\xi_{12} - \xi_{11})\big)_+ \wedge (\xi_{22} - \xi_{21}) \end{cases},$$

ii) *if $f_1(s) < 0 < f_2(s)$ for all $s_1 \leq s \leq s_2$, then*
$$\begin{cases} R_1^*[X;\underline{\xi}] = \big(X - VaR_{s_1}(X) + \xi_{11}\big)_+ \wedge \xi_{12} \\ R_2^*[X;\underline{\xi}] = \xi_{21} + \big(X - VaR_{s_2}(X) + \xi_{22} - \xi_{21}\big)_+ \end{cases},$$

iii) *if $0 < f_1(s) < f_2(s)$ for all $s_1 \leq s \leq s_2$, then*
$$\begin{cases} R_1^*[X;\underline{\xi}] = \xi_{11} + \big(X - VaR_{s_2}(X) + \xi_{12} - \xi_{11} + \xi_{22} - \xi_{21}\big)_+ \wedge (\xi_{12} - \xi_{11}) \\ R_2^*[X;\underline{\xi}] = \xi_{21} + \big(X - VaR_{s_2}(X) + \xi_{22} - \xi_{21}\big)_+ \end{cases},$$

where $\underline{\xi} = (\xi_{11}, \xi_{12}, \xi_{21}, \xi_{22}) \in \mathcal{D}$ is a vector of constants and

$$\mathcal{D} := \begin{cases} 0 \leq \xi_{11} \leq \xi_{12}, 0 \leq \xi_{21} \leq \xi_{22}, \xi_{11} + \xi_{21} \leq VaR_{s_1}(X), \xi_{12} + \xi_{22} \leq VaR_{s_2}(X), \\ \xi_{12} - \xi_{11} + \xi_{22} - \xi_{21} \leq VaR_{s_2}(X) - VaR_{s_1}(X). \end{cases}$$

Proof. Let $(R_1, R_2) \in \mathcal{G}$ such that the boundary conditions are satisfied, as required in our optimisation problem from 3.2.

Part i) is first investigated. If (R_1^*, R_2^*) is not the unique solution of 3.2, then it may happen that

$$\int_{s_1}^{s_2}\Big(f_1(s)\Big(R_1\big(VaR_s(X)\big)-R_1^*\big(VaR_s(X);\underline{\xi}\big)\Big)+f_2(s)\Big(R_2\big(VaR_s(X)\big)-R_2^*\big(VaR_s(X);\underline{\xi}\big)\Big)\Big)ds<0.\quad(3.3)$$

Note that $R_1(x) - R_1^*(x;\underline{\xi}) \leq 0$ holds for all $x \in [VaR_{s_1}(X), VaR_{s_2}(X)]$, which is a result of Proposition 2.1. The latter and the fact that $f_1(s) - f_2(s) < 0$ is true for all $s_1 \leq s \leq s_2$ imply that the left hand side of 3.3 is greater than

$$\int_{s_1}^{s_2} f_2(s)\Big(R_1\big(VaR_s(X)\big)+R_2\big(VaR_s(X)\big)-R_1^*\big(VaR_s(X);\underline{\xi}\big)-R_2^*\big(VaR_s(X);\underline{\xi}\big)\Big)ds$$
$$=\int_{s_1}^{s_2} f_2(s)\Big(I^*\big(VaR_s(X);\underline{\xi}\big)-I\big(VaR_s(X)\big)\Big)ds \geq 0,\quad(3.4)$$

where the last step is a result of Proposition 2.1, i.e. $I^*(x;\underline{\xi}) \leq I(x)$ is true on $[VaR_{s_1}(X), VaR_{s_2}(X)]$. Thus, 3.4 contradicts 3.3, and therefore (R_1^*, R_2^*) is indeed the unique optimal solution of 3.2 for part i).

Part ii) simply follows from the fact that

$$R_1(x) \leq R_1^*(x;\underline{\xi}) \text{ and } R_2(x) \geq R_2^*(x;\underline{\xi}), \text{ for all } x \in [VaR_{s_1}(X), VaR_{s_2}(X)],$$

which are consequences of Proposition 2.1.

Part iii) can be proven in the same manner as Part i), and therefore is left to the reader. Thus, the proof is now complete. □

We need to specify the premium calculation used by each reinsurance company in order to evaluate the insurer's optimal decision. Our two-stage optimisation problem can be applied to any premium principle that is quantile-based, i.e. the premium charged by reinsurer is a function of the quantile function corresponding to its paid amount. We derive the ideal reinsurance contract among all VaR-based decisions, but one can extend our findings for ES-based decisions, or any other objective function that is based on a risk measure which is a function of the insurer's risk quantile. Note that all distorted and spectral risk measures are risk quantile functions.

A major class of quantile-based risk measures is the distorted class, for which the following definition is needed.

Definition 3.1. *A distortion function is a non-decreasing function* $g : [0, 1] \to [0, 1]$ *such that* $g(0) = 0$ *and* $g(1) = 1$.

Any distortion function defines a distorted expectation, which represents a risk measure and can be useful in premium calculations (see for example, Jones and Zitikis, 2003). The insurer chooses its optimal VaR-based contract by transferring the risk to one reinsurer that uses the expected value principle, i.e. $\mathbf{P}\big(R_1[X]\big) = (1+\rho)\mathbf{E}\big(R_1[X]\big)$, while the second reinsurer prefers the *distorted* premium principle. Thus,

$$\mathbf{P}\big(R_2[X]\big) = \int_0^{x_F} g\Big(\Pr\big(R_2[X] > x\big)\Big)\, dx = \int_0^1 VaR_s\big(R_2[X]\big)\Phi(s)\, ds,$$

where $g(\cdot)$ is a distorted function and $\Phi(s) := g^{'}(1-s)$. Note that $g(\cdot)$ is not assumed to be differentiable on its domain. Since $g(\cdot)$ is concave, its right and left derivatives always exist and the usual derivative exists almost everywhere. Thus, the above integral does not change its value due to the fact that at most a null-measure set could possibly be removed (for details, see Dhaene *et al.*, 2012).

The specific optimisation problem that is elaborated in this section can be summarised as follows:

$$\min_{(R_1,R_2)\in\mathcal{G}} VaR_\alpha\big(I[X]\big) + (1+\rho)\int_0^1 VaR_s\big(R_1[X]\big)\,ds + \int_0^1 VaR_s\big(R_2[X]\big)\Phi(s)\,ds. \tag{3.5}$$

Closed form solutions are obtained if the premium principles are fully specified. It should not be surprising if the optimal arrangement is heavily sensitive to the reinsurer pricing strategy. The first result allows for diversification gain, while the presence of the second reinsurer becomes futile from the insurer point of view in Theorem 3.2. The next set of assumptions are sufficient to describe our findings from Theorem 3.1.

Assumption 3.1. *Let $g(\cdot)$ be a distortion function with corresponding positive function $\Phi(\cdot)$ such that $\limsup_{s\to 0} g(s)=0$ and $g'_+(0)>1+\rho$, where $g'_+(\cdot)$ represents the right-derivative function.*

We now discuss our assumptions. Clearly, the concavity of $g(\cdot)$ implies that $\Phi(\cdot)$ is non-negative and non-decreasing. The positivity assumptions allows us to remove the possibility of having infinitely many optimal solutions for our first stage optimisation problem. The first assumption on function $g(\cdot)$ tells us that a jump at 0 for $g(\cdot)$ is excluded, which could be possible, since the concavity guarantees continuity only on $(0,1)$. The key condition $g'_+(0)>1+\rho$ implies that there exist $s^{**}\le s^*<1$ such that:

i) $s^* := inf\{s:\Phi(s)\ge 1+\rho\}\ge \rho^*$;

ii) s^{**} is the unique non-trivial solution of $G(1-s^{**})=0$, where $G(t):=(1+\rho)t-g(t)$.

We only need to argue that $\rho^*\le s^*$ holds. Note that $G(t)<0$ for all $0<t<1-s^{**}$, and therefore if $s^*<\rho^*$, then $G(1-\rho^*)<0$. The latter implies that $g\big(1/(1+\rho)\big)>1$, which is false, and in turn one finds that $\rho^*\le s^*$ is true.

The next theorem establishes situations in which the second reinsurer sets its premium not only on the tail risk, and consequently takes into consideration the entire loss spectrum.

Theorem 3.1. *Let us assume that Assumption 3.1 holds such that $s^{**}\le\alpha$ Then, the VaR-based optimal reinsurance contract described in 3.5 is as follows:*

$$R_1^*[X] = \big(X - VaR_{s^{**}}(X)\big)_+ \wedge \big(VaR_\alpha(X) - VaR_{s^{**}}(X)\big),\ R_2^*[X] = X\wedge VaR_{s^{**}}(X), \tag{3.6}$$

and the corresponding insurer risk becomes

$$VaR_\alpha\Big(\mathbf{L}\big(R_1^*[X],R_2^*[X]\big)\Big) = (1+\rho)\int_{VaR_{s^{**}}(X)}^{VaR_\alpha(X)} \bar F(x)\,dx + \int_0^{VaR_{s^{**}}(X)} g\big(\bar F(x)\big)\,dx. \tag{3.7}$$

Note 3.1. *Theorem 3.1 shows that one is not able to simply apply a sequential optimisation for high dimensional problems. That is, the optimal solution cannot be incrementally obtained by initially finding the insurer's allocation in the presence of the first reinsurance only, as shown in Theorem 2.1, and then the insurer to look for sharing the remaining risk with the second reinsurer.*

Proof. The proof is composed of two stages, and initially additional restrictions are added to our main optimisation problem defined in 3.5. Thus, provided that $s^*<\alpha$,

$$\left\{\begin{array}{l}\displaystyle\min_{(R_1,R_2)\in\mathcal{G}} VaR_\alpha\big(I[X]\big)+(1+\rho)\int_0^1 VaR_s\big(R_1[X]\big)\,ds+\int_0^1 VaR_s\big(R_2[X]\big)\Phi(s)\,ds\\ \text{subject to}\ \ R_1\big(VaR_{s^*}(X)\big)=\xi_{11},\,R_1\big(VaR_\alpha(X)\big)=\xi_{12},\\ \qquad\qquad\ \ R_2\big(VaR_{s^*}(X)\big)=\xi_{21},\,R_2\big(VaR_\alpha(X)\big)=\xi_{22},\end{array}\right. \tag{3.8}$$

is solved at the moment, where $\underline{\xi}=(\xi_{11},\xi_{12},\xi_{21},\xi_{22})\in\mathcal{D}_1$ is a vector of constants and

$$\mathcal{D}_1 := \left\{\begin{array}{l}0\le\xi_{11}\le\xi_{12},0\le\xi_{21}\le\xi_{22},\xi_{11}+\xi_{21}\le VaR_{s^*}(X),\xi_{12}+\xi_{22}\le VaR_\alpha(X),\\ \xi_{12}-\xi_{11}+\xi_{22}-\xi_{21}\le VaR_\alpha(X)-VaR_{s^*}(X).\end{array}\right.$$

We are now able to solve 3.8 by applying Proposition 3.2 twice, but keeping in mind Assumption 3.1 and its consequences. A graphical representation of the optimal solution is provided in Figure 3.1, where without loss of generality $\xi_{11} \leq \xi_{21}$ and $\xi_{12} \leq \xi_{22}$ are assumed to hold.

