

## Optimal Risk Transfer: A Numerical Optimisation Approach

ALEXANDRU V. ASIMIT<sup>1</sup>

*Cass Business School, City, University of London, London EC1Y 8TZ,  
United Kingdom. E-mail: asimit@city.ac.uk*

TAO GAO

*Department of Mathematical Sciences, University of Bath, Bath BA2 7AY, United Kingdom.  
E-mail: t.gao@bath.ac.uk*

JUNLEI HU

*Cass Business School, City, University of London, London EC1Y 8TZ,  
United Kingdom. E-mail: Junlei.Hu.1@city.ac.uk*

EUN-SEOK KIM

*School of Business and Management, Queen Mary, University of London,  
London E1 4NS, United Kingdom. E-mail: e.kim@qmul.ac.uk*

December 15, 2017

**Abstract:** Capital efficiency and asset/liability management are part of the Enterprise Risk Management Process of any insurance/reinsurance conglomerate and serve as quantitative methods to fulfill the strategic planning within an insurance organisation. There has been a considerable amount of work in this ample research field, but invariably one of the last questions is whether or not, numerically, the method is practically implementable, which is our main interest. The numerical issues are dependent upon the traits of the optimisation problem and therefore, we plan to focus on the optimal reinsurance design, which has been a very dynamic topic in the last decade. The existing literature is focused on finding closed-form solutions that are usually possible when economic, solvency, etc constraints are not included in the model. Including these constraints, the optimal contract can only be found numerically. The efficiency of these methods is extremely good for some well-behaved convex problems, such as the Second-Order Conic Problems. Specific numerical solutions are provided in order to better explain the advantages of appropriate numerical optimisation methods chosen to solve various risk transfer problems. The stability issues are also investigated together with a case study performed for an insurance group that aims capital efficiency across the entire organisation.

*Keywords and phrases:* Linear Programming, Optimal Reinsurance/Risk Transfer, Risk Measure, Second-Order Conic Programming.

---

<sup>1</sup>Corresponding author. Phone: +44(0)2070405282. Fax: +44(0)2070408572.

## 1. INTRODUCTION

Various actuarial problems involve decision-making procedures that evaluate the most favourable risk position of an insurance company. For example, capital efficiency and asset/liability management are part of the Enterprise Risk Management Process of any insurance/reinsurance conglomerate and serve as quantitative methods to fulfill the strategic planning within the organisation. The decision-makers are prone to combine expert judgment with core quantitative methods, which involve numerical optimisation and often, intensive computing skills. Therefore, many optimisation problems are not practically implementable in a straightforward manner to practitioners and academics that are not operation research inclined. Unfortunately, numerical issues are anecdotally disregarded and for this reason, we aim to implement optimisation algorithms that are hardly accessible to non-specialists in this field. In order to better communicate the advantages and caveats of possible solutions, we plan to focus on optimal risk transfer problems, but the numerical methods are transferrable skills when implementing other actuarial problems.

Consider a two-player insurance setting where the first player is the risk holder who transfers a portion of its risk to the second player. At the same time, the second player charges the first player to cover its cost of transfer. This setting includes a very common risk transfer, known as *reinsurance*, where the two players are known as the insurer and the reinsurer respectively; the transfer cost is known as the *reinsurance premium*. The primary aim of optimal reinsurance design is to find the ‘best’ trade-off between risk-bearing and profit-earning. The Stop-Loss reinsurance contract is optimal when minimising the variance of an insurer’s retained risk is considered (see Borch, 1960) or when maximising the expected utility of insurer’s final wealth (see Arrow, 1963). After the two pioneering works, optimal reinsurance problems have been widely studied in various settings. For example, Gajek and Zagrodny (2000) and Kaluszka (2001 and 2004) have studied the variance-minimisation problem with premium principles including standard deviation premium principle and mean-variance premium principle. In recent years, the risk-measure based optimal reinsurance model has received vast attention. Two commonly used risk measures, *Value-at-risk* (VaR) and *Conditional Value-at-risk* (CVaR), are treated by Cai and Tan (2007), Cai *et al.* (2008), Cheung (2010), Asimit *et al.* (2013c) and Cui *et al.* (2013). On the other hand, optimal reinsurance problems modeled with multiple risks or counterparty default risks are investigated by Bernard and Ludkovski (2012), Cai and Wei (2012), Asimit *et al.* (2013a) and Cai *et al.* (2014).

As anticipated, a risk transfer is not always made in between two players, which is possible if multiple reinsurance players accept to share the risk with one insurer or a pool of insurers. Alternatively, one can place this risk sharing problem in the context of an insurance group that intends to achieve capital efficiency (see Asimit *et al.*, 2013b and Asimit *et al.*, 2016). This standard risk management tool aims to reduce the capital requirements for an insurance group that has multiple legal entities with specific regulatory requirements, where many economic constraints are usually imposed internally and externally. The problem is not tractable in closed-form solutions due its complexity, and therefore, it is rarely discussed in the literature.

So far, the omnipresent model risk, i.e. the risk of choosing a wrong model, has been fashionably ignored; data scarcity is a common problem in practice, which is a standard source for model error. The

same problem arises if data do not reflect the future predictions, especially in extreme scenarios where expert opinion is the usual choice for practitioners. In any of these cases, the underlying model is not unknowable, but unknown, and the decision-maker is exposed to a higher level of uncertainty when robust decisions are sought. Robust optimisation is precisely the standard method that helps in reducing the model risk and it has been translated in the optimal reinsurance context by Balbás *et al.* (2015) and Asimit *et al.* (2016). At a superficial level, one may think that the risk level increases as a result of conservative/robust decisions, which in fact is not true, since the main aim is to reduce the change in the chosen decision, if the “true” model would have been known. Thus, this naïve perception is actually not true for robust optimised decisions and it would be interesting to show how one can take advantage of computational methods when searching for robust decisions, rather than focusing on model-specific closed-form solutions.

We have tried to justify, by discussing various strands of research on optimal risk transfers, but beyond the optimal reinsurance problem, why numerical methods are of interest. Our goal is now to demonstrate that computations are possible in an efficient manner. In the context of optimal reinsurance design, the efficiency of the numerical methods is high for some particular convex problems as justified in Tan and Weng (2014). This excellent paper provides the way forward in solving optimisation problems of that type when multiple business lines and/or other risks (such as background, counterparty default and model risks) are included in the model.

The remainder of this paper is organised as follows. Section 2 introduces the background knowledge of this paper including some key definitions. Formulations of intended optimisation problems are explained in Section 3. The uncertainty of model choice is discussed by analysing worst-case models in Section 4. Section 5 provides numerical examples, while the main conclusions of the paper are summarised in Section 6. All proofs are relegated in Section 7.

## 2. BACKGROUND

Consider a two-player (only for simplicity, but further generalisations are possible) insurance setting where the first player holds  $m$  lines of business (LOB) with random risks, denoted as  $X_1, \dots, X_m$ , which is defined on a probability space  $(\Omega, \mathcal{E}, \mathcal{P})$  with  $\Omega \subseteq \mathfrak{R}_+$ . Assume that for each of the risks  $X_k$ ,  $k \in \mathcal{M} := \{1, \dots, m\}$ , the first insurance player transfers  $I_{2k}[X_k]$  to a secondary player, and its remaining risk becomes  $I_{1k}[X_k] = X_k - I_{2k}[X_k]$ . There also exists a cost of transfer denoted by  $P$ . Consequently, the after-transfer risk positions of the two players becomes

$$\sum_{k=1}^m I_{1k}[X_k] + P \quad \text{and} \quad \sum_{k=1}^m I_{2k}[X_k] - P.$$

The risk preferences of the first and second player are represented by a *risk measure (RM)* denoted as  $\rho_1$  and  $\rho_2$ , respectively. Thus, there is a rationality constraint that both players’ after-transfer risk position should be no worse than their before-transfer position. That is,

$$\rho_1 \left( \sum_{k=1}^m I_{1k}[X_k] + P \right) \leq \rho_1 \left( \sum_{k=1}^m X_k \right) \quad \text{and} \quad \rho_2 \left( \sum_{k=1}^m I_{2k}[X_k] - P \right) \leq \rho_2(0). \quad (2.1)$$

Here are some properties of a RM that will later be recalled.

**Definition 2.1.** Let  $(\Omega, \mathcal{E}, \mathcal{P})$  and  $\mathcal{X}$  be a probability space and a linear space of random variables on  $\Omega$ , respectively. A RM  $\rho : \mathcal{X} \rightarrow \mathfrak{R}$  is translation invariant if

$$\rho(X + a) = \rho(X) + a \text{ for any } X \in \mathcal{X} \text{ and } a \in \mathfrak{R}.$$

The comonotonic property of  $\rho$  requires that  $\rho(X + f(X)) = \rho(X) + \rho(f(X))$  for any  $X \in \mathcal{X}$  and non-decreasing function  $f$ .

If  $\rho_1$  and  $\rho_2$  have the translation invariant property, the rationality constraints in (2.1) become

$$\max \left\{ \underline{P}, \rho_2 \left( \sum_{k=1}^m I_{2k}[X_k] \right) \right\} \leq P \leq \min \left\{ \bar{P}, \rho_1 \left( \sum_{k=1}^m X_k \right) - \rho_1 \left( \sum_{k=1}^m I_{1k}[X_k] \right) \right\}, \quad (2.2)$$

where  $\underline{P}$  is the minimum charge that the second player will agree in a transfer, while  $\bar{P}$  is the maximum price that the first player is prepared to pay in exchange for a risk transfer.