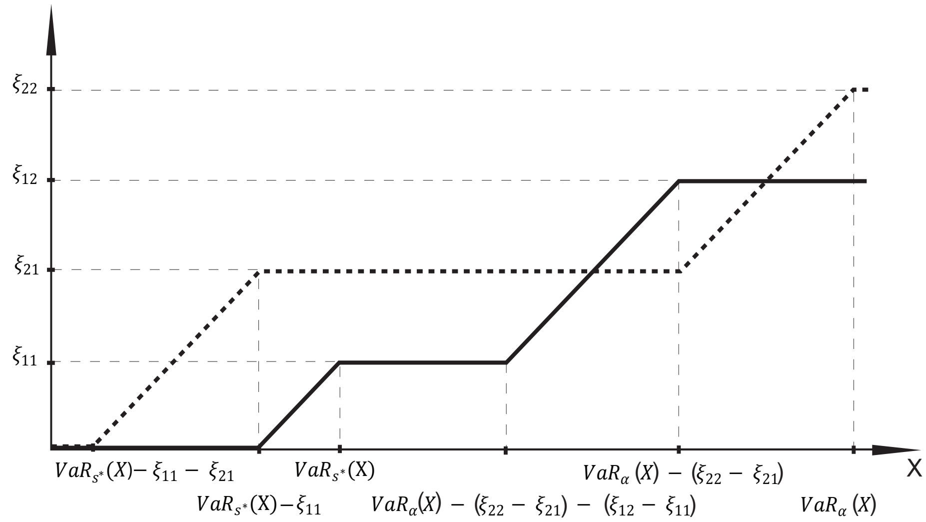

FIGURE 3.1. The construction of $R_1^*[X;\underline{\xi}]$ (solid line) and $R_2^*[X;\underline{\xi}]$ (dotted line) in Theorem 3.1 when $s^* < \alpha$.

As a result, the solution of 3.8 is given by

$$R_1^*[X;\underline{\xi}] := \begin{cases} \big(X - VaR_{s^*}(X) + \xi_{11}\big)_+ \wedge \xi_{11}, X \leq VaR_{\alpha}(X) - (\xi_{12} - \xi_{11}) - (\xi_{22} - \xi_{21}) \\ \big(X - VaR_{\alpha}(X) + (\xi_{22} - \xi_{21}) + \xi_{12}\big)_+ \wedge \xi_{12}, \text{ otherwise} \end{cases}$$

and

$$R_2^*[X;\underline{\xi}] := \begin{cases} \big(X - VaR_{s^*}(X) + \xi_{11} + \xi_{21}\big)_+ \wedge \xi_{21}, X \leq VaR_{\alpha}(X) - (\xi_{22} - \xi_{21}) \\ \big(X - VaR_{\alpha}(X) + \xi_{22}\big)_+ \wedge \xi_{22}, \text{ otherwise} \end{cases}.$$

Note that the above are true even if either $\xi_{11} \leq \xi_{21}$ or $\xi_{12} \leq \xi_{22}$ is not satisfied. Consequently, the second step optimisation problem is equivalent to

$$\min_{\mathcal{D}_1} G_1(\underline{\xi}) := VaR_{\alpha}\big(X\big) - \xi_{12} - \xi_{22} + \int_{VaR_{s^*}(X)-\xi_{11}-\xi_{21}}^{VaR_{s^*}(X)-\xi_{11}} g\big(\bar{F}(x)\big)\,dx + \int_{VaR_{\alpha}(X)-(\xi_{22}-\xi_{21})}^{VaR_{\alpha}(X)} g\big(\bar{F}(x)\big)\,dx$$
$$+(1+\rho)\left(\int_{VaR_{s^*}(X)-\xi_{11}}^{VaR_{s^*}(X)} \bar{F}(x)\,dx + \int_{VaR_{\alpha}(X)-(\xi_{12}-\xi_{11})-(\xi_{22}-\xi_{21})}^{VaR_{\alpha}(X)-(\xi_{22}-\xi_{21})} \bar{F}(x)\,dx\right). \tag{3.9}$$

Now, $\xi_{11} \leq \xi_{12} \leq VaR_{\alpha}(X) - VaR_{s^*}(X) - (\xi_{22} - \xi_{21}) + \xi_{11}$ suggests that $\mathcal{D}_1$ is a simple region with respect to ξ_{12}, since the boundary curves remain the same when varying $\big\{\xi_{11}, \xi_{21}, \xi_{22}\big\}$. Next,

$$\frac{dG_1}{d\xi_{12}} = -1 + (1+\rho)\bar{F}\big(VaR_{\alpha}(X) - (\xi_{12} - \xi_{11}) - (\xi_{22} - \xi_{21})\big) \leq -1 + (1+\rho)\bar{F}\big(VaR_{s^*}(X)\big),$$

which is non-positive if $VaR_{\rho^*}(X) \leq VaR_{s^*}(X)$. The latter is obviously ensured by Assumption 3.1, and therefore 3.9 is the same as minimising

$$\begin{aligned} G_2(\xi_{11}, \xi_{21}, \xi_{22}) &:= G_1\big(\xi_{11}, VaR_{\alpha}(X) - VaR_{s^*}(X) - (\xi_{22} - \xi_{21}) + \xi_{11}, \xi_{21}, \xi_{22}\big) \qquad (3.10) \\ &= VaR_{s^*}\big(X\big) - \xi_{11} - \xi_{21} + (1+\rho)\int_{VaR_{s^*}(X)-\xi_{11}}^{VaR_{\alpha}(X)-(\xi_{22}-\xi_{21})} \bar{F}(x)\,dx \\ &\quad + \int_{VaR_{s^*}(X)-\xi_{11}-\xi_{21}}^{VaR_{s^*}(X)-\xi_{11}} g\big(\bar{F}(x)\big)\,dx + \int_{VaR_{\alpha}(X)-(\xi_{22}-\xi_{21})}^{VaR_{\alpha}(X)} g\big(\bar{F}(x)\big)\,dx, \end{aligned}$$

over $\mathcal{D}_2 := \big\{0 \le \xi_{11}, 0 \le \xi_{21}, \xi_{11}+\xi_{21} \le VaR_{s^*}(X), 0 \le \xi_{22}-\xi_{21} \le VaR_\alpha(X) - VaR_{s^*}(X)\big\}$. Note that $\mathcal{D}_2$ is a simple region with respect to ξ_{22} since $\xi_{21} \le \xi_{22} \le VaR_\alpha(X) - VaR_{s^*}(X) + \xi_{21}$. Thus,

$$\begin{aligned}\frac{dG_2}{d\xi_{22}} &= -(1+\rho)\bar{F}\big(VaR_\alpha(X) - (\xi_{22}-\xi_{21})\big) + g\Big(\bar{F}\big(VaR_\alpha(X) - (\xi_{22}-\xi_{21})\big)\Big)\\ &= -G\Big(\bar{F}\big(VaR_\alpha(X) - (\xi_{22}-\xi_{21})\big)\Big).\end{aligned}$$

Clearly, $G(0)=0$ and $G(1)=\rho>0$. Moreover, Assumption 3.1 implies that $G(\cdot)$ attains its global minimum at $1-s^*$. Thus, $s^{**}<s^*$ and $G(t)\le 0$ for all $0\le t\le 1-s^{**}$. The latter and the fact that $\bar{F}\big(VaR_\alpha(X)-(\xi_{22}-\xi_{21})\big) \le \bar{F}\big(VaR_{s^*}(X)\big) \le 1-s^*$ suggest the non-decreasing property in ξ_{22} of $G_2(\cdot)$. This allows us to further conclude that 3.10 is equivalent to solving

$$\begin{aligned}&\min_{\mathcal{D}_3} G_3(\xi_{11},\xi_{21}) := G_2\big(\xi_{11},\xi_{21},\xi_{21}\big) \qquad (3.11)\\ &= VaR_{s^*}(X) - \xi_{11} - \xi_{21} + (1+\rho)\int_{VaR_{s^*}(X)-\xi_{11}}^{VaR_\alpha(X)} \bar{F}(x)\,dx + \int_{VaR_{s^*}(X)-\xi_{11}-\xi_{21}}^{VaR_{s^*}(X)-\xi_{11}} g\big(\bar{F}(x)\big)\,dx\end{aligned}$$

over $\mathcal{D}_3 := \big\{0\le \xi_{11}, 0\le\xi_{21}, \xi_{11}+\xi_{21}\le VaR_{s^*}(X)\big\}$. Obviously,

$$\frac{dG_3}{d\xi_{21}} = -1 + g\Big(\bar{F}\big(VaR_{s^*}(X)-\xi_{11}-\xi_{21}\big)\Big) \le 0,$$

and therefore 3.11 is reduce to finding the solution of

$$\begin{aligned}\min_{0\le\xi_{11}\le VaR_{s^*}(X)} G_4(\xi_{11}) &:= G_3(\xi_{11}, VaR_{s^*}(X)-\xi_{11})\\ &= (1+\rho)\int_{VaR_{s^*}(X)-\xi_{11}}^{VaR_\alpha(X)} \bar{F}(x)\,dx + \int_0^{VaR_{s^*}(X)-\xi_{11}} g\big(\bar{F}(x)\big)\,dx.\end{aligned}$$

Simple calculations lead to $\frac{dG_4}{d\xi_{11}} = G\Big(\bar{F}\big(VaR_{s^*}(X)-\xi_{11}\big)\Big)$. Recall that $G(\cdot)$ takes non-positive values whenever $0\le t\le 1-s^{**}$. The latter and the fact that

$$\bar{F}\big(VaR_{s^*}(X)-\xi_{11}\big) \le 1-s^{**} \Leftrightarrow VaR_{s^{**}}(X) \le VaR_{s^*}(X)-\xi_{11}$$

suggest that $G_4(\cdot)$ is non-increasing on $[0, VaR_{s^*}(X)-VaR_{s^{**}}(X)]$. Similarly, one may find that $G_4(\cdot)$ is non-decreasing on $[VaR_{s^*}(X)-VaR_{s^{**}}(X), VaR_{s^*}(X)]$. Thus, the global minimum of $G_4(\cdot)$ is attained at $VaR_{s^*}(X)-VaR_{s^{**}}(X)$, and simple calculations yield 3.6 and 3.7 in this setting. Therefore, the $s^*<\alpha$ case is fully explained.

We now assume that $s^{**}\le\alpha\le s^*$. The first stage optimisation problem is as given by 3.8 where the positions of s^* and α are swapped. Its solution is provided by Figure 3.2, and may be obtained via applying Proposition 3.2 twice.

The proof continues with similar derivations to the ones previously displayed when the $s^*<\alpha$ scenario was investigated. The derivatives are taken in the same order and yield the following global minimum solution

$$\xi_{11}^* = \xi_{12}^* = VaR_\alpha(X) - VaR_{s^{**}}(X),\;\; \xi_{21}^* = \xi_{22}^* = VaR_{s^{**}}(X),$$

which concludes the $s^{**}\le\alpha\le s^*$ scenario. Thus, the proof is now complete. □

Theorem 3.1 assumes that the second reinsurance company charges a non-purely tail risk based premium, and consequently $\Phi(s)>0$ over the whole domain of definition. If the latter is not satisfied, then there exists s_0 such that $\Phi(s)>0$ on $(s_0,1]$, and it stays at level zero for smaller values than s_0. In other words, the corresponding $g(t)\equiv 1$ whenever $t\ge 1-s_0$. One may construct multiple examples satisfying this property, but probably the most natural distorted premium principle example would be the ES one, where $g(t) := (t/(1-\beta))\wedge 1$ with $0\le\beta<1$. This scenario is discussed in Theorem 3.2. In addition, Theorem 3.1 always allows for diversification, while

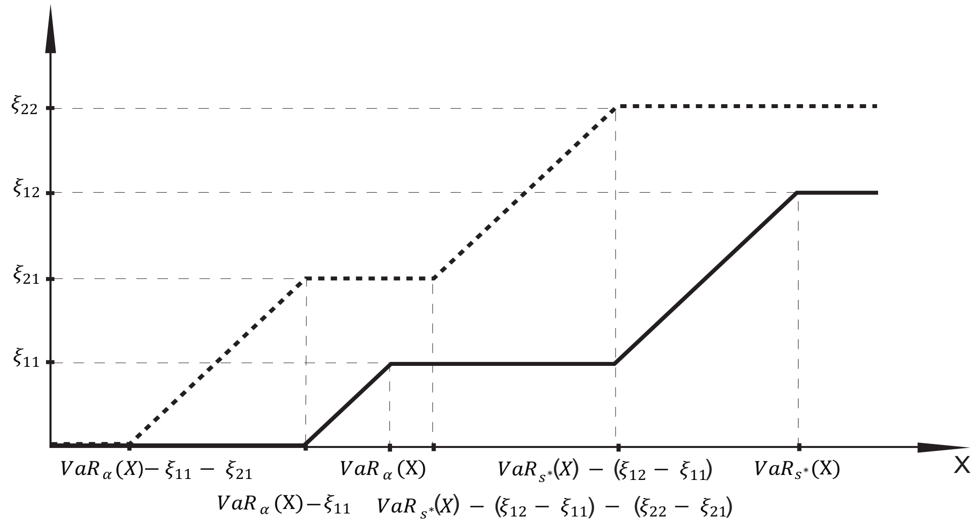

FIGURE 3.2. The construction of $R_1^*[X;\underline{\xi}]$ (solid line) and $R_2^*[X;\underline{\xi}]$ (dotted line) in Theorem 3.1 when $\alpha \le s^*$.

Theorem 3.2 shows an improved diversification only if $\beta < \rho^*$. Moreover, Theorem 3.2 provides infinitely many optimal solutions (due to the tail risk pricing choice made by the second reinsurer), as compared to Theorem 3.1, and we are able to find all of them as a result of our constructive proposed methodology of finding the ideal arrangement(s).

Theorem 3.2. *The VaR-based insurer risk problem*

$$\min_{(R_1,R_2)\in\mathcal{G}} VaR_\alpha\big(I[X]\big) + (1+\rho)\mathbf{E}\big(R_1[X]\big) + ES_\beta\big(R_2[X]\big), \tag{3.12}$$

is considered such that $\max\{\rho^*,\beta\} \le \alpha$. *Then, its optimal solution has the following composition:*

a) *whenever* $\beta < \rho^*$

$$R_1^*[X] \equiv 0; R_2^*[X] = \begin{cases} h_2^*(X), & X \le VaR_\beta(X) \\ X \wedge VaR_\alpha(X) - VaR_\beta(X) + a, & X > VaR_\beta(X) \end{cases},$$

b) *whenever* $\beta > \rho^*$

$$R_1^*[X] = \big(X - VaR_{\rho^*}(X)\big)_+ \wedge \big(VaR_\alpha(X) - VaR_{\rho^*}(X)\big); R_2^*[X] = h_3^*(X) \wedge b$$

with $a \in [0, VaR_\beta(X)]$ *and* $b \in [0, VaR_{\rho^*}(X)]$ *some parameters. In addition,* $h_2^*(\cdot)$ *and* $h_3^*(\cdot)$ *are non-decreasing Lipschitz functions with unit constants such that*

$$h_2^*(0) = h_3^*(0) = 0,\ h_2^*\big(VaR_\beta(X)\big) = a,\ h_3^*\big(VaR_{\rho^*}(X)\big) = b.$$

Therefore, the corresponding insurer risk becomes

$$VaR_\alpha\Big(\mathbf{L}\big(R_1^*[X], R_2^*[X]\big)\Big) = VaR_\beta(X) + \frac{1}{1-\beta}\int_{VaR_\beta(X)}^{VaR_\alpha(X)} \bar{F}(x)\,dx$$

if $\beta < \rho^*$, *and*

$$VaR_\alpha\Big(\mathbf{L}\big(R_1^*[X], R_2^*[X]\big)\Big) = VaR_{\rho^*}(X) + (1+\rho)\int_{VaR_{\rho^*}(X)}^{VaR_\alpha(X)} \bar{F}(x)\,dx,$$

if $\beta > \rho^*$.

Note 3.2.

i) *Essentially, the insurer and second reinsurer are able to share the layer* $[0, VaR_{\beta\wedge\rho^*}]$ *in any possible way due to the tail-based pricing offered by the second risk absorber. Therefore, the insurer may decide to achieve its optimal arrangement, among the infinitely many optimal solutions given in Theorem 3.2, by imposing an additional criterion. For example, one may show, via our proposed two-stage method, that maximising/minimising insurer's expected profit/loss among this set of VaR-based optimal solutions would be possible. This additional criterion leads to an overall unique optimal solution as follows:*

$$I^*[X] = X \wedge VaR_{\beta\wedge\rho^*}(X) + \big(X - VaR_\alpha(X)\big)_+,$$

which is attained if the insurer keeps the maximal possible risk for itself.

ii) *If* $\beta < \rho^*$ *then, for any possible risk layer, the first reinsurance company charges a higher premium than the second reinsurer, and therefore it is not surprising why it is not optimal to allocate any layer to the first reinsurer.*

iii) *Theorem 3.2 also includes the result from Theorem 2.1 if* $\beta > \rho^*$.

Proof. We derive the optimal solution of 3.12 in two steps. Initially,

$$\begin{cases} \min\limits_{(R_1,R_2)\in\mathcal{G}} VaR_\alpha\big(I[X]\big) + (1+\rho)\displaystyle\int_0^1 VaR_s\big(R_1[X]\big)\,ds + \frac{1}{1-\beta}\int_0^1 VaR_s\big(R_2[X]\big)\,ds \\ \text{subject to} \quad R_1\big(VaR_\beta(X)\big) = \xi_{11},\, R_1\big(VaR_\alpha(X)\big) = \xi_{12}, \\ \qquad\qquad\quad R_2\big(VaR_\beta(X)\big) = \xi_{21},\, R_2\big(VaR_\alpha(X)\big) = \xi_{22}, \end{cases} \tag{3.13}$$

is solved, where $\underline{\xi} = (\xi_{11}, \xi_{12}, \xi_{21}, \xi_{22}) \in \mathcal{E}_1$ is a vector of constants and

$$\mathcal{E}_1 := \begin{cases} 0 \le \xi_{11} \le \xi_{12}, 0 \le \xi_{21} \le \xi_{22}, \xi_{11} + \xi_{21} \le VaR_\beta(X), \xi_{12} + \xi_{22} \le VaR_\alpha(X), \\ \xi_{12} - \xi_{11} + \xi_{22} - \xi_{21} \le VaR_\alpha(X) - VaR_\beta(X). \end{cases}$$

Let us now assume that $\beta < \rho^*$. Thus, $1+\rho > 1/(1-\beta)$, and one may further apply Proposition 3.2 twice. A graphical representation of the optimal solution from 3.13 is provided in Figure 3.3, where without loss of generality $\xi_{11} \le \xi_{21}$ and $\xi_{12} \le \xi_{22}$ are assumed to hold.

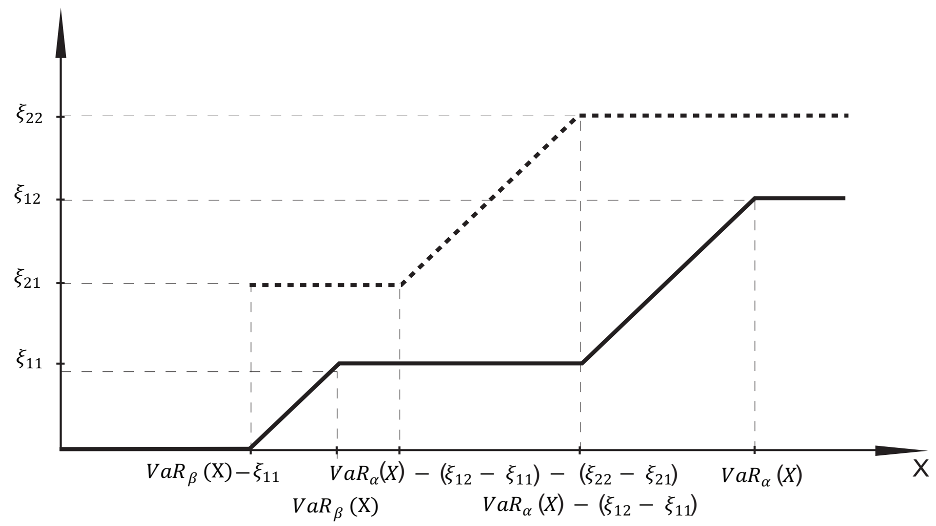

FIGURE 3.3. The construction of $R_1^*[X;\underline{\xi}]$ (solid line) and $R_2^*[X;\underline{\xi}]$ (dotted line) in Theorem 3.2 when $\beta < \rho^*$.