The set of possible ceding functions must satisfy  $I_{1k} \in \mathcal{I} := \{f : 0 \leq f(x) \leq x\}$ , for all  $1 \leq k \leq m$  and the set of feasible contracts,  $\mathcal{A}$  consists of elements  $(I_{11}, \dots, I_{1m}, P) \in \mathcal{I}^m \times \mathfrak{R}$ . Alternatively, in order to avoid potential moral hazard issues, the set of feasible contracts can be defined by  $\mathcal{A}_{co}$  consisting of elements  $(I_{11}, \dots, I_{1m}, P) \in \mathcal{I}_{co}^m \times \mathfrak{R}$ , where ceding functions are comonotone

$$\mathcal{I}_{co} := \{f : 0 \leq f(x) \leq x, f(x) \text{ and } x - f(x) \text{ are non-decreasing in } x \forall x\}.$$

The risk transfer cost  $P$  is based on the second player's risk preference and is usually based on another RM,  $\Pi(\cdot)$  such that  $\sum_{k=1}^m \Pi(I_{2k}[X_k]) \leq P$ . Without loss of generality, it is assumed that the second player has the same preference for all LOBs, but otherwise, the numerical technologies can be easily adapted. There is a vast literature on RMs and it would be impossible to discuss all of them in one paper, but we will focus on the most sensible ones that are socially accepted among academics and practitioners: the classical standard deviation, the class of *coherent* RMs discussed in the seminal paper of Artzner *et al.* (1999) and the well-known VaR. All of these RMs are formally defined now.

**Definition 2.2.** Let  $b \in \mathfrak{R}_+$  and  $Z$  be a random risk defined on  $(\Omega, \mathcal{E}, \mathcal{P})$ . The standard deviation RM is given by  $\rho(Z) = EZ + b \text{Sd}(Z)$ , where  $\text{Sd}(\cdot)$  is the usual standard deviation.

A RM is *law-invariant* if the risk measurement is solely based on the risk distribution, i.e. risks equally distributed are measured with the same exposure and all RMs are assumed to satisfy this property. One of the most used RMs in practice that is not coherent is VaR. By definition, the VaR of a generic loss variable  $Z$  at a confidence level  $a$ ,  $\text{VaR}_a(Z)$ , is the  $a\%$ -quantile, i.e.

$$\text{VaR}_a(Z) := \inf\{z \geq z_0 : \Pr(Z \leq z) \geq a\},$$

where  $z_0 := \sup\{z \in \mathfrak{R} : \Pr(Z \leq z) = 0\}$  represents the left-end point of the distribution function of  $Z$ . By convention,  $\inf \emptyset := +\infty$  and  $\sup \emptyset := -\infty$ . The class of coherent RMs consists of law-invariant RMs that are *monotone, translation invariant, positive homogeneous and sub-additive* (for details, see Artzner *et al.*, 1999). The subclass of coherent RMs that possesses the *comonotonic* property is known as the *distortion* RMs (see Wang *et al.*, 1997).

**Definition 2.3.** Let  $g : [0, 1] \leftarrow [0, 1]$  be a distortion function, i.e. non-decreasing and concave function such that  $g(0) = 0$  and  $g(1) = 1$ . Denote  $\phi(\cdot) := g'(1 - \cdot)$ . Then a distortion RM has the following representation  $\rho_g(Z) := \int_0^1 \phi(a) \text{VaR}_a(Z) da$ , where the random risk  $Z$  is on  $(\Omega, \mathcal{E}, \mathcal{P})$ .

Note that the derivative of  $g$  is understood as the right derivative, whenever it does exist, which may not exist for at most a countable set, due to the concavity property of  $g$ . A well-known distortion RM is CVaR with distortion function  $g(x) = \frac{x}{1-a} \wedge 1$ , where  $x \wedge z := \min(x, z)$ . CVaR has appeared in various forms in the literature (see Acerbi and Tasche, 2002), but we prefer using the following representation

$$\text{CVaR}_a(Z) := \inf_{t \in \mathfrak{R}} \left\{ t + \frac{1}{1-a} E(Z - t)_+ \right\}, \quad \text{where } (x)_+ := \max(x, 0), \quad (2.3)$$

as first defined in Rockafeller and Uryasev (2000), due to its computational advance.

Not all coherent RMs are comonotonic, and a large parametric subclass of coherent RMs that are not comonotonic is introduced by Fischer (2003) as follows:

$$\Psi_\alpha(Z) := EZ + \lambda \left( E(Z - EZ)_+^\alpha \right)^{1/\alpha} \quad \text{with } 0 \leq \lambda \leq 1, 1 \leq \alpha \leq \infty,$$

where  $\alpha = \infty$  in the above is read as  $EZ + \lambda \left( \text{VaR}_1(Z) - E(Z) \right)$ .

### 3. EMPIRICAL RISK TRANSFER MODELS

This section first describes the main optimal risk transfer problem to be solved by using our numerical methods. We further explain the one LOB case and extend these results for the multiple LOBs by providing formulations that can be implemented and solved in an efficient manner.

**3.1. Main optimisation problem.** The main problem that we propose to investigate is the weighted average of two insurance players under various economic and/or rationality constraints like (2.2); its mathematical formulation is as follows:

$$\begin{aligned} \min_{\substack{(I_{11}, \dots, I_{1m}) \in \\ \mathcal{A} \text{ or } \mathcal{A}_{co}; P \in \mathfrak{R}}} \left\{ \delta_1 \rho_1 \left( \sum_{k=1}^m I_{1k}[X_k] + P \right) + \delta_2 \rho_2 \left( \sum_{k=1}^m (X_k - I_{1k}[X_k]) - P \right) \right\} \\ \text{s.t. } \sum_{k=1}^m \Pi(X_k - I_{1k}[X_k]) \leq P, \underline{P} \leq P \leq \bar{P}, \\ \rho_1 \left( \sum_{k=1}^m I_{1k}[X_k] \right) + P \leq \rho_1 \left( \sum_{k=1}^m X_k \right), \rho_2 \left( \sum_{k=1}^m (X_k - I_{1k}[X_k]) \right) \leq P. \end{aligned} \quad (3.1)$$

where  $\delta_1, \delta_2 \geq 0$ . The above contains the risk sharing problems, where the players cooperate to reduce their risks, known as *Pareto efficient* transfers, and under certain conditions, the set of these transfers are as given in the above-mentioned problem (see, for example, Aase 2002). The risk sharing problem may require to optimise over  $\mathcal{A}_{co}$  if the moral hazard is highly likely. Risk transfers within an insurance group do not involve moral hazard, since each legal entity is owned in a certain proportion by the same owner, and therefore, the feasibility set is  $\mathcal{A}$ . Setting,  $(\delta_1, \delta_2)$  equal to  $(1, 0)$  or  $(0, 1)$ , then (3.1) becomes a standard optimal reinsurance contract evaluated from the point of view of the first and second insurance player, respectively. We believe that formulation (3.1) is a fairly general problem and it worth investigating.

The optimisation problem (3.1) has infinite dimension, as one needs to search over all possible ceding functions of  $I_{1k}$ ,  $k \in \mathcal{M}$ . This issue can be overcome by reformulating (3.1) into a discretised setting. Assume that a sample of size  $n$ , denoted by  $\mathbf{x}_k = (x_{k1}, \dots, x_{kn})^T$ , is available for  $X_k$ ,  $k \in \mathcal{M}$ . This is possible if historical data is available, otherwise, a socially accepted proxy model can be used to draw those samples. For each observed loss  $x_{ki}$ ,  $i = 1, \dots, n$ , we assume that the first player transfers  $z_{ki}$  and retains  $y_{ki} := x_{ki} - z_{ki}$ . Consequently, the decision variables,  $I_{1k}, I_{2k}$ ,  $k \in \mathcal{M}$ , in (3.1) can be replaced by vectors  $\mathbf{y}_k = (y_{k1}, \dots, y_{kn})^T$  and  $\mathbf{z}_k = (z_{k1}, \dots, z_{kn})^T$  respectively, and hence, the infinite-dimensional optimisation problem (3.1) can be translated into the following  $(n \times 2m + 1)$ -dimensional problem.

Now, if we optimise over  $\mathcal{A}$ , the risk transfers should satisfy  $\mathbf{y}_k \leq \mathbf{x}_k$ ,  $\mathbf{z}_k \leq \mathbf{x}_k$  and  $\mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k$  for all  $k \in \mathcal{M}$ . Recall that the inequality (equality) between two vectors is understood componentwise, i.e.  $\mathbf{a} \leq \mathbf{b}$  ( $\mathbf{a} = \mathbf{b}$ ) if and only if  $a_i \leq b_i$  ( $a_i = b_i$ ) for all  $i$ . For a given  $k \in \mathcal{M}$ , an allocation  $(\mathbf{y}_k, \mathbf{z}_k)$  of  $\mathbf{x}_k$  is feasible in  $\mathcal{A}_{co}$  if the vectors are increasingly ordered. That is,

$$\begin{cases} 0 \leq y_{ki} \leq x_{ki}, 0 \leq y_{k,i+1} - y_{ki}, & \forall i = 1, 2, \dots, n, \\ 0 \leq z_{ki} \leq x_{ki}, 0 \leq z_{k,i+1} - z_{ki}, & \forall i = 1, 2, \dots, n, \\ y_{ki} + z_{ki} = x_{ki}, & \forall i = 1, 2, \dots, n, \end{cases}$$

or equivalently,

$$\begin{cases} 0 \leq y_{ki} \leq x_{ki}, y_{k,i+1} - y_{ki} \leq x_{k,i+1} - x_{ki}, & \forall i = 1, 2, \dots, n, \\ 0 \leq z_{ki} \leq x_{ki}, z_{k,i+1} - z_{ki} \leq x_{k,i+1} - x_{ki}, & \forall i = 1, 2, \dots, n, \\ y_{ki} + z_{ki} = x_{ki}, & \forall i = 1, 2, \dots, n. \end{cases}$$

The latter could be written as follows:

$$\mathbf{A}\mathbf{y}_k \leq \mathbf{A}\mathbf{x}_k, \mathbf{A}\mathbf{z}_k \leq \mathbf{A}\mathbf{x}_k, \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k,$$

where the  $n \times n$  matrix  $\mathbf{A}$  is defined by

$$\mathbf{A} := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ -1 & 1 & \cdots & 0 \\ & \ddots & \ddots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Various problems that we further investigate are convex and reducible to *Second Order Conic Programming* (SOCP). SOCP problems are convex optimization problems in which a linear function is minimised over the intersection of various second-order (quadratic) cones. The standard form of SOCP has the following representation

$$\min_{\mathbf{x}} \{ \mathbf{a}_1^T \mathbf{x} \} \quad \text{s.t.} \quad \|\mathbf{B}_l \mathbf{x} + \mathbf{b}_l\| \leq \mathbf{c}_l^T \mathbf{x} + \mathbf{d}_l, \quad l = 1, \dots, L, \quad \mathbf{B} \mathbf{x} = \mathbf{a}_2,$$

where  $\mathbf{a}_1 \in \mathbb{R}^n$ ,  $\mathbf{B}_l \in \mathbb{R}^{n_l \times n}$ ,  $\mathbf{b}_l \in \mathbb{R}^{n_l}$ ,  $\mathbf{c}_l \in \mathbb{R}^n$ ,  $\mathbf{d}_l \in \mathbb{R}$ ,  $\mathbf{B} \in \mathbb{R}^{p \times n}$ ,  $\mathbf{a}_2 \in \mathbb{R}^p$ . In addition,  $\|\cdot\|$  represents the *Euclidean distance*, i.e.  $\|\mathbf{x}\| := \left( \sum_{i=1}^n x_i^2 \right)^{1/2}$ . A more concise introduction and summary on SOCP can be found in the Appendix of Tan and Weng (2014).

SOCP problems have received considerable attention from researchers due to its wide range of applications (see Lobo *et al.*, 1998, and Alizadeh and Goldfarb, 2003) and very efficient primal-dual interior-point

methods that have been developed for such optimisation problems. The study of primal-dual interior point methods for SOCP problems started with Nesterov and Todd (1997, 1998). These authors presented their results in the context of optimisation over self-scaled cones, which includes the class of second-order cones as a special case. Their work culminated in the development of a particular primal-dual method called the NT method. Adler and Alizadeh (1995) studied the relationship between semi-definite problems, SOCP and the specialised so-called  $XZ + ZX$  method of Alizadeh *et al.* (1998) to SOCP. Then, Alizadeh and Schmieta (1997) gave non-degeneracy conditions for SOCP and developed a numerically stable implementation of the  $XZ + ZX$  method; this implementation is used in the SDPpack software package. Subsequently, Monteiro and Tsuchiya (2000) proved that this method, and hence, all members of the Monteiro-Zhang family of methods, have a polynomial iteration complexity. The computational effort of solving SOCP with a standard form is  $O(L^{1/2})$  iterations of  $O(n^2 \sum n_i)$  in the case of primal-dual interior-point method. In practice, the special structure (e.g., sparsity) often allows solving the SOCP more efficiently. There are now several packages available that may handle SOCP problems such as the SDPpack package noted above and the Sturm's SeDuMi, which is another widely available package that is based on the Nesterov-Todd method.

Because of SOCP's broad applicability and computational tractability, there are many studies showing classes of optimisation problems that could be reformulated as SOCP problems. Although it may be possible to formulate an optimisation problem as an SOCP, the task may be far from being trivial. For example, problems with *hyperbolic constraints* are SOCP representable by using the fact that a constraint of the form

$$\mathbf{w}^T \mathbf{w} \leq xy, \quad \text{where } \mathbf{w} \in \mathfrak{R}^n, \quad x, y \in \mathfrak{R}_+,$$

is equivalent to the following second-order cone constraint

$$\left\| \begin{pmatrix} 2\mathbf{w} \\ x - y \end{pmatrix} \right\| \leq x + y.$$

For further details about transforming many kinds of constraints into second-order cone inequalities, we refer the reader to Lobo *et al.* (1998) and Alizadeh and Goldfarb (2003).

The RMs in (3.1) are replaced by their sample estimators. Let  $Z$  be a discrete random variable with possible outcomes  $\{z_1, \dots, z_n\}$ , each having the same chance to occur. Thus, the standard deviation RM from Definition 2.2 is replaced by its unbiased estimator,  $\frac{1}{n} \mathbf{1}^T \mathbf{z} + b \frac{\|\mathbf{Qz}\|}{\sqrt{n-1}}$ , where the  $n \times n$  matrix  $\mathbf{Q}$  is defined as

$$\mathbf{Q} := \begin{pmatrix} 1 - \frac{1}{n} & -\frac{1}{n} & \dots & -\frac{1}{n} \\ -\frac{1}{n} & 1 - \frac{1}{n} & \dots & -\frac{1}{n} \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{n} & -\frac{1}{n} & \dots & 1 - \frac{1}{n} \end{pmatrix}.$$

By definition,  $\mathbf{0}$  and  $\mathbf{1}$  are column vectors of zeros and ones, respectively, where the dimension is not explicitly given. The unbiased estimator is advisable for relatively large values of  $n$ , while the biased estimator (with  $\sqrt{n-1}$  replaced by  $\sqrt{n}$ ) is more appropriate for small  $n$ .

The VaR and all other distortion RMs dependent on the ordering of  $z_i$ 's. Let  $z_{(1)} \leq z_{(2)} \leq \dots \leq z_{(n)}$  be the order statistics sample. Then, the VaR-estimator for  $\text{VaR}_a(Z)$  is  $z_{(\lceil na \rceil)}$  with  $\lceil \cdot \rceil$  being the ceiling

function. Now, distortion RMs are a tad more cumbersome and its computational complexity largely depends on whether or not the sample is ordered. That is, if the sample is increasingly ordered,  $z_{(i)} = z_i$  for all  $1 \leq i \leq n$ , then

$$\rho_g(Z) = \sum_{i=1}^n \left( g((n-i+1)/n) - g((n-i)/n) \right) z_i. \quad (3.2)$$

Otherwise,

$$\begin{aligned} \rho_g(Z) &= \sum_{i=1}^n \left( g((n-i+1)/n) - g((n-i)/n) \right) z_{(i)} \\ &= \sum_{i=1}^n \left( g((n-i+1)/n) - g((n-i)/n) \right) \left( (n-i+1) \text{CVaR}_{(i-1)/n}(Z) - (n-i) \text{CVaR}_{i/n}(Z) \right) \\ &= b_0 \mathbf{1}^T \mathbf{z} + \sum_{i=1}^{n-1} b_i \text{CVaR}_{i/n}(Z), \end{aligned}$$

where  $b_0 = 1 - g((n-1)/n)$  and  $b_i = (n-i) \left( 2g((n-i)/n) - g((n-i+1)/n) - g((n-i-1)/n) \right)$  for all  $1 \leq i \leq n-1$ .

**3.2. One LOB case.** It is first assumed that  $m = 1$  in order to elaborate the simplest (3.1)-type optimisation problems that one may have. Recall that ordered samples simplify the estimators that are based on VaR or any other distortion RM.

**Problem 3.1.** Assume that (3.1) is optimised over  $\mathcal{A}_{co}$  and without loss of generality  $\mathbf{x}_1$  is assumed to be increasingly ordered. Moreover,  $\rho_1(\cdot) = \text{VaR}_a(\cdot)$ ,  $\rho_2(\cdot) = E(\cdot) + b \text{Sd}(\cdot)$  and  $\Pi(\cdot) = \rho_g(\cdot)$ . Therefore, the following optimisation problem needs to be solved:

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, P) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}}} & \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Q}\mathbf{z}_1\|}{\sqrt{n-1}} \right) \right\} \\ \text{s.t.} & \quad \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Q}\mathbf{z}_1\|}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\ & \quad y_{1a^*} + P \leq x_{1a^*}, \mathbf{a}^T \mathbf{z}_1 \leq P, \\ & \quad \mathbf{A}\mathbf{y}_1 \leq \mathbf{A}\mathbf{x}_1, \mathbf{A}\mathbf{z}_1 \leq \mathbf{A}\mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1, \end{aligned} \quad (3.3)$$

where  $a^* = \lceil na \rceil$  is a scalar and  $a_i = g((n-i+1)/n) - g((n-i)/n)$  for all  $1 \leq i \leq n$  due to equation (3.2).

It is not difficult to show that Problem 3.1 is in fact of the SOCP-type. One may show (as explained in the beginning of the proof of Theorem 3.1) that (3.3) is solved by

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, P, t) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}}} & \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\ \text{s.t.} & \quad \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\ & \quad y_{1a^*} + P \leq x_{1a^*}, \mathbf{a}^T \mathbf{z}_1 \leq P, \|\mathbf{Q}\mathbf{z}_1\| \leq t, \\ & \quad \mathbf{A}\mathbf{y}_1 \leq \mathbf{A}\mathbf{x}_1, \mathbf{A}\mathbf{z}_1 \leq \mathbf{A}\mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1. \end{aligned}$$

Problem 3.1 can be easily adapted when  $\rho_1$  is a distortion RM, since  $\mathbf{y}_1$  is increasingly ordered and thus, its contribution is just a linear function due to (3.2).





while the optimal  $P$  is any value that satisfies the first constraint. Then, (3.6) is solved by the following SOCP:

$$\begin{aligned}
& \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, \mathbf{u}_1, \mathbf{w}, s_1, r_1, t, \mathbf{e}, \mathbf{c}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \\ \times \mathbb{R}^{n \times (n-1)} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}}} \left\{ r_1 + \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\
& \text{s.t. } s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 \leq r_1, \mathbf{0} \leq \mathbf{u}_1, \mathbf{y}_1 - s_1 \mathbf{1} \leq \mathbf{u}_1, \|\mathbf{Qz}_1\| \leq t, \\
& s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 \leq \text{CVaR}_a(X_1) - \underline{P}, \\
& s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 + \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq \text{CVaR}_a(X_1), \\
& s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 + b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i e_i \leq \text{CVaR}_a(X_1), \\
& \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq \bar{P}, b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i e_i \leq \bar{P}, \\
& c_i + \frac{1}{n-i} \mathbf{1}^T \mathbf{w}_i \leq e_i, \mathbf{0} \leq \mathbf{w}_i, \mathbf{z}_1 - c_i \mathbf{1} \leq \mathbf{w}_i, \quad \forall i = 1, \dots, n-1, \\
& \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{z}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1.
\end{aligned} \tag{3.7}$$

The latter equivalence could be proved in a similar manner as Theorem 3.1 is justified and therefore, its proof is left to the reader.