The solution of 3.13 can be formally written in the following fashion:

$$R_1^*[X;\underline{\xi}] := \begin{cases} \big(X - VaR_\beta(X) + \xi_{11}\big)_+ \wedge \xi_{11}, X \le VaR_\alpha(X) - (\xi_{12} - \xi_{11}) \\ \big(X - VaR_\alpha(X) + \xi_{12}\big) \wedge \xi_{12}, \text{ otherwise} \end{cases}$$

and

$$R_2^*[X;\underline{\xi}] := \begin{cases} h_2^*\big(X;\underline{\xi}\big) \wedge \xi_{21}, X \leq VaR_\alpha(X) - (\xi_{12}-\xi_{11}) - (\xi_{22}-\xi_{21}) \\ \big(X - VaR_\alpha(X) + \xi_{12} - \xi_{11} + \xi_{22}\big) \wedge \xi_{22}, \text{ otherwise} \end{cases},$$

where $h_2^*(\cdot;\underline{\xi})$ is a non-decreasing Lipschitz function with unit constant such that $h_2^*(0;\underline{\xi}) = 0$ and $h_2^*\big(VaR_\beta(X) - \xi_{11};\underline{\xi}\big) = \xi_{21}$. Consequently, the second step optimisation problem is equivalent to

$$\min_{\mathcal{E}_1} G_1(\underline{\xi}) := VaR_\alpha(X) - \xi_{12} - \xi_{22} + (1+\rho)\left(\int_{VaR_\beta(X)-\xi_{11}}^{VaR_\beta(X)} \bar{F}(x)\,dx + \int_{VaR_\alpha(X)-(\xi_{12}-\xi_{11})}^{VaR_\alpha(X)} \bar{F}(x)\,dx\right)$$
$$+\xi_{21} + \frac{1}{1-\beta}\int_{VaR_\alpha(X)-(\xi_{12}-\xi_{11})-(\xi_{22}-\xi_{21})}^{VaR_\alpha(X)-(\xi_{12}-\xi_{11})} \bar{F}(x)\,dx. \tag{3.14}$$

Note that $\xi_{21} \leq \xi_{22} \leq VaR_\alpha(X) - VaR_\beta(x) - (\xi_{12}-\xi_{11}) + \xi_{21}$ always holds, suggesting that $\mathcal{E}_1$ represents a ξ_{22}-simple region. The latter and the fact that

$$\frac{dG_1}{d\xi_{22}} = -1 + \frac{1}{1-\beta}\bar{F}\big(VaR_\alpha(X) - (\xi_{12}-\xi_{11}) - (\xi_{22}-\xi_{21})\big) \leq -1 + \frac{1}{1-\beta}\bar{F}\big(VaR_\beta(X)\big) \leq 0$$

imply that 3.14 is reduced to minimising

$$\begin{aligned} G_2(\xi_{11},\xi_{12},\xi_{21}) &:= G_1\big(\xi_{11},\xi_{12},\xi_{21}, VaR_\alpha(X) - VaR_\beta(X) - (\xi_{12}-\xi_{11}) + \xi_{21}\big) \\ &= VaR_\beta(X) - \xi_{11} + (1+\rho)\left(\int_{VaR_\beta(X)-\xi_{11}}^{VaR_\beta(X)} + \int_{VaR_\alpha(X)-(\xi_{12}-\xi_{11})}^{VaR_\alpha(X)}\right)\bar{F}(x)\,dx \\ &\quad + \frac{1}{1-\beta}\int_{VaR_\beta(X)}^{VaR_\alpha(X)-(\xi_{12}-\xi_{11})} \bar{F}(x)\,dx, \end{aligned} \tag{3.15}$$

over $\mathcal{E}_2 := \big\{0 \leq \xi_{11}, 0 \leq \xi_{21}, \xi_{11}+\xi_{21} \leq VaR_\beta(X), 0 \leq \xi_{12}-\xi_{11} \leq VaR_\alpha(X) - VaR_\beta(X)\big\}$. The above is a ξ_{12}-simple region since $\xi_{11} \leq \xi_{12} \leq VaR_\alpha(X) - VaR_\alpha(X) + \xi_{11}$ holds. Next,

$$\frac{dG_2}{d\xi_{12}} = \left(1 + \rho - \frac{1}{1-\beta}\right)\bar{F}\big(VaR_\alpha(X) - (\xi_{12}-\xi_{11})\big) \geq 0,$$

and consequently 3.15 is further reduced to

$$\begin{aligned} \min_{\mathcal{E}_3} G_3(\xi_{11},\xi_{21}) &:= G_2\big(\xi_{11},\xi_{11},\xi_{21}\big) \\ &= VaR_\beta(X) - \xi_{11} + (1+\rho)\int_{VaR_\beta(X)-\xi_{11}}^{VaR_\beta(X)} \bar{F}(x)\,dx + \frac{1}{1-\beta}\int_{VaR_\beta(X)}^{VaR_\alpha(X)} \bar{F}(x)\,dx, \end{aligned}$$

where $\mathcal{E}_3 := \{0 \leq \xi_{11}, 0 \leq \xi_{21}, \xi_{11}+\xi_{21} \leq VaR_\beta(X)\}$. For any fixed ξ_{21}, the above function is increasing in ξ_{11}. The latter is true since its partial derivative with respect to ξ_{11} is positive as a result of $VaR_\beta(X) - \xi_{11} \leq VaR_\beta(X) \leq VaR_{\rho^*}(X)$. Therefore, the minimum is attained at $\xi_{11}^* = 0$ and ξ_{21}^* may take any value from $[0, VaR_\beta(X)]$. Some simple calculations allow one to recover the results stated in this theorem whenever $\beta < \rho^*$.

Only the main steps of the other case, i.e. $\rho^* < \beta$, are given, since the proof is very similar to what we have seen earlier. The solution of 3.13 changes slightly only for losses between $VaR_\beta(X)$ and $VaR_\alpha(X)$, in the sense that $R_1(\cdot)$ reaches its upper level as late as possible, but not before $R_2(\cdot)$ does. These are illustrated in Figure 3.4, and could be found via Proposition 3.2.

The mathematical formulation of the risk allocation is then given by

$$R_1^*[X;\underline{\xi}] := \begin{cases} \big(X - VaR_\beta(X) + \xi_{11}\big)_+ \wedge \xi_{11}, X \leq VaR_\alpha(X) - (\xi_{12}-\xi_{11}) - (\xi_{22}-\xi_{21}) \\ \big(X - VaR_\alpha(X) + \xi_{22} - \xi_{21} + \xi_{12}\big) \wedge \xi_{12}, \text{ otherwise} \end{cases}$$

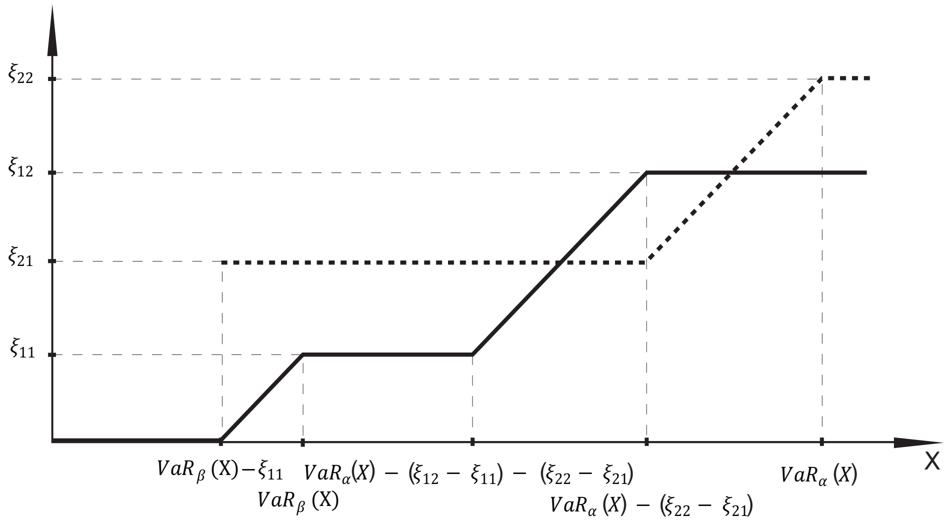

FIGURE 3.4. The construction of $R_1^*[X;\underline{\xi}]$ (solid line) and $R_2^*[X;\underline{\xi}]$ (dotted line) in Theorem 3.2 when $\rho* < \beta$.

and

$$R_2^*[X;\underline{\xi}] := \begin{cases} h_3^*\big(X;\underline{\xi}\big) \wedge \xi_{21}, X \leq VaR_\alpha(X) - (\xi_{22} - \xi_{21}) \\ \big(X - VaR_\alpha(X) + \xi_{22}\big) \wedge \xi_{22}, \text{ otherwise} \end{cases},$$

where $h_3^*(\cdot;\underline{\xi})$ is a non-decreasing Lipschitz function with unit constant such that $h_3^*(0;\underline{\xi}) = 0$ and $h_3^*\big(VaR_\beta(X) - \xi_{11};\underline{\xi}\big) = \xi_{21}$.

One may easily find that the second stage optimisation problem is then given by

$$\begin{cases} \min_{\mathcal{E}_1} VaR_\alpha(X) - \xi_{12} - \xi_{22} + \xi_{21} + \dfrac{1}{1-\beta} \displaystyle\int_{VaR_\alpha(X)-(\xi_{22}-\xi_{21})}^{VaR_\alpha(X)} \bar{F}(x)\,dx \\ \qquad +(1+\rho)\left(\displaystyle\int_{VaR_\beta(X)-\xi_{11}}^{VaR_\beta(X)} \bar{F}(x)\,dx + \int_{VaR_\alpha(X)-(\xi_{12}-\xi_{11})-(\xi_{22}-\xi_{21})}^{VaR_\alpha(X)-(\xi_{22}-\xi_{21})} \bar{F}(x)\,dx \right) \end{cases}$$

All other steps follow by taking appropriate partial derivatives in the same order as in the proof of the previous case. The proof is now complete. □

4. Robustness issues

In this section, we briefly investigate and discuss some robust statistical approaches to identify the optimal reinsurance arrangement. Therefore, it is important to estimate the risk measures in a robust manner, which boils down to robustly estimating the p-quantiles of a distribution (for different values of p between 0 and 1).