Unfortunately, the  $0 < \delta_1 < \delta_2$  case makes the Problem 3.2 not tractable. The lack of convexity leads only to reformulations that are not computationally efficient and for this reason, we do not further discuss such a setting. In fact, all the remaining optimisation problems assume that  $\delta_1 > \delta_2 > 0$ , but one could follow the same reasons as used in the reformulation (3.7), whenever the  $0 < \delta_1 = \delta_2$  setting is in place.

It would be interesting to point out that solving Problem 3.2 over  $\mathcal{A}_{co}$  is a tad simpler. Thus, if  $\mathbf{x}_1$  and  $\mathbf{y}_1$  are increasingly ordered, then Problem 3.2 is simplified to solving:

$$\begin{aligned}
& \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, P) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}}} \left\{ (\delta_1 - \delta_2)P + \frac{\delta_1}{n(1-a)} \left( (a^* - na)y_{1a^*} + \sum_{i=a^*+1}^n y_{1i} \right) + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Qz}_1\|}{\sqrt{n-1}} \right) \right\} \\
& \text{s.t. } \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Qz}_1\|}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\
& \frac{1}{n(1-a)} \left( (a^* - na)y_{1a^*} + \sum_{i=a^*+1}^n y_{1i} \right) + P \leq \text{CVaR}_a(X_1), \\
& b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathbb{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 - c_i \mathbf{1})_+ \right\} \leq P, \\
& \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1.
\end{aligned}$$

As before, the above is solved by the following SOCP:

$$\begin{aligned}
& \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, \mathbf{w}, P, t, \mathbf{e}, \mathbf{c}) \in \mathbb{R}^n \times \mathbb{R}^n \\ \times \mathbb{R}^{n \times (n-1)} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}}} \left\{ (\delta_1 - \delta_2)P + \frac{\delta_1}{n(1-a)} \left( (a^* - na)y_{1a^*} + \sum_{i=a^*+1}^n y_{1i} \right) + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\
& \text{s.t. } \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \|\mathbf{Q}\mathbf{z}_1\| \leq t, b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i e_i \leq P, \\
& \frac{1}{n(1-a)} \left( (a^* - na)y_{1a^*} + \sum_{i=a^*+1}^n y_{1i} \right) + P \leq \text{CVaR}_a(X_1), \\
& c_i + \frac{1}{n-i} \mathbf{1}^T \mathbf{w}_i \leq e_i, \mathbf{0} \leq \mathbf{w}_i, \mathbf{z}_1 - c_i \mathbf{1} \leq \mathbf{w}_i, \quad \forall i = 1, \dots, n-1, \\
& \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1,
\end{aligned}$$

where the proof is once again left to the reader, since it requires similar arguments as we have seen before.

**Problem 3.3.** Assume that (3.1) is optimised over  $\mathcal{A}$  and the moral hazard is only removed for the first player. Thus, without loss of generality  $\mathbf{x}_1$  and  $\mathbf{y}_1$  are assumed to be increasingly ordered. Moreover,  $\rho_1(\cdot) = \text{VaR}_a(\cdot)$ ,  $\rho_2(\cdot) = E(\cdot) + b \text{Sd}(\cdot)$  and  $\Pi(\cdot) = \rho_g(\cdot)$ . Therefore, the following optimisation problem needs to be solved:

$$\begin{aligned}
& \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, P) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}}} \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Q}\mathbf{z}_1\|}{\sqrt{n-1}} \right) \right\} \\
& \text{s.t. } \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Q}\mathbf{z}_1\|}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, y_{1a^*} + P \leq x_{1a^*}, \\
& b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathbb{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 - c_i \mathbf{1})_+ \right\} \leq P, \\
& \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1,
\end{aligned}$$

which can be solved efficiently as shown in Theorem 3.2.

**Theorem 3.2.** If  $\delta_1 > \delta_2 > 0$ , then Problem 3.3 is solved by an SOCP given by:

$$\begin{aligned}
& \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, \mathbf{w}, P, t, \mathbf{e}, \mathbf{c}) \in \mathbb{R}^n \times \mathbb{R}^n \\ \times \mathbb{R}^{n \times (n-1)} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}}} \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\
& \text{s.t. } \|\mathbf{Q}\mathbf{z}_1\| \leq t, \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\
& b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i e_i \leq P, y_{1a^*} + P \leq x_{1a^*}, \\
& c_i + \frac{1}{n-i} \mathbf{1}^T \mathbf{w}_i \leq e_i, \mathbf{0} \leq \mathbf{w}_i, \mathbf{z}_1 - c_i \mathbf{1} \leq \mathbf{w}_i, \quad \forall i = 1, \dots, n-1, \\
& \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1.
\end{aligned}$$

The next setting aims to illustrate how to deal with non-comonotone RMs and we choose those defined by Fischer (2003).

**Problem 3.4.** Assume that (3.1) is optimised over  $\mathcal{A}$  and the moral hazard issues are removed from the first player's behaviour. Therefore, without loss of generality  $\mathbf{x}_1$  and  $\mathbf{y}_1$  are assumed to be increasingly

ordered. Moreover,  $\rho_1(\cdot) = \text{VaR}_a(\cdot)$ ,  $\rho_2(\cdot) = E(\cdot) + b \text{Sd}(\cdot)$  and  $\Pi(\cdot) = \Psi_\alpha(\cdot)$ . Therefore, the following optimisation problem needs to be solved:

$$\begin{aligned} & \min_{\substack{(y_1, z_1, P) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}}} \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Q}\mathbf{z}_1\|}{\sqrt{n-1}} \right) \right\} \\ & \text{s.t. } \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + b \frac{\|\mathbf{Q}\mathbf{z}_1\|}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, y_{1a^*} + P \leq x_{1a^*}, \\ & \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \lambda \left( \frac{1}{n} \sum_{i=1}^n \left( z_{1i} - \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 \right)_+^\alpha \right)^{1/\alpha} \leq P, \\ & \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1. \end{aligned}$$

Problem 3.4 can be efficiently solved if  $\alpha \in \{1, 2\}$  and the corresponding reformulations are given in Theorem 3.3.

**Theorem 3.3.** Assume that  $\delta_1 > \delta_2 > 0$ . If  $\alpha = 1$ , then Problem 3.4 is solved by the following SOCP:

$$\begin{aligned} & \min_{\substack{(y_1, z_1, P, t, u) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n}} \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\ & \text{s.t. } \|\mathbf{Q}\mathbf{z}_1\| \leq t, \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\ & \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{\lambda}{n} \sum_{i=1}^n u_i \leq P, y_{1a^*} + P \leq x_{1a^*}, \\ & 0 \leq u_i, z_{1i} - \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 \leq u_i, \quad \forall i = 1, \dots, n, \\ & \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1. \end{aligned}$$

If  $\alpha = 2$ , then Problem 3.4 is solved by an SOCP given by:

$$\begin{aligned} & \min_{\substack{(y_1, z_1, P, t, u) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n}} \left\{ (\delta_1 - \delta_2)P + \delta_1 y_{1a^*} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\ & \text{s.t. } \|\mathbf{Q}\mathbf{z}_1\| \leq t, \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\ & \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{\lambda}{n} \|\mathbf{u}\| \leq P, y_{1a^*} + P \leq x_{1a^*}, \\ & \mathbf{0} \leq \mathbf{u}, z_1 - \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 \mathbf{1} \leq \mathbf{u}, \quad \forall i = 1, \dots, n, \\ & \mathbf{0} \leq \mathbf{A}\mathbf{y}_1, \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1. \end{aligned}$$

**3.3. Multiple LOBs case.** When the insurance players have more than one LOB, the numerical methods are similar to what we have seen earlier. For the ease of exposition, assume now that  $m = 2$ . Unfortunately, even if  $\mathbf{x}_k$ ,  $\mathbf{y}_k$  and  $\mathbf{z}_k$  are increasingly ordered for  $k = 1, 2$ , none of  $\mathbf{x}_1 + \mathbf{x}_2$ ,  $\mathbf{y}_1 + \mathbf{y}_2$  or  $\mathbf{z}_1 + \mathbf{z}_2$  are ordered. Therefore, unlike the one LOB case, the empirical representation of *VaR* with multiple LOBs cannot be written as a linear function of decision variables even if the optimisation problem is solved over  $\mathcal{A}_{co}$ . In fact, if  $\rho_1$  or  $\rho_2$  is *VaR*, our model (3.1) becomes a non-convex optimisation problem, and hence, it cannot be reduced to an SOCP and requires special algorithms such as the *Difference of Convex Functions Algorithm* (DCA). The DCA can be applied to solve a wide class of non-convex problems whose

non-convex component can be written as a difference of two convex functions (including VaR). In technical terms, iterations between the primal space and the dual space finally lead to a convergence (see Tao and An, 1998). The DCA has not been used in the actuarial literature, but has been successfully applied by Wozabal *et al.* (2010) in the classical portfolio investment optimisation. However, as emphasised in Tao and An (1998), the nature of the optimum obtained from DCA depends on the starting point chosen by the algorithm, and hence, it is not computationally efficient in solving our optimisation model (3.1) and we choose not to provide details on them..

**Problem 3.5.** *Assume that  $\rho_1(\cdot) = \text{CVaR}_a(\cdot)$ ,  $\rho_2(\cdot) = E(\cdot) + b \text{Sd}(\cdot)$  and  $\Pi(\cdot) = \rho_g(\cdot)$ . Moreover, the portfolio consists of two independent LOBs and (3.1) is optimised over  $\mathcal{A}$ . Therefore, the following optimisation problem needs to be solved:*

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2, P) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}}} & \left\{ (\delta_1 - \delta_2)P + \delta_1 \min_{s \in \mathbb{R}} \left\{ s + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2 - s\mathbf{1})_+ \right\} \right. \\ & \left. + \delta_2 \left( \frac{1}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) + b \sqrt{\frac{\|\mathbf{Q}\mathbf{z}_1\|^2 + \|\mathbf{Q}\mathbf{z}_2\|^2}{n-1}} \right) \right\} \\ \text{s.t.} & \frac{1}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) + b \sqrt{\frac{\|\mathbf{Q}\mathbf{z}_1\|^2 + \|\mathbf{Q}\mathbf{z}_2\|^2}{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \\ & \min_{s \in \mathbb{R}} \left\{ s + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2 - s\mathbf{1})_+ \right\} + P \leq \text{CVaR}_a(X_1 + X_2), \\ & b_0 \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathbb{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2 - c_i \mathbf{1})_+ \right\} \leq P, \\ & \mathbf{y}_k \leq \mathbf{x}_k, \mathbf{z}_k \leq \mathbf{x}_k, \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k \in \{1, 2\}, \end{aligned}$$

where  $\text{CVaR}_a(X_1 + X_2)$  is a scalar and it can be found via (3.2).

**Theorem 3.4.** *If  $\delta_1 > \delta_2 > 0$ , then Problem 3.5 is solved by the following SOCP:*

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2, \mathbf{u}, \mathbf{w}, P, s, r, t, e, c) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \\ \times \mathbb{R}^n \times \mathbb{R}^{n \times (n-1)} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^{n-1} \times \mathbb{R}^{n-1}}} & \left\{ (\delta_1 - \delta_2)P + \delta_1 r + \delta_2 \left( \frac{1}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) + \frac{bt}{\sqrt{n-1}} \right) \right\} \\ \text{s.t.} & \frac{1}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \|\hat{\mathbf{Q}}\mathbf{z}_{12}\| \leq t, \\ & s + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u} \leq r, \mathbf{0} \leq \mathbf{u}, \mathbf{y}_1 + \mathbf{y}_2 - s\mathbf{1} \leq \mathbf{u}, \\ & r + P \leq \text{CVaR}_a(X_1 + X_2), b_0 \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) + \sum_{i=1}^{n-1} b_i e_i \leq P, \\ & c_i + \frac{1}{n-i} \mathbf{1}^T \mathbf{w}_i \leq e_i, \mathbf{0} \leq \mathbf{w}_i, \mathbf{z}_1 + \mathbf{z}_2 - c_i \mathbf{1} \leq \mathbf{w}_i, \quad i = 1, \dots, n-1, \\ & \mathbf{y}_k \leq \mathbf{x}_k, \mathbf{z}_k \leq \mathbf{x}_k, \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k \in \{1, 2\}, \end{aligned}$$

where  $\mathbf{z}_{12}^T = (\mathbf{z}_1^T, \mathbf{z}_2^T)$  and  $\hat{\mathbf{Q}}$  is an  $2n \times 2n$  block matrix given by:

$$\hat{\mathbf{Q}} := \begin{pmatrix} \mathbf{Q} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} \end{pmatrix}.$$

As a final note, recall that  $\mathbf{y}_1 + \mathbf{y}_2$  is not ordered, even if  $\mathbf{y}_1$  and  $\mathbf{y}_2$  are increasingly ordered. Consequently, further simplifications to Problem 3.5 (when optimisation is made over  $\mathcal{A}_{co}$ ) are not possible in the same way that we have found for Problem 3.2.

#### 4. ROBUST OPTIMISATION

It has been assumed throughout Section 3 that  $\mathbf{x}_k$ ,  $k \in \mathcal{M}$ , is drawn from a proxy model. In general, uncertainty related to the underlying probability measure corresponding to the proxy model exists, and there may be many possible choices to model the corresponding probabilities. It is assumed in the current section that the buyer's and seller's set of beliefs (about the underlying probability measure) belong to  $\{\mathcal{P}_1, \dots, \mathcal{P}_N\}$ . That is, the buyer and seller model the uncertainty in a way that their risk measurements are calculated with respect to some index sets  $\mathcal{S}_1, \mathcal{S}_2 \subseteq \{1, \dots, N\}$ , which is indicated by

$$\rho_1(\cdot; \mathcal{P}_i) \quad \text{and} \quad \rho_2(\cdot; \mathcal{P}_j) \quad \text{with} \quad i \in \mathcal{S}_1, j \in \mathcal{S}_2.$$

Let  $p_{ikl} := \mathcal{P}_i(X_k = x_{kl})$ ,  $\mathbf{x}_k = (x_{k1}, \dots, x_{kn})^T$  and  $\mathbf{p}_{ik} = (p_{ik1}, \dots, p_{ikn})^T$ . Note that  $\mathbf{p}_{ik} = \frac{1}{n}\mathbf{1}$ , i.e.  $\mathcal{P}_i$  is the empirical distribution where each possible outcome  $\{x_{k1}, x_{k2}, \dots, x_{kn}\}$  has the same chance to occur. That is, the proxy model assumed throughout Section 3 is a potential candidate of  $\mathcal{P} \in \{\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_N\}$ . Models other than the empirical distribution can be obtained by discretising various fitted models. For example, let  $F_i(\cdot) = \mathcal{P}_i(X_k \leq \cdot)$  with  $k \in \mathcal{M}$  denoting the cumulative distribution function (*cdf*) of  $X_k$  under  $\mathcal{P}_i$ , then

$$p_{ikl} = F_i\left(\frac{x_{k,l+1} + x_{kl}}{2}; \hat{\nu}\right) - F_i\left(\frac{x_{kl} + x_{k,l-1}}{2}; \hat{\nu}\right) \quad \text{for all } i = 1, 2, \dots, N, k \in \mathcal{M} \text{ and } l = 1, 2, \dots, n,$$

where  $\hat{\nu}$  is the set of parameter estimates obtained via various estimation methods such as the Maximum Likelihood estimation. By convention,  $x_{k0} = -\infty$  and  $x_{k,n+1} = +\infty$ . More details about the construction of candidate models can be found in Asimit *et al.* (2017). The LOBs are assumed to be independent of each other, which helps us to write the model in a simpler way, but the lack of independence does not create additional computational effort. Thus, instead of solving the optimisation problem (3.1), the following *worst-case* model is proposed:

$$\begin{aligned} \min_{\substack{(I_{11}, \dots, I_{1m}) \in \\ \mathcal{A} \text{ or } \mathcal{A}_{co}; P \in \mathbb{R}}} \max_{\substack{(i,j) \in \\ \mathcal{S}_1 \times \mathcal{S}_2}} & \left\{ \delta_1 \rho_1\left(\sum_{k=1}^m I_{1k}[X_k] + P; \mathcal{P}_i\right) + \delta_2 \rho_2\left(\sum_{k=1}^m (X_k - I_{1k}[X_k]) - P; \mathcal{P}_j\right) \right\} \quad (4.1) \\ \text{s.t.} & \max_{l \in \mathcal{S}_2} \sum_{k=1}^m \Pi(X_k - I_{1k}[X_k]; \mathcal{P}_l) \leq P, \underline{P} \leq P \leq \bar{P}, \\ & \rho_1\left(\sum_{k=1}^m I_{1k}[X_k]; \mathcal{P}_i\right) + P \leq \rho_1\left(\sum_{k=1}^m X_k\right), \quad \forall i \in \mathcal{S}_1, \\ & \rho_2\left(\sum_{k=1}^m (X_k - I_{1k}[X_k]); \mathcal{P}_j\right) \leq P, \quad \forall j \in \mathcal{S}_2. \end{aligned}$$

More details about the worst-case formulation, its advantages and caveats can be found in Asimit *et al.* (2017). Recall that without loss of generality, sample sizes are the same amongst all LOBs, an assumption that enables reduction in the notation abuse that we inevitably rely on during this paper.

Before giving some examples of (4.1) we need to define the risk measurements under various probability measure  $\mathcal{P}_i$ ,  $i = 1, \dots, N$ . Let  $Z$  be a generic discrete random variable with possible outcomes  $\{z_1, \dots, z_n\}$

such that  $\mathcal{P}(Z = z_l) = p_l$ . One may show that the standard deviation under  $\mathcal{P}$  is

$$Sd(Z; \mathcal{P}) = \|\tilde{\mathbf{Q}}\mathbf{z}\|, \quad \text{where } \tilde{q}_{l_1 l_2} = \sqrt{p_{l_1}}(1_{\{l_1=l_2\}} - p_{l_2}) \quad \text{for all } 1 \leq l_1, l_2 \leq n,$$

where by definition,  $1_A$  represents the indicator operator corresponding to set  $A$  that equals to 1 if  $A$  is true and 0 otherwise. Recall that when we work with the empirical measure, i.e.  $p_i = 1/n$  for all  $1 \leq i \leq n$ , then standard deviation has a correction factor of  $\sqrt{n/(n-1)}$  in order to recover the use of the standard sample variance unbiased estimator, which appears in all examples from Section 3. The correction factor is advisable for reasonable sized samples and not for small samples. The current section discusses robust formulations and therefore, it is expected to be in a situation where data scarcity is present; the correction factor does not further appear in the case of empirical probability measure. Thus, the standard deviation RM of  $X_k$  under  $\mathcal{P}_i$  is given by

$$\mathbf{p}_{ik}^T \mathbf{x}_k + b \|\tilde{\mathbf{Q}}_{ik} \mathbf{x}_k\|, \quad \text{where } \tilde{q}_{l_1 l_2 ik} = \sqrt{p_{ikl_1}}(1_{\{l_1=l_2\}} - p_{ikl_2}) \quad \text{for all } 1 \leq l_1, l_2 \leq n.$$

The CVaR formulation from (2.3) now becomes

$$\text{CVaR}_a(X_k; \mathcal{P}_i) = \inf_{t \in \mathfrak{R}} \left\{ t + \frac{1}{1-a} \mathbf{p}_{ik}^T (\mathbf{x}_k - t\mathbf{1})_+ \right\}.$$

Other distortion RMs than CVaR depend on the knowledge of sample ordering. Whenever the sample is known to be in ascending order, we may write

$$\rho_g(X_k; \mathcal{P}_i) = \sum_{l=1}^n \left( g \left( 1 - \sum_{s=1}^{l-1} p_{iks} \right) - g \left( 1 - \sum_{s=1}^l p_{iks} \right) \right) x_{kl}.$$

If there is no piece of information regarding the sample ordering, then distortion risk measures are more cumbersome and for this reasons, this section avoids dealing with non ordered samples when distortion risk measures other than CVaR are involved.