For a distribution F and a general $0 < p < 1$, the p-quantile is denoted $Q_p = inf\,\{x \in \mathbb{R} | F(x) \geq p\}$ with $0 < p < 1$. For a given F, these quantiles can often be computed analytically. When the underlying distribution is unknown, we have to estimate it from the observed data set $X_n = \{x_1, x_2, \ldots, x_n\}$. The empirical distribution function has a jump of size $1/n$ at each of the n data points and is given by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I(x \leq x_i)$$

where $I(\cdot)$ is the indicator function. Naturally, the empirical quantile is $\hat{Q}_p = inf\{x | F_n(x) \geq p\}$.

Classical parametric statistical procedures work well if the underlying assumptions hold, but may become extremely unreliable if the shape of the true underlying model deviates from the assumed parametric model. A typical violation which can have a big influence on the result is the presence of outliers in the data. Outliers are observations that do not follow the pattern indicated by the majority of the data. Sometimes these outlying observations are errors resulting from data input mistakes, but other reasons may also explain their occurrence in real data, such as belonging to different populations or observations made under exceptional circumstances. Therefore, it is interesting to always investigate the outliers after detection. In practice one often tries to detect outliers using diagnostics starting from a classical fitting method. However, as classical statistical techniques are known to be extremely sensitive to outliers, it is possible that these outliers affect the estimates so strongly that the resulting fitted model does not allow detection of the deviating observations (*masking effect*). Additionally, some reliable data points might even appear as outliers (*swamping effect*). Since it is well-known that outliers can lead to very misleading results when applying traditional statistical methods, robust alternatives have been developed. These techniques search for the model fitted by the majority of the data and hence are more robust against the possibility that the data contain one or more expected outliers. This has been the starting point of the robust statistical literature for the past 40 years, from the initial work in the univariate setting to recent work for high-dimensional data analysis (see e.g. Huber, 1981, Hampel et al., 1986 and Maronna et al., 2006 for more information). Recently, outlier detection and robustness have also become important in various financial applications such as asset allocation models (e.g. Welsch and Zhou, 2007), interest-rate models (e.g. Czellar et al., 2007), ruin probabilities (e.g. Loisel et al., 2008), time series modelling (e.g. Croux et al., 2010) and claims reserving (e.g. Verdonck and Debruyne, 2011). Therefore, robust methods are helpful to gain insight in the data and this knowledge often yields a significant improvement to the classical techniques. When there are no outliers in the data, the classical robust methods yield similar results, so the deviation between the two fits can be used to correctly identify outliers.

When examining the obtained closed formulas, it is clear that robust estimation of the risk measure will make the methodology more robust, and hence the goal of this section is to robustly estimate the p-quantiles of a data set (for different values of $0 < p < 1$). However, many univariate robust estimators are based purely on quantiles. Examples include the median instead of the classical mean for location and the interquartile range (IQR) rather than the classical standard deviation for scale. A popular tool to measure robustness is the breakdown point, which is the smallest amount of contamination that may cause an estimator to take on arbitrarily large values. Both classical estimators, mean and variance, have a breakdown point of 0%, whereas the median and IQR, respectively have a breakdown point of 50% and 25%. It is clear that the p-quantile estimator has a breakdown point of $100(1-p)$%. For example, for the 0.75-quantile, we have to alter (at least) the 25% largest observations of the data set to break down the estimator (if we let all these observations go to infinity, then the estimated quantile will also go to infinity). Therefore, if we only need to calculate p-quantiles for $p < 0.75$, robustness would probably not be an important issue, since most real data sets contain less than 25% outliers. Due to the fact that we are also interested in p-quantiles for $p > 0.75$ and even $p > 0.95$, a smaller amount of outliers are already able to significantly influence the result. Note that extremely small observations have no impact on this estimator, since only outliers in the right tail of the distribution have an effect on the estimation of the p-quantile.

Recently, robustness and sensitivity analysis of risk measures have been studied by Cont et al. (2010) and Kou et al. (2012). They also emphasize the importance of robustness of risk estimators in practice. Cont et al. (2010) have proposed an alternative for ES, namely the *Truncated Tail Value-at-Risk (TrTVaR)*, which is defined as follows:

$$TrTVaR_{\alpha_1,\alpha_2}(Z) := \frac{1}{\alpha_2-\alpha_1}\int_{\alpha_1}^{\alpha_2} VaR_s(Z)\,ds = \frac{(1-\alpha_1)ES_{\alpha_1}(Z)-(1-\alpha_2)ES_{\alpha_2}(Z)}{\alpha_2-\alpha_1}, \quad (4.1)$$

and is simply the average of VaR levels across a range of loss probabilities. The robustness and efficiency properties depend on the choice for $0 < \alpha_1 < \alpha_2 < 1$ (the higher α_2, the more efficient but also the less robust the estimator will be). It can easily be seen that this measure incorporates both the VaR and the ES risk measure. It is well-known that ES is a convex risk measure, while VaR is not convex (see for example, Acerbi and Tasche, 2002 and Denuit *et al.*, 2005). TrTVaR represents a compromise between VaR and ES, in the sense that only a part of the tail behaviour is measured by it, and as expected, it has similar properties to VaR. That is, TrTVaR is a non-convex and robust risk measure (see Cont *et al.*, 2010), and has the advantage of being more tail sensitive than VaR. Since it might also be interesting to give explicit derivations when the insurer's risk position is evaluated by this more robust measure, we provide in Theorem 4.1 the optimal risk allocation whenever the insurer decision is based on the TrTVaR risk measure. It is further assumed that $1+\rho < 1/\big(\alpha_2-\alpha_1\big)$, which represents the most plausible scenario as α_1 and α_2 are expected to have values close to 1.

Theorem 4.1. *If* $1+\rho < 1/\big(\alpha_2-\alpha_1\big)$, *then the TrTVaR-based optimal reinsurance contract is given by*

$$R^*[X] = \begin{cases} \big(X - VaR_{\alpha^{**}}(X)\big)_+ \wedge \big(VaR_{\alpha^*}(X) - VaR_{\alpha^{**}}(X)\big)_+, & \alpha^* > \alpha_1 \\ \big(X - VaR_{\rho^*}(X)\big)_+ \wedge \big(VaR_{\alpha_1}(X) - VaR_{\rho^*}(X)\big)_+, & \alpha^* < \alpha_1 \end{cases}, \tag{4.2}$$

where $\alpha^* = 1 - \frac{1-\alpha_2}{1-(1+\rho)\big(\alpha_2-\alpha_1\big)}$ *and* $\alpha^{**} = \min(\rho^*, \alpha_1)$.

Proof. Keeping in mind equation 4.1, the function that needs to be minimised over $\mathcal{F}$ is given by

$$\begin{aligned} TrTVaR_{\alpha_1,\alpha_2}\Big(\mathbf{L}\big(R[X]\big)\Big) &= TrTVaR_{\alpha_1,\alpha_2}(X) - TrTVaR_{\alpha_1,\alpha_2}\big(R[X]\big) + (1+\rho)\mathbf{E}\big(R[X]\big) \\ &= TrTVaR_{\alpha_1,\alpha_2}(X) + (1+\rho)\int_0^{\alpha_1} R\big(VaR_s(X)\big)\,ds \\ &+ \left(1+\rho-\frac{1}{\alpha_2-\alpha_1}\right)\int_{\alpha_1}^{\alpha_2} R\big(VaR_s(X)\big)\,ds + (1+\rho)\int_{\alpha_2}^{1} R\big(VaR_s(X)\big)\,ds. \end{aligned}$$

As before, the above is solved via a two-stage procedure, where the first step becomes

$$\begin{cases} \min\limits_{R\in\mathcal{F}}\ (1+\rho)\left(\int_0^{\alpha_1} + \int_{\alpha_2}^{1}\right) R\big(VaR_s(X)\big)\,ds + \left(1+\rho-\frac{1}{\alpha_2-\alpha_1}\right)\int_{\alpha_1}^{\alpha_2} R\big(VaR_s(X)\big)\,ds \\ \quad \text{subject to}\quad R\big(VaR_{\alpha_1}(X)\big) = \xi_1,\, R\big(VaR_{\alpha_2}(X)\big) = \xi_2. \end{cases} \tag{4.3}$$

Note that $(\xi_1, \xi_2) \in \mathcal{C}_1$ is a vector of constants with

$$\mathcal{C}_1 := \Big\{0 \le \xi_2 - \xi_1 \le VaR_{\alpha_2}(X) - VaR_{\alpha_1}(X), 0 \le \xi_1 \le VaR_{\alpha_1}(X), 0 \le \xi_2 \le VaR_{\alpha_2}(X)\Big\}.$$

Now, $1+\rho-\frac{1}{\alpha_2-\alpha_1}$ is assumed to be negative. Similar arguments to the one used in the proof of Theorems 2.1 and 2.2 show that 4.3 is solved by $R_1^*[X;\xi_1,\xi_2] := \big(X - VaR_{\alpha_1}(X) + \xi_1\big)_+ \wedge \xi_2$. Equations 1.1 and 4.1, and the fact that

$$\mathbf{E}\big(R_1^*[X;\xi_1,\xi_2]\big) = \int_0^{\xi_2} \Pr\left(R_1^*[X;\xi_1,\xi_2] > x\right) dx = \int_{VaR_{\alpha_1}(X)-\xi_1}^{VaR_{\alpha_1}(X)+\xi_2-\xi_1} \Pr(X > x)\,dx,$$

show that the second step is to minimise over $\mathcal{C}_1$

$$\begin{aligned} H_4(\xi_1,\xi_2) &:= (1+\rho)\int_{VaR_{\alpha_1}(X)-\xi_1}^{VaR_{\alpha_1}(X)+\xi_2-\xi_1} \Pr(X > x)\,dx \\ &\quad - \frac{1}{\alpha_2-\alpha_1}\left((1-\alpha_1)\xi_1 - (1-\alpha_2)\xi_2 + \int_{VaR_{\alpha_1}(X)}^{VaR_{\alpha_1}(X)+\xi_2-\xi_1} \Pr(X > x)\,dx\right). \end{aligned}$$

Clearly, $\frac{dH_4}{d\xi_2} = \frac{1-\alpha_2}{\alpha_2-\alpha_1} + \Big(1+\rho-\frac{1}{\alpha_2-\alpha_1}\Big)\bar{F}\big(VaR_{\alpha_1}(X)+\xi_2-\xi_1\big)$. The latter is negative if and only if $F\big(VaR_{\alpha_1}(X)+\xi_2-\xi_1\big) < \alpha^*$, which is equivalent to $VaR_{\alpha_1}(X)+\xi_2-\xi_1 < VaR_{\alpha^*}(X)$.

Note that $\alpha^* < \alpha_2$ always holds. Let us assume first that $\alpha^* > \alpha_1$. Keeping in mind that $\xi_2 \in \big[\xi_1, \xi_1 + VaR_{\alpha_2}(X) - VaR_{\alpha_1}(X)\big]$, then for any fixed $\xi_1 \in \big[0, VaR_{\alpha_1}(X)\big]$ we have that $\frac{dH_4}{d\xi_2} < 0$ if $\xi_2 \in \big[\xi_1, \xi_1 + VaR_{\alpha^*}(X) - VaR_{\alpha_1}(X)\big)$, and $\frac{dH_4}{d\xi_2} \geq 0$ if

$$\xi_2 \in \big[\xi_1 + VaR_{\alpha^*}(X) - VaR_{\alpha_1}(X), \xi_1 + VaR_{\alpha_2}(X) - VaR_{\alpha_1}(X)\big].$$

Thus,

$$H_4(\xi_1,\xi_2) \geq H_4\big(\xi_1, \xi_1 + VaR_{\alpha^*}(X) - VaR_{\alpha_1}(X)\big) = (1+\rho)\int_{VaR_{\alpha_1}(X)-\xi_1}^{VaR_{\alpha^*}(X)} \Pr(X>x)\,dx - \xi_1 + K_1,$$

where K_1 is a constant with respect to ξ_1. Taking the derivative with respect to ξ_1 of the right hand side function from above, one may recover 4.2 for this case, by following similar steps as used in the proof of Theorem 2.1.

It only remains to show the $\alpha^* < \alpha_1$ case. Thus, $\frac{dH_4}{d\xi_2} > 0$ holds for any fixed $\xi_1 \in \big[0, VaR_{\alpha_1}(X)\big]$, and any $\xi_2 \in \big[\xi_1, \xi_1 + VaR_{\alpha_2}(X) - VaR_{\alpha_1}(X)\big]$. Consequently,

$$H_4(\xi_1,\xi_2) \geq H_4\big(\xi_1,\xi_1\big) = (1+\rho)\int_{VaR_{\alpha_1}(X)-\xi_1}^{VaR_{\alpha_2}(X)} -\xi_1 + K_2,$$

where K_2 is a constant with respect to ξ_1. As before, one may find the global optimal solution for this case. The proof is now complete. □

It is expected that the limiting cases from Theorem 4.1 can be used to recover the optimal VaR and ES-based decisions found in Theorems 2.1 and 2.2, which are further explained. The VaR case is obtained as $\alpha_1 \nearrow \alpha_2$, which implies that $\alpha^* \nearrow \alpha_2$. It is simple to show that if α_1 is sufficiently close to α_2, then $\alpha^* < \alpha_1$ ($\alpha^* > \alpha_1$) as long as $\rho^* \geq \alpha_2$ ($\rho^* < \alpha_2$), which in turn makes the optimal reinsurance contract from 4.2 to be $R^*[X] = \big(X - VaR_{\rho^*}(X)\big)_+ \wedge \big(VaR_{\alpha_2}(X) - VaR_{\rho^*}(X)\big)_+$. Clearly, the latter recovers the VaR result from Theorem 2.1. The ES case is obtained when $\alpha_2 \nearrow 1$, which yields that $\alpha^* \nearrow 1$. Thus, $\alpha^* > \alpha_1$ and 4.2 becomes $R^*[X] = \big(X - VaR_{\rho^*}(X)\big)_+$, as long as $1+\rho < 1/(1-\alpha_1)$, i.e. $\rho^* < \alpha_1$, which aligns with our findings from Theorem 2.2.

It is interesting to point out that an ideal contract from Theorem 4.1 does not always require for layers higher than $VaR_{\alpha_1}(X)$ to be covered by the insurance company, as it has been seen for VaR-based optimal arrangement, even though the insurer risk is measured by quantifying a portion of the worst $1-\alpha_1$ events. For example, the reinsurer covers in full the $[VaR_{\rho^*}(X), VaR_{\alpha^*}(X)]$ layer, whenever $\rho^* \leq \alpha_1 < \alpha^* < \alpha_2$. Further, if $\rho = 1$, $\alpha = 95\%$, $\alpha_1 = 93\%$ and $\alpha_2 = 97\%$, then $\rho^* = \alpha^{**} = 50\%$ and $\alpha^* = 96.7\%$. Consequently, the optimal 95% VaR decision makes the reinsurer to cover the $[VaR_{50\%}(X), VaR_{95\%}(X)]$ layer, while the optimal $[93\%, 97\%]$ TrTVaR reinsurance arrangement allows the reinsurer to pay only the $[VaR_{50\%}(X), VaR_{96.7\%}(X)]$ layer. The given example makes the TrTVaR to be blind to the worst 3% events, but the reinsurer pays the worst 3.3% events, which includes some of the extreme events embedded in the $[93\%, 97\%]$ TrTVaR risk measure. The latter can be explained by the more complex structure of the TrTVaR risk measure.

Actually, the TrTVaR estimator simply deletes a certain percentage, $1-\alpha_2$, of observations before computing the ES in the usual way. A drawback of this methodology is that the value α_2 has to be chosen beforehand and does not depend on the observed data. Hence, the proportion of large observations that is removed is specified before computing the risk measure on the remaining observations. If a prudent value for α_2 is taken, and in reality there are no outliers in the data (or only a very few), the result will not be very efficient, because a significant amount of "good" observations is not taken into account for estimation. To overcome this problem, one could also follow another approach, namely empirically determining the amount of outliers in the data. Therefore, one might try to automatically screen the data for outliers and remove them by using robust methods.

In general, one will try to estimate the distribution of the data before computing the p-quantiles. Since we deal with univariate data, it might certainly be interesting to start with some exploratory data analysis and visualise the data by constructing a few plots. To have an idea about the underlying distribution of the data, it is useful to construct QQ plots for possible theoretical distributions. The most common form is the normal QQ plot, which is used to test graphically whether the observed data stems from a normal distribution. The shape of the QQ plot can also be very useful to highlight distributional asymmetry, outliers, heavy tails, multi-modality or other interesting data characteristics. Besides the normal QQ plot, it is also possible to construct QQ plots for other reference distributions (see for example the `qqplot` command in R which creates a QQ plot for most theoretical distributions). One could also use the classical boxplot in order to detect outliers, which has been introduced by Tukey (1977). However, the classical boxplot rejection rule inherently assumes normality of the data, and therefore, it usually classifies too many points as outlying when the data are skewed, which is more the rule than an exception in financial data. To overcome this drawback, Hubert and Vandervieren (2008) have proposed the adjusted boxplot, which includes a robust measure of skewness and gives a more accurate representation of possible outliers.

One could also opt for kernel density estimation methods (see Silverman, 1986 for an introductory text) and hence, no prior belief about the distribution is needed. Fernholz (1997) has already shown that smoothing the empirical distribution function with an appropriate kernel and bandwidth can reduce the variance and mean squared error of some quantile-based estimators. Recently, Hubert et al. (2012) have constructed methods for reducing the variance and mean squared error of different univariate robust estimators that are based on quantiles, by implementing a kernel smoothed distribution function instead of the empirical distribution function. This smoothing procedure also improves the ability to detect outliers with the adjusted boxplot. A range of kernel functions have to be chosen, but popular choices are the Gaussian basis functions. Another important issue is the selection of the bandwidth of the kernel, which exhibits a strong influence on the resulting estimate. Optimal choices for this bandwidth exist (see for example, Sheather and Jones, 1991, and Jones et al., 1996). A robust bandwidth selection procedure is also proposed in Hubert et al. (2012). The kernel density estimation plot is helpful to notice multi-modality, outliers and other data anomalies.

After visual data inspection, it is plausible to believe that a good understanding about the distribution is achieved. We then propose to use the following strategy to obtain robust estimates for the quantile estimation.

(1) Robustly estimate the parameters of the distribution. For most common distributions, the following references are useful:
   - Exponential distribution (see Gather, 1986 and Gather and Schultze, 1999).
   - Gamma distribution (see Marazzi and Ruffieux, 1996 and 1999, and Marazzi and Barbati 2003).
   - LogNormal distributon (see Serfling, 2002).
   - Pareto distribution (see Brazauskas and Serfling, 2000 and 2003).
   - Symmetric distributions (see Maronna et al., 2006).
   - Weibull distribution (see Boudt et al., 2011).

   Many of these techniques are already implemented in R.

(2) Before computing the required p-quantiles with the obtained robust parameter estimates, one should check whether the assumption about the distribution is appropriate. This can be done by using a goodness-of-fit test. For some of the above-mentioned distributions, there exist particular goodness-of-fit tests (see for example, Rizzo, 2009 for the Pareto distribution). Otherwise, the Kolmogorov-Smirnov, Cramér-von Mises or Anderson-Darling test on the regular points are well-known alternatives. Most methods will automatically indicate which

observations are the non-outlying ones. Otherwise, we propose to theoretically calculate the 99%-quantile with the robust parameter estimates and flag all observations larger than this value as outlying.

(3) If the goodness-of-fit test is not rejected, the required theoretical quantiles are obtained. Otherwise, one should try the entire procedure with another distribution.

When calculating p-quantiles for very large p (say $p \geq 0.95$), it might be better to follow an approach which is based on the idea of combining robust statistics and extreme value statistics. At first sight, it might look inappropriate to combine these two research fields, since it is hard to decide whether one or a group of very large observations are outliers or not when dealing with heavy-tailed distributions. However, it is already shown that robustness may play an important role in extreme value theory (see e.g. Dell'Aquila and Embrechts, 2006, Vandewalle et al., 2007, and Hubert et al., 2012). For Pareto-type distributions, Hubert et al. (2012) have proposed a robust estimator of the extreme value index, because it is shown that classical estimators, such as the Hill estimator, tend to overestimate this parameter in the presence of outliers. Based on this robust estimator, they have constructed a diagnostic tool to detect automatically the observations that have an unusual large influence on the Hill estimator. This method also allows the robust estimation of large p-quantiles, as described by Weissman (1978).