**4.1. One LOB case.** The case in which there is one single LOB, i.e.  $m = 1$ , is not difficult and not very different than the their non-robust counterpart. To make our point, only one example is discussed in details and is given by Problem 4.1.

**Problem 4.1.** Assume that (4.1) is optimised over  $\mathcal{A}_{co}$  and without loss of generality,  $\mathbf{x}_1$  is assumed to be increasingly ordered. Moreover,  $\rho_1(\cdot) = \rho_g(\cdot)$ ,  $\rho_2(\cdot) = E(\cdot) + b Sd(\cdot)$  and  $\Pi(\cdot) = (1 + \theta)E(\cdot)$ ,  $\theta > 0$ , implying the following optimisation:

$$\min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, P) \in \\ \mathfrak{R}^n \times \mathfrak{R}^n \times \mathfrak{R}}} \max_{\substack{(i,j) \in \\ \mathcal{S}_1 \times \mathcal{S}_2}} \left\{ (\delta_1 - \delta_2)P + \delta_1 \tilde{\mathbf{a}}_{i1}^T \mathbf{y}_1 + \delta_2 \left( \mathbf{p}_{j1}^T \mathbf{z}_1 + b \|\tilde{\mathbf{Q}}_{j1} \mathbf{z}_1\| \right) \right\}$$

$$s.t. \quad \max_{l \in \mathcal{S}_2} \{ \mathbf{p}_{l1}^T \mathbf{z}_1 \} \leq \frac{P}{1 + \theta}, \quad P \leq P \leq \bar{P},$$

$$\tilde{\mathbf{a}}_{i1}^T \mathbf{y}_1 + P \leq \tilde{\mathbf{a}}_{i1}^T \mathbf{x}_1, \quad \forall i \in \mathcal{S}_1,$$

$$\mathbf{p}_{j1}^T \mathbf{z}_1 + b \|\tilde{\mathbf{Q}}_{j1} \mathbf{z}_1\| \leq P, \quad \forall j \in \mathcal{S}_2,$$

$$\mathbf{A} \mathbf{y}_1 \leq \mathbf{A} \mathbf{x}_1, \quad \mathbf{A} \mathbf{z}_1 \leq \mathbf{A} \mathbf{x}_1, \quad \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1,$$

where  $\tilde{\mathbf{a}}_{i1} := (a_{i11}, \dots, a_{i1n})^T$  with  $a_{i1l} = g \left( 1 - \sum_{s=1}^{l-1} p_{iks} \right) - g \left( 1 - \sum_{s=1}^l p_{iks} \right)$  for all  $1 \leq l \leq n$ .

We are now able to transform the above into an SOCP that can be efficiently solved using standard commercial solvers.

**Theorem 4.1.** *If  $\delta_1 > \delta_2 > 0$ , then Problem 4.1 is solved by*

$$\begin{aligned} & \min_{\substack{(y_1, z_1, P, r, t_1) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^n}} r \\ & \text{s.t. } (\delta_1 - \delta_2)P + \delta_1 \tilde{\mathbf{a}}_{i_1}^T \mathbf{y}_1 + \delta_2 (\mathbf{p}_{j_1}^T \mathbf{z}_1 + bt_{j_1}) \leq r, \quad \forall (i, j) \in \mathcal{S}_1 \times \mathcal{S}_2, \\ & \tilde{\mathbf{a}}_{i_1}^T \mathbf{y}_1 + P \leq \tilde{\mathbf{a}}_{i_1}^T \mathbf{x}_1, \underline{P} \leq P \leq \bar{P}, \|\tilde{\mathbf{Q}}_{j_1} \mathbf{z}_1\| \leq t_{j_1}, \quad \forall (i, j) \in \mathcal{S}_1 \times \mathcal{S}_2, \\ & \mathbf{p}_{j_1}^T \mathbf{z}_1 + bt_{j_1} \leq P, \mathbf{p}_{j_1}^T \mathbf{z}_1 \leq \frac{P}{1 + \theta}, \quad \forall j \in \mathcal{S}_2, \\ & \mathbf{A} \mathbf{y}_1 \leq \mathbf{A} \mathbf{x}_1, \mathbf{A} \mathbf{z}_1 \leq \mathbf{A} \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1, \end{aligned}$$

where  $m = |\mathcal{S}_2|$  is the cardinality of  $\mathcal{S}_2$ .

**4.2. Multiple LOBs case.** Robust optimisation with multiple LOBs is a difficult problem when dealing with distortion RMs since the ordering is not preserved when summing the risks. The mathematical formulation of the latter statement is that  $(X_1 + Y_1, X + Y)$  may not be a comonotone vector even if  $(X_1, X)$  and  $(Y_1, Y)$  are comonotone vectors. Reformulations are possible, though a tad more cumbersome than what we have seen by now, if CVaR is not used to order the buyer and/or seller risk preferences. Once again, for the ease of exposition, it is assumed an insurance portfolio consisting of two LOBs, i.e.  $m = 2$ .

**Problem 4.2.** *Assume that  $\rho_1(\cdot) = \text{CVaR}_a(\cdot)$ ,  $\rho_2(\cdot) = E(\cdot) + b \text{Sd}(\cdot)$  and  $\Pi(\cdot) = (1 + \theta)E(\cdot)$ ,  $\theta > 0$ . Moreover, the portfolio consists of two independent LOBs and (4.1) is optimised over  $\mathcal{A}$ . Thus, the following optimisation problem must be solved:*

$$\begin{aligned} & \min_{\substack{(y_1, y_2, z_1, z_2, P) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}}} \max_{\substack{(i_1, i_2, j_1, j_2) \in \\ \mathcal{S}_1 \times \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_2}} \left\{ (\delta_1 - \delta_2)P + \delta_1 \min_{s \in \mathbb{R}} \left\{ s + \frac{1}{1 - a} \mathbf{q}_{i_1 i_2}^T (\mathbf{y}_1 + \mathbf{y}_2 - s \mathbf{1})_+ \right\} \right. \\ & \quad \left. + \delta_2 \left( \mathbf{p}_{j_1}^T \mathbf{z}_1 + \mathbf{p}_{j_2}^T \mathbf{z}_2 + b \sqrt{\|\tilde{\mathbf{Q}}_{j_1} \mathbf{z}_1\|^2 + \|\tilde{\mathbf{Q}}_{j_2} \mathbf{z}_2\|^2} \right) \right\} \\ & \text{s.t. } \max_{\substack{(l_1, l_2) \in \\ \mathcal{S}_2 \times \mathcal{S}_2}} \left\{ \mathbf{p}_{l_1}^T \mathbf{z}_1 + \mathbf{p}_{l_2}^T \mathbf{z}_2 \right\} \leq \frac{P}{1 + \theta}, \underline{P} \leq P \leq \bar{P}, \\ & \min_{s \in \mathbb{R}} \left\{ s + \frac{1}{1 - a} \mathbf{q}_{i_1 i_2}^T (\mathbf{y}_1 + \mathbf{y}_2 - s \mathbf{1})_+ \right\} + P \leq \text{CVaR}_a(X_1 + X_2; \mathcal{Q}_{i_1 i_2}), \quad \forall (i_1, i_2) \in \mathcal{S}_1 \times \mathcal{S}_1, \\ & \mathbf{p}_{j_1}^T \mathbf{z}_1 + \mathbf{p}_{j_2}^T \mathbf{z}_2 + b \sqrt{\|\tilde{\mathbf{Q}}_{j_1} \mathbf{z}_1\|^2 + \|\tilde{\mathbf{Q}}_{j_2} \mathbf{z}_2\|^2} \leq P, \quad \forall (j_1, j_2) \in \mathcal{S}_2 \times \mathcal{S}_2, \\ & \mathbf{y}_k \leq \mathbf{x}_k, \mathbf{z}_k \leq \mathbf{x}_k, \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k \in \{1, 2\}, \end{aligned}$$

where the probability measure  $\mathcal{Q}_{i_1 i_2}$  assigns the mass as follows

$$\mathbf{q}_{i_1 i_2} := (q_{i_1 i_2 1}, \dots, q_{i_1 i_2 n})^T \quad \text{with} \quad q_{i_1 i_2 l} = p_{i_1 1l} \times p_{i_2 2l}, \quad \forall l \leq n.$$



**Theorem 4.2.** *If  $\delta_1 > \delta_2 > 0$ , then Problem 4.2 is solved by the following SOCP:*

$$\begin{aligned}
& \min_{\substack{(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2, \mathbf{m}, s \in P, r, t) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \\ \times \mathbb{R}^{n \times m^2} \times \mathbb{R}^{m^2} \times \mathbb{R}^{m^2} \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}}} r \\
& \text{s.t. } (\delta_1 - \delta_2)P + \delta_1 e_{i_1 i_2} + \delta_2 (\mathbf{p}_{j_1 1}^T \mathbf{z}_1 + \mathbf{p}_{j_2 2}^T \mathbf{z}_2 + bt) \leq r, \quad \forall (i_1, i_2, j_1, j_2) \in \mathcal{S}_1 \times \mathcal{S}_1 \times \mathcal{S}_2 \times \mathcal{S}_2, \\
& \underline{P} \leq P \leq \bar{P}, \quad \mathbf{p}_{l_1 1}^T \mathbf{z}_1 + \mathbf{p}_{l_2 2}^T \mathbf{z}_2 \leq \frac{P}{1 + \theta}, \quad \forall (l_1, l_2) \in \mathcal{S}_2 \times \mathcal{S}_2, \\
& e_{i_1 i_2} + P \leq CVaR_a(X_1 + X_2; \mathcal{Q}_{i_1 i_2}), s_{i_1 i_2} + \frac{1}{1-a} \mathbf{q}_{i_1 i_2}^T \mathbf{w}_{i_1 i_2} \leq e_{i_1 i_2}, \quad \forall (i_1, i_2) \in \mathcal{S}_1 \times \mathcal{S}_1, \\
& \mathbf{0} \leq \mathbf{w}_{i_1 i_2}, \mathbf{y}_1 + \mathbf{y}_2 - s_{i_1 i_2} \mathbf{1} \leq \mathbf{w}_{i_1 i_2}, \quad \forall (i_1, i_2) \in \mathcal{S}_1 \times \mathcal{S}_1, \\
& \mathbf{p}_{j_1 1}^T \mathbf{z}_1 + \mathbf{p}_{j_2 2}^T \mathbf{z}_2 + bt \leq P, \quad \|\hat{\mathbf{Q}}_{j_1 j_2} \mathbf{z}_{12}\| \leq t, \quad \forall (j_1, j_2) \in \mathcal{S}_2 \times \mathcal{S}_2, \\
& \mathbf{y}_k \leq \mathbf{x}_k, \quad \mathbf{z}_k \leq \mathbf{x}_k, \quad \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k \in \{1, 2\},
\end{aligned}$$

where  $m = |\mathcal{S}_1|$  and  $\hat{\mathbf{Q}}_{j_1 j_2}$  is an  $2n \times 2n$  block matrix given by:

$$\hat{\mathbf{Q}}_{j_1 j_2} := \begin{pmatrix} \tilde{\mathbf{Q}}_{j_1 1} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{Q}}_{j_2 2} \end{pmatrix}.$$