5. Examples and Numerical Results

This section provides a numerical analysis related to our main results described in Theorems 2.1 and 3.1. We start by enumerating some of the well known distorted premium principle examples.

i) Dual-power function principle: $g(t) = 1 - (1-t)^{\beta}$, $\beta > 1$;
ii) ES principle: $g(t) = \min(t/(1-\beta), 1)$, $0 < \beta < 1$;
iii) Gini principle: $g(t) = (1+\beta)t - \beta t^2$, $0 < \beta \leq 1$;
iv) Proportional hazard transform (PHT) principle: $g(t) = t^{1-\beta}$, $0 < \beta < 1$;
v) Wang Transform: $g(t) = F_N\big(F_N^{-1}(t) + \lambda\big)$, $\lambda > 0$,

where $F_N(\cdot)$ and $F_N^{-1}(\cdot)$ represent the standard normal cumulative distribution function and its inverse respectively. The Φ's functions can be easily calculated, and therefore we only state the form for the Wang Transform premium, which is

$$\Phi(s) = \exp\left\{\lambda F_N^{-1}(s) - \frac{\lambda^2}{2}\right\},$$

(for more details, see Jones and Zitikis, 2003 and 2007). Note that, except for ES, all other examples are potential candidates for Theorem 3.1.

We assume three particular distributions for the total loss, X, and all of them having an expected loss amount of 1,000.

a) Exponential ($\lambda = 1,000$);
b) LogNormal ($\mu = 6.4; \sigma = 1.00773$);
c) Pareto ($\theta_1 = 3, \theta_2 = 2,000$) with the survival function given by $\bar{F}(x) = \left(1 + \frac{x}{\theta_2}\right)^{-\theta_1}, x \geq 0$.

It is also assumed that the first reinsurer prefers an expected value principle with $\rho = 1$, while the other reinsurer opts out for the PHT principle with $\beta \in \{0.5, 0.6\}$. Thus, $\rho^* = 0.5$ and

$$(s^*, s^{**}) = \begin{cases} (93.75\%, 75\%), & \beta = 0.5 \\ (93.16\%, 68.50\%), & \beta = 0.6 \end{cases}.$$

The insurer sets its confidence level at $\alpha = 99.5\%$, which aligns with the capital requirements designed within the Solvency II, that applies to any insurance or reinsurance company that operates in the European Union. We are now able to provide some numerical results for Theorem 3.1 where the diversification gain takes place. Recall that the risk allocation has the following form

$$\begin{cases} I^*[X] &= \big(X - VaR_\alpha(X)\big)_+ \\ R_1^*[X] &= \big(X - VaR_{s^{**}}(X)\big)_+ \wedge \big(VaR_\alpha(X) - VaR_{s^{**}}(X)\big) \\ R_2^*[X] &= X \wedge VaR_{s^{**}}(X) \end{cases},$$

which is now compared with the non-diversified scenario that is allocated as in Theorem 2.1

$$\begin{cases} I^*[X] &= X \wedge VaR_{\rho^*}(X) + \big(X - VaR_\alpha(X)\big)_+ \\ R_1^*[X] &= \big(X - VaR_{\rho^*}(X)\big)_+ \wedge \big(VaR_\alpha(X) - VaR_{\rho^*}(X)\big) \\ R_2^*[X] &= 0 \end{cases}.$$

The corresponding insurer total VaR-based risk is then quantified by 3.7 and 2.2, respectively.

Note that the premium charged by the second reinsurance company increases as β increases, and the diversification gain should decrease as well. Our numerical results capture this pattern and are given in Table 5.1. The values that appear in the parentheses reflect the diversification gain

Table 5.1. VaR of the Insurer Total loss at 0.995 level: Theorem 3.1 setting

	Exponential	*LogNormal*	*Pareto*
One Reinsurer	1,683.15	1,650.24	1,721.28
Two Reinsurers $\beta = 0.5$	1,490 (11.48%)	1,500.75 (9.06%)	1,560.42 (9.35%)
Two Reinsurers $\beta = 0.6$	1,545.06 (8.20%)	1,544.92 (6.38%)	1,608.65 (6.54%)

corresponding to each setting. It is quite apparent that the insurer may benefit from clever risk allocation. One may find it a tad intriguing that for $\beta = 0.6$, the Exponential loss replicates a higher total risk than the LogNormal scenario. It is very true that the LogNormal distribution has a heavier tail than the Exponential one, but they are not ordered stochastically over the entire loss spectrum, which explains this potential misunderstanding.

We have displayed the optimal risk allocation only for $\alpha = 0.995$, and further explanations underpin our choice. It is not surprising that the VaR corresponding to the insurer total loss at various levels from 3.7 increases as the targeted level $\alpha \in [s^{**}, 1]$ becomes more conservative. Therefore, it is futile to analyse the cedent efficient frontier from this perspective. No matter what risk level is used, the insurer is facing the same capital requirement, and it is at a 99.5% level for any insurance or reinsurance company that operates in the European Union (EU), which will soon become mandatory. Under this framework, we argue that the global optimality is attained if the cedent chooses the confidence level requested by Solvency II.

The risk capital generated by any insurance company business is usually sustained by its very own shareholders, who in turn expect a capital return. The latter aligns with the well known *Cost of Capital* (CoC) approach, which is especially used whenever traditional evaluations available within the financial markets are not applicable for quantifying the cost of transferring the liabilities to another counter-party. Specifically, the CoC evaluation of a generic liability, Z, is evaluated by $(1-\delta)\mathbf{E}(Z) + \delta VaR_{0.995}(Z)$ in our setting. The value of δ is usually set as the sum of the risk free interest and a risk premium naturally expected by the shareholders. Thus, any risk allocation from the efficient frontier described in 3.6 has a liability evaluated via the CoC method given by

$$\begin{aligned} CoC(\gamma;0.995) &:= (1-\delta)\mathbf{E}\Big(\mathbf{L}\big(R_1^*[X],R_2^*[X]\big)\Big)+\delta VaR_{0.995}\Big(\mathbf{L}\big(R_1^*[X],R_2^*[X]\big)\Big)\\ &= (1-\delta)\int_{VaR_\gamma(X)}^{x_F}\bar{F}(x)\,dx+\delta\Big(VaR_{0.995}(X)-VaR_\gamma(X)\Big)_+\\ &\quad +(1+\rho)\int_{VaR_{s^{**}}(X)}^{VaR_\gamma(X)}\bar{F}(x)\,dx+\int_0^{VaR_{s^{**}}(X)}g\big(\bar{F}(x)\big)\,dx. \end{aligned}$$

for any targeted confidence level $\gamma\in[s^{**},1]$. The latter is illustrated in Figure 5.1 for all three loss distributions discussed in the beginning of the current section whenever $\delta=10\%$. It can be

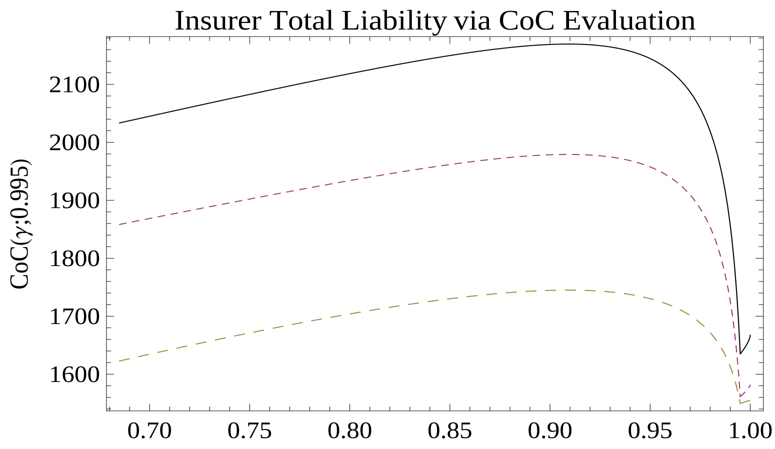

FIGURE 5.1. CoC evaluation of the insurer total liability for various values of γ and $\beta=0.6$ for respectively, Pareto (solid), LogNormal (short dash) and Exponential (long dash) loss distributions.

observed that the cedent would prefer to choose from its efficient frontier the risk allocation that is optimised for the 99.5% confidence level. In addition, the most expensive choice is attained in the case in which the insurance company prefers to minimise its total VaR-based risk at $\gamma=90.9\%$. This underpins the fact that the diversification through reinsurance is sensible to the regulatory capital requirements that the insurance company should meet.

TABLE 5.2. VaR of the Insurer Total loss at 0.995 level: Theorem 3.2 setting

	Exponential	*LogNormal*	*Pareto*
One Reinsurer	1,683.15	1,650.24	1,721.28
Two Reinsurers $\beta=0.4$	1,502.49 (10.73%)	1,463.79 (11.30%)	1,508.16 (12.38%)
Two Reinsurers $\beta=0.3$	1,349.53 (19.82%)	1,313.10 (20.43%)	1,336.97 (22.33%)
Two Reinsurers $\beta=0.2$	1,216.89 (27.70%)	1,187.14 (28.06%)	1,195.10 (30.57%)
Two Reinsurers $\beta=0.1$	1,099.80 (34.66%)	1,078.76 (34.63%)	1,074.74 (37.56%)

Finally, this section is concluded with some numerical examples that show the diversification gain provided by the strategy given in Theorem 3.2 as compared to Theorem 2.1. That is, we assume the same particular parametric loss distribution assumptions as used earlier. As before, the first reinsurer chooses a security loading factor $\rho = 1$. In addition, the second reinsurer charges its premium based on the average worst $1-\beta$ events, where $\beta \in \{0.1, 0.2, 0.3, 0.4\}$, i.e. $\beta < \rho^* = 0.5$ always holds. All results are summarised in Table 5.2. Clearly, a reduction in the value of β reduces the level of reinsurance premium, and therefore, the overall risk decreases, and the diversification gain (given in parentheses) increases.

Acknowledgements

The authors are grateful to two anonymous referees for a number of insightful questions and suggestions that have guided our work on revising the paper. The second author acknowledges his research support by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

References

Acerbi, C. 2002. "Spectral Measure of Risks: A Coherent Representation of Subjective Risk Aversion," *Journal of Banking & Finance*, 26(7), 1505-1518.

Acerbi, C. and Tasche, D. 2002. "On the Coherence of Expected Shortfall," *Journal of Banking & Finance*, 26(7), 1487-1503.

Arrow, K.J. 1963. "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, 53(5), 941–973.