## 5. NUMERICAL ILLUSTRATIONS

The current section is dedicated to providing numerical illustrations related to our previous optimisation problems and not only. We first discuss the stability and consistency of the numerical results in Subsection 5.1, while Subsection 5.2 focuses on a case study about capital efficiency within an IG.

**5.1. Stability and consistency.** The main aim is to provide analyses on the stability and consistency of the empirical results for Problem 3.1. We adopt a procedure which is similar to the one presented in Tan and Weng (2014). Recall that when we solve the empirical optimisation model (3.3),  $\mathcal{M} = \{1\}$  and for each given sample  $\mathbf{x}_1 = (x_{11}, \dots, x_{1n})^T$ , we find the optimal solutions  $\mathbf{y}_1^* = (y_{11}^*, \dots, y_{1n}^*)^T$  and  $\mathbf{z}_1^* = (z_{11}^*, \dots, z_{1n}^*)^T$ . The solutions are called stable if we expect that the optimal solutions  $\mathbf{y}_1^*$  and  $\mathbf{z}_1^*$  always have the same functional form for a given  $\mathbf{x}_1$ . Consistency occurs when we expect that  $\mathbf{y}_1^*$  and  $\mathbf{z}_1^*$  converge as the sample size  $n$  increases. Since  $\mathbf{y}_1^* + \mathbf{z}_1^* = \mathbf{x}_1$ , it is only sufficient to study stability and consistency for  $\mathbf{z}_1^*$ . All optimisations are implemented on a desktop with 4 core Intel i7-3770 at 3.40GHz, 16GB RAM, running Linux x64, MATLAB R2014b, CVX 2.1.

The parametrisation for the Problem 3.1 assumes a Pareto loss  $X_1$  with  $\Pr(X_1 > x) = \left(\frac{2,000}{x+2,000}\right)^3$  for all  $x > 0$ ,  $\delta_1 = 0.8$ ,  $\delta_2 = 0.2$ ,  $b = 0.5$ ,  $\underline{P} = 0$  and  $\bar{P} = 5,000$ . The scatter plots of  $\mathbf{z}_1^*$  against  $\mathbf{x}_1$  are shown in Figure 5.1 for Pareto loss samples and various sample sizes  $n$ . Our first observation is that the empirical solutions mimics the functional form of  $z_{1i}^* = x_{1i} \wedge d$ , where  $d$  is a constant for  $i = 1, \dots, n$ . Thus, it is useful to fit  $c(x_{1i} - d_1)_+ \wedge c(d_2 - d_1)$  based on our data  $(x_{1i}, z_{1i}^*)$  for all  $i = 1, \dots, n$  and estimate the three parameters using the *Ordinary Least Square* (OLS) regression. Note that we expect the estimates of  $\hat{c}$  and  $\hat{d}_1$  to be very close to 1 and 0, respectively. Furthermore, the goodness of fit is tested by introducing an admissibility criterion. That is, the estimated pair of  $(\hat{c}, \hat{d}_1, \hat{d}_2)$  is said to be *admissible* whenever

$$\left| z_{1i}^* - \hat{c}(x_{1i} - \hat{d}_1)_+ \wedge \hat{c}(\hat{d}_2 - \hat{d}_1) \right| < \varepsilon \quad \forall i = 1, \dots, n,$$

otherwise, the fitted pair of  $(\hat{c}, \hat{d}_1, \hat{d}_2)$  is said to be inadmissible.

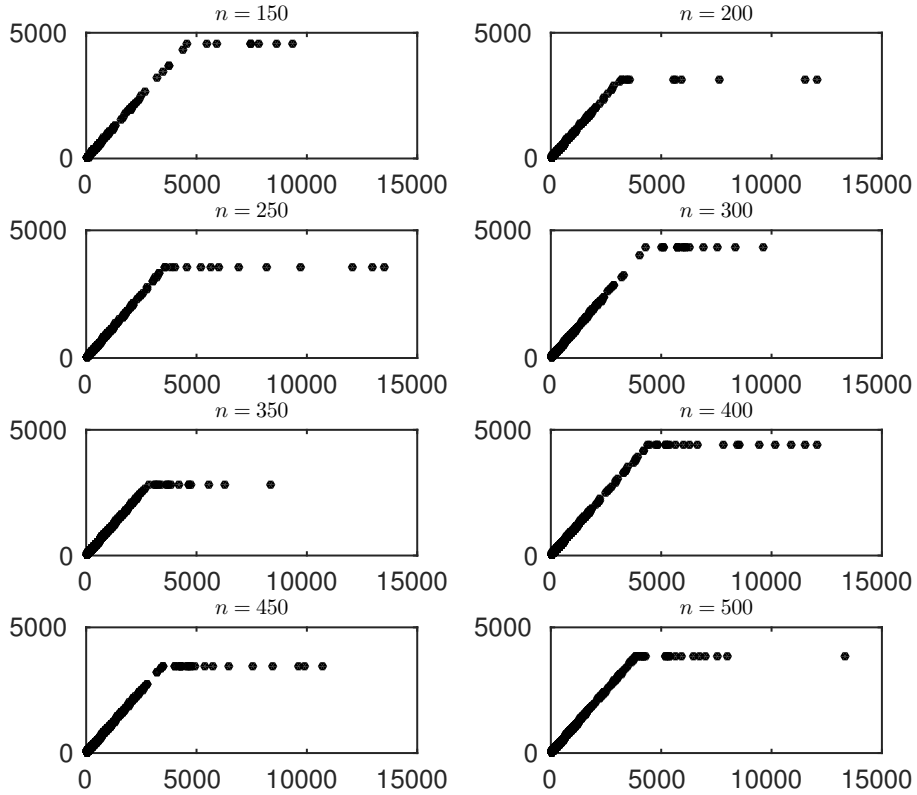


FIGURE 5.1. Empirical solutions of  $\mathbf{z}_1^*$  for Problem 3.1 with Pareto loss and various sample sizes  $n$ .

We draw 1,000 random samples from  $X_1$  for various sample sizes  $\{150, 200, 250, 300, 350, 400, 450, 500\}$  and estimate the three parameters. Boxplots of those estimates are shown in Figure 5.2. The lower and upper edge of each box represents respectively the lower and upper quartile, while the short line inside the box represents the median. The end-point of each whisker represents the minimum or maximum point estimate that is not an outlier, while outliers are marked as crosses inside the plot. Note that an outlier is a point estimate which is away from the box for more than 1.5 times the interquartile range. While Figure 5.2 gives a graphical illustration of  $(\hat{c}, \hat{d}_1, \hat{d}_2)$ , Table 5.1 provides a numerical summary.

The evidence provided by Figure 5.2 and Table 5.1 is overwhelming. First of all, we notice that when  $\varepsilon = 1\%$ , the admissibility is 100% for all sample sizes. If  $\varepsilon$  is reduced to 0.1%, admissibility is still over 92% for all sample sizes. Recall that we expect the estimates of  $\hat{c}$  and  $\hat{d}_1$  to be close to 1 and 0, respectively. The slope estimates are perfect, while a small variation from the null value is observed for  $d_1$ . Thus, we can conclude that our empirical results are very stable in terms of the shape of the functional form of  $z_{i1}^* = x_{i1} \wedge d$ . Also, we notice that the standard errors are continuously decreasing as the sample size  $n$  increases and therefore, we may conclude the consistency.

Finally, the average run-times for solving Problem 3.1 is summarised in Table 5.2.