Bellini,F. and Rosazza Gianin, E. 2012. "Haezendonck-Goovaerts Risk Measures and Orlicz Quantiles," *Insurance: Mathematics and Economics*, 51(1), 107–114.

Borch, K., 1960. "An Attempt to Determine the Optimum Amount of Stop Loss Reinsurance," *Transactions of the 16th International Congress of Actuaries*, vol. I, 597–610.

Boudt, K., Caliskan, D., and Croux, C. 2011. "Robust Explicit Estimators for Weibull Parameters," *Metrika*, 73(2), 187–209.

Brazauskas, V. and Serfling, R. 2000. "Robust Estimation of Tail Parameters for Two-parameter Pareto and Exponential Models via Generalized Quantile Statistics," *Extremes*, 3(3), 231–249.

Brazauskas, V. and Serfling, R. 2003. "Favorable Estimators for Fitting Pareto models: A Study using Goodness-of-fit Measures with Actual Data," *ASTIN Bulletin*, 33(2), 365–381.

Cai, J., Tan, K.S., Weng, C. and Zhang, Y. 2008. "Optimal Reinsurance under VaR and CTE Risk Measures," *Insurance: Mathematics and Economics*, 43(1), 185–196.

Centeno, M.L. and Guerra, M. 2010. "The Optimal Reinsurance Strategy - The Individual Claim Case," *Insurance: Mathematics and Economics*, 46(3), 450–460.

Cheung, K.C. 2010. "Optimal Reinsurer Revisited - A Geometric Approach ," *Astin Bulletin*, 40(1), 221–239.

Chi, Y. and Tan, K.S. 2011. "Optimal Reinsurance under VaR and CVaR Risk Measures: A Simplified Approach," *Astin Bulletin*, 42(2), 487–509.

Cont, R., Deguest, R. and Scandalo, G. 2010. "Robustness and Sensitivity Analysis of Risk Measurement Procedures," *Quantitative Finance*, 10(6), 593–606.

Croux, C., Gelper, G., and Mahieu, K. 2010. "Robust Exponential Smoothing of Multivariate Time Series," *Computational Statistics and Data Analysis*, 54, 2999–3006.

Czellar, V., Karolyi, G.A. and Ronchetti, E. 2007. "Indirect Robust Estimation of the Short-term Interest Rate Process," *Journal of Empirical Finance*, 14, 546–563.

Dell'Aquila, R., Embrechts, P. 2006. "Extremes and Robustness: A Contradiction?," *Financial Markets and Portfolio Management*, 20, 103–118.

Denuit, M., Dhaene, J., Goovaerts, M.J. and Kaas, R. 2005. *Actuarial Theory for Dependent Risks: Measures, Orders and Models*, Wiley, Chichester.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas R. and Vyncke, D. 2002 a. "The Concept of Comonotonicity in Actuarial Science and Finance: Theory," *Insurance: Mathematics and Economics*, 31(1), 3–33.

Dhaene, J., Denuit, M., Goovaerts, M.J., Kaas R. and Vyncke, D. 2002 b. "The Concept of Comonotonicity in Actuarial Science and Finance: Applications," *Insurance: Mathematics and Economics*, 31(2), 133–161.

Dhaene, J., Kukush, A., Linders, D. and Qihe. 2012. "Remarks on Quantiles and Distortion Risk Measures," *European Actuarial Journal*, 2(2), 319–328.

Fernholz, L.T. 1997. "Reducing the Variance by Smoothing," *Journal of Statistical Planning and Inference*, 57(1), 29-38.

Gather, U. 1986. "Robust Estimation of the Mean of the Exponential Distribution in Outlier Situations," *Communications in Statistics A*, 15, 2323–2345.

Gather, U., and Schultze, V. 1999. "Robust Estimation of Scale of an Exponential Distribution," *Statistica Neerlandica*, 53, 327–341.

Goovaerts, M.J., Linders, D., Van Weert, K. and Tank, F. 2012. "On the Interplay Between Distortion, Mean Value and Haezendonck-Goovaerts Risk Measures," *Insurance: Mathematics and Economics*, 51(1), 10–18.

Goovaerts, M.J., Kaas, R., Dhaene, J. and Tang, Q. 2004. "Some New Classes of Consistent Risk Measures," *Insurance: Mathematics and Economics*, 34(3), 505–516.

Guerra, M. and Centeno, M.L. 2008. "Optimal Reinsurance Policy: The Adjustment Coefficient and the Expected Utility Criteria," *Insurance: Mathematics and Economics*, 42(2), 529–539.

Guerra, M. and Centeno, M.L. 2010. "Optimal Reinsurance for Variance Related Premium Calculation Principles," *Astin Bulletin*, 40(1), 97–121.

Haezendonck, J. and Goovaerts, M.J. 1982. "A New Premium Calculation Principle Based on Orlicz Norms," *Insurance: Mathematics and Economics*, 1(1), 41–53.

Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. 1986. *Robust Statistics: The Approach Based on Influence Functions*. New York: Wiley.

Hartigan, J.A. and Hartigan, P.M. 1985. "The Dip Test of Unimodality," *Annals of Statistics*, 13(1), 70–84.

Huber, P.J. 1981. *Robust Statistics*. New York: Wiley.

Hubert, M. and Vandervieren, E. 2008. "An Adjusted Boxplot for Skewed Distributions," *Computational Statistics and Data Analysis*, 52(12), 5186–5201.

Hubert, M., Dierckx, G. and Vanpaemel, D. 2012. "Detecting Influential Data Points for the Hill Estimator in Pareto-type Distributions," *Computational Statistics and Data Analysis*, in press.

Hürlimann, V. 2003. "Conditional Value-at-Risk Bounds for Compound Poisson Risks and a Normal Approximation," *Journal of Applied Mathematics*, Issue 3, 141-153.

Jones, M.C., Marron, J.S. and Sheather, S.J. 1996. "A Brief Survey of Bandwidth Selection for Density Estimation," *Journal of the American Statistical Association*, 91(433), 401–407.

Jones, B.L. and Zitikis, R. 2003 "Empirical Estimation of Risk Measures and Related Quantities ," *North American Actuarial Journal*, 7(4), 44–54.

Jones, B.L. and Zitikis, R. 2007 "Risk Measures, Distortion Parameters, and Their Empirical Estimation," *Insurance: Mathematics and Economics*, 41(2), 279–297.

Kaluszka, M., 2001. "Optimal Reinsurance under Mean-variance Premium Principles," *Insurance: Mathematics and Economics*, 28(1), 61–67.

Kaluszka, M., 2005. "Truncated Stop Loss as Optimal Reinsurance Agreement in One-period Models," *Astin Bulletin*, 35(2), 337–349.

Kaluszka, M. and Okolewski, A., 2008. "An Extension of Arrow's Result on Optimal Reinsurance Contract," *The Journal of Risk and Insurance*, 75(2), 275–288.

Kou, S.G., Peng, X. and Heyde, C. 2012. "External Risk Measures and Basel Accords," *Mathematics of Operations Research*, in press.

Loisel, S., Mazza, C., Rullire, D. 2008. “Robustness Analysis and Convergence of Empirical Finite-time Ruin Probabilities and Estimation Risk Solvency Margin,” *Insurance: Mathematics and Economics*, 42(2), 746–762.

Ludrovski, M. and Young, V.R. 2009. “Optimal Risk Sharing under Distorted Probabilities,” *Mathematics and Financial Economics*, 2(2), 87-105.

Marazzi, A. and Ruffieux, C. 1996. “Implementing M-estimators of the Gamma distribution,” *Robust Statistics, Data Analysis, and Computer Intensive Methods, in Honor of Peter Hubers 60th Birthday, H. Rieder (Ed.). Lecture Notes in Statistics, Springer*, 109, 277–297.

Marazzi, A. and Ruffieux, C. 1999. “The Truncated Mean of an Asymmetric Distribution,” *Computational Statistics and Data Analysis*, 32, 79-100.

Marazzi, A. and Barbati, G. 2003. “Robust Parametric Means of Asymmetric Distributions: Estimation and Testing,” *Estadistica*, 54(162-163), 47–72.

Maronna, R., Martin, D. and Yohai, V. 2006. Robust Statistics - Theory and Methods. New York: Wiley.

Rizzo, M.L. 2009. “New Goodness-of-Fit Tests for Pareto Distributions,” *Astin Bulletin*, 39(2), 691–715.

Serfling, R. 2002. “Efficient and Robust Fitting of LogNormal Distributions,” *North American Actuarial Journal*, 6, 95–109.

Sheather, S. J. and Jones, M. C. 1991. “A Reliable Data-based Bandwidth Selection Method for Kernel Density Estimation,” *Journal of the Royal Statistical Society series B*, 53, 683–690.

Silverman, B.W. 1986. *Density estimation for statistics and data analysis*, New York: Chapman and Hall.

Tukey, J.W. 1977. *Exploratory Data Analysis*, Reading, Massachusetts: Addison-Wesley.

Vandewalle, B., Beirlant, J., Christmann, A., Hubert, M. 2007. “A Robust Estimator for the Tail Index of Pareto-type Distributions,” *Computational Statistics and Data Analysis*, 51, 6252–6268.

Van Heerwaarden, A.E., Kaas, R. and Goovaerts, M.J. 1989. “Optimal Reinsurer in the Relation to Ordering of Risk,” *Insurance: Mathematics and Economics*, 8(1), 11–17.

Verdonck, T. and Debruyne, M. 2011. “The Influence of Individual Claims on the Chain-ladder Estimates: Analysis and Diagnostic Tool,” *Insurance: Mathematics and Economics*, 48(1), 85–98.

Verlaak, R. and Beirlant, J. 2003. “Optimal Reinsurance Programs: An Optimal Combination of Several Reinsurance Protections on a Heterogeneous Insurance Portfolio,” *Insurance: Mathematics and Economics*, 33(2), 381–403.

Wang, S. and Young, V.R. 1998. “Ordering Risks: Expected Utility Theory versus Yaari’s Dual Theory of risk,” *Insurance: Mathematics and Economics* 22(2), 145–161.

Weissman, I. 1978. “Estimation of Parameters and Large Quantiles Based on the k Largest Observations,” *Journal of the American Statistical Association*, 73, 812–815.

Welsch, R. and Zhou, X. 2007. “Application of Robust Statistics to Asset Allocation Models”, *REVSTAT - Statistical Journal*, 5(1), 97–114.

Young, V.R., 1999. “Optimal Insurance under Wang’s Premium Principle,” *Insurance: Mathematics and Economics* 25(2), 109–122.