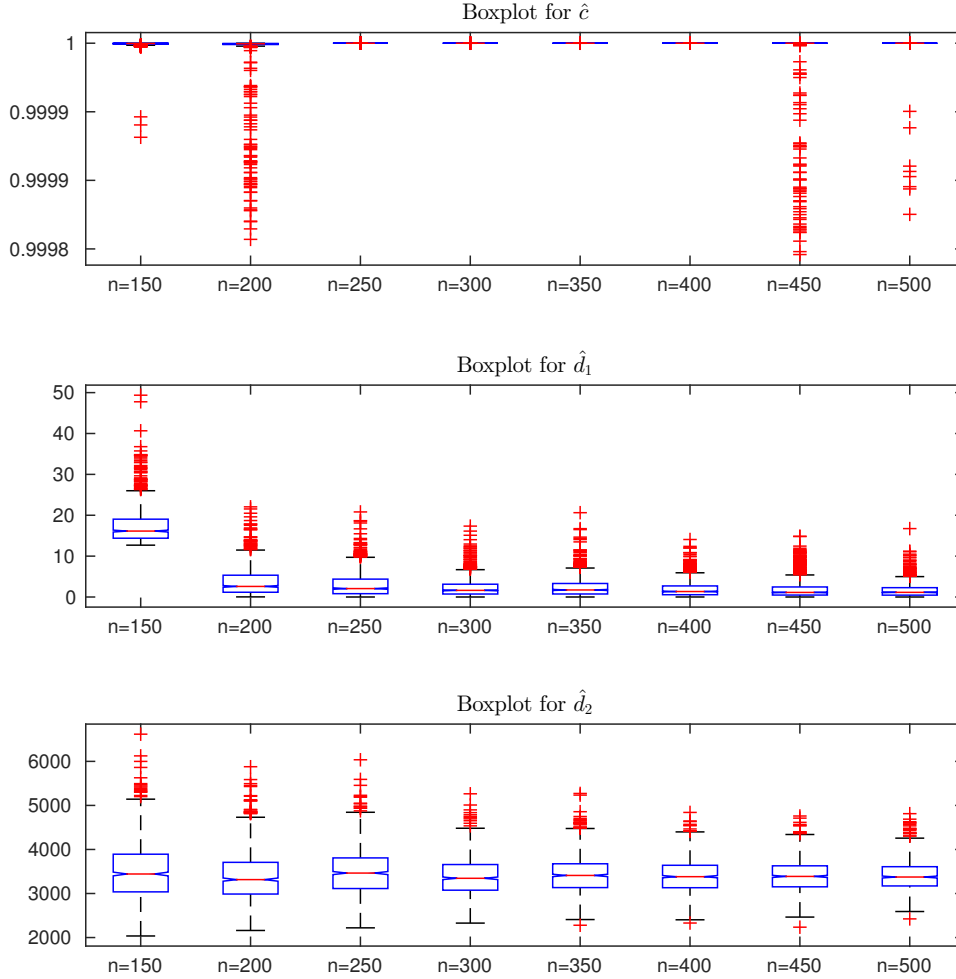


FIGURE 5.2. Boxplots of fitted  $\hat{c}$ ,  $\hat{d}_1$  and  $\hat{d}_2$  for empirical solutions to Problem 3.1 with Pareto loss distribution and various sample sizes  $n$ .

**5.2. Case Study.** Consider an *Insurance group (IG)* with two *legal entities (LE)*, LE1 and LE2. Their total future liabilities are denoted as  $X_1$  and  $X_2$ , respectively. With the opportunity of risk sharing amongst the LEs, LE1 transfers  $I_{21}[X_1]$  to LE2 and retains  $I_{11}[X_1] = X_1 - I_{21}[X_1]$ , while LE2 transfers  $I_{12}[X_2]$  to LE1 and retains  $I_{22}[X_2] = X_2 - I_{12}[X_2]$ . Also, assume that both LE1 and LE2 are CVaR-regulated, and hence, their risk preferences are given by

$$\rho_k(\cdot) := E(\cdot) + \lambda_k(\text{CVaR}_a(\cdot) - E(\cdot)) \quad \text{for } k \in \{1, 2\}.$$

Recall that the CVaR-based risk capital approach is implemented within the Swiss Solvency Test (the regulatory environment in Switzerland) and our setting is realistic for such IG. This setting aims to produce capital efficiency for an IG that operates in various insurance markets that are CVaR-regulated or the regulatory environments that are considered equivalent with the CVaR-based risk capital requirements.

$n$	Admissibility		mean of $\hat{c}$	mean of $\hat{d}_1$	mean of $\hat{d}_2$
	( $\varepsilon = 0.01$ )	( $\varepsilon = 0.001$ )			
150	100%	96.8%	1.00 (0.00)	4.37 (4.39)	3,469.7 (661.41)
200	100%	97.7%	1.00 (0.00)	3.20 (2.94)	3,377.4 (542.61)
250	100%	97.2%	1.00 (0.00)	2.73 (2.67)	3,457.6 (519.35)
300	100%	96.0%	1.00 (0.00)	2.26 (2.18)	3,381.3 (449.74)
350	100%	94.6%	1.00 (0.00)	2.07 (2.12)	3,450.3 (427.89)
400	100%	94.7%	1.00 (0.00)	1.88 (1.87)	3,401.1 (387.71)
450	100%	94.3%	1.00 (0.00)	1.67 (1.63)	3,442.7 (381.75)
500	100%	92.7%	1.00 (0.00)	1.56 (1.58)	3,386.9 (336.63)

TABLE 5.1. Proportion of admissible fitted solutions based on 1,000 samples from a Pareto loss and various sample sizes  $n$ . The last three columns summarise the mean values and standard errors of the estimates.

$n$	150	200	250	300	350	400	450	500
Time (seconds)	0.90	1.24	2.75	3.33	6.14	7.32	12.59	17.36

TABLE 5.2. Average run-time for various sample sizes  $n$

For example, Solvency II (the European Union regulatory regime) is a VaR-based environment, but Swiss Solvency Test and Solvency II are considered equivalent by both parties. An IG has to meet the regulatory capital requirements at the IG level (based on the consolidated balance sheet) and at the individual level. Therefore, we intend to analyse the aggregate capital efficiency by meeting the requirements at the individual level, which is more stringent due to the IG diversification effect that is in place when risk capital is set based on the consolidated balance sheet. For mathematical purposes, the confidence level  $a$  could take different values, but due to the equivalence principle, it would be unrealistic to assume such a scenario. Moreover, the risk preferences from above follow the well-known Cost-of-Capital approach and  $\lambda_1, \lambda_2$  are some parameters (usually greater than 6%) that reflect the cost of using the shareholders' capital that is very sensitive to the tax regime. The Cost-of-Capital is very sensible to the level of taxation and therefore, it is realistic to assume that  $\lambda_1 \neq \lambda_2$ . Thus, our risk transfer problem is given by

$$\begin{aligned} \min_{(I_{11}, I_{12}) \in \mathcal{I} \times \mathcal{I}} \{ & \rho_1(I_{11}[X_1] + I_{12}[X_2]) + \rho_2(X_1 + X_2 - I_{11}[X_1] - I_{12}[X_2]) \}, \\ \text{s.t. } & \rho_1(X_1 - I_{11}[X_1]) \leq C_1, \rho_2(I_{12}[X_2]) \leq C_2, \end{aligned} \quad (5.1)$$

where  $C_1$  and  $C_2$  are the asset fungibility limit for LE1 and LE2 respectively. A standard regulatory constraint is to impose limits for asset transfer, which makes the assets to not be fully fungible, as we actually assume. Note that the risk transfers are not necessarily comonotone as it is assumed in Asimit *et al.* (2013b), where a similar, but unconstrained problem, is investigated. The objective function (5.1) can be further written as

$$E(X_1 + X_2) + \lambda_1(\text{CVaR}_a(T_1) - ET_1) + \lambda_2(\text{CVaR}_a(T_2) - ET_2),$$

with  $T_1 = I_{11}[X_1] + I_{12}[X_2]$  and  $T_2 = X_1 + X_2 - I_{11}[X_1] - I_{12}[X_2]$ . Note that  $E(X_1 + X_2)$  is a constant, and hence solving (5.1) is equivalent to solve the following infinite dimensional optimisation problem:

$$\begin{aligned} \min_{(I_{11}, I_{12}) \in \mathcal{I} \times \mathcal{I}} & \left\{ \lambda_1 (\text{CVaR}_a(T_1) - ET_1) + \lambda_2 (\text{CVaR}_a(T_2) - ET_2) \right\}, \\ \text{s.t.} & \quad \rho_1(X_1 - I_{11}[X_1]) \leq C_1, \quad \rho_2(I_{12}[X_2]) \leq C_2. \end{aligned}$$

This means that after sampling from  $X_1$  and  $X_2$  and using similar notations to those used in Sections 3 and 4, our aim is to solve the following finite dimensional optimisation problem

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n}} & \left\{ \lambda_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2 - s_1 \mathbf{1})_+ \right\} - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2) \right. \\ & \left. + \lambda_2 \min_{s_2 \in \mathbb{R}} \left\{ s_2 + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2 - s_2 \mathbf{1})_+ \right\} - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) \right\} \quad (5.2) \\ \text{s.t.} & \quad \frac{1 - \lambda_1}{n} \mathbf{1}^T \mathbf{z}_1 + \lambda_1 \min_{s_3 \in \mathbb{R}} \left\{ s_3 + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{z}_1 - s_3 \mathbf{1})_+ \right\} \leq C_1 \\ & \quad \frac{1 - \lambda_2}{n} \mathbf{1}^T \mathbf{y}_2 + \lambda_2 \min_{s_4 \in \mathbb{R}} \left\{ s_4 + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{y}_2 - s_4 \mathbf{1})_+ \right\} \leq C_2, \\ & \quad \mathbf{y}_k \leq \mathbf{x}_k, \quad \mathbf{z}_k \leq \mathbf{x}_k, \quad \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k \in \{1, 2\}. \end{aligned}$$

As it has been done earlier, we further reformulate (5.2) in a way to enable efficient computations.

**Theorem 5.1.** *Problem (5.2) is solved by the next LP:*

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, s_1, s_2) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^4 \times \mathbb{R}^4}} & \left\{ M \left( \lambda_1 r_1 + \lambda_2 r_2 - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2) - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) \right) + r_3 + r_4 \right\} \quad (5.3) \\ \text{s.t.} & \quad \frac{1 - \lambda_1}{n} \mathbf{1}^T \mathbf{z}_1 + \lambda_1 r_3 \leq C_1, \quad \frac{1 - \lambda_2}{n} \mathbf{1}^T \mathbf{y}_2 + \lambda_2 r_4 \leq C_2, \\ & \quad s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 \leq r_1, \quad \mathbf{0} \leq \mathbf{u}_1, \quad \mathbf{y}_1 + \mathbf{y}_2 - s_1 \mathbf{1} \leq \mathbf{u}_1, \\ & \quad s_2 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_2 \leq r_2, \quad \mathbf{0} \leq \mathbf{u}_2, \quad \mathbf{z}_1 + \mathbf{z}_2 - s_2 \mathbf{1} \leq \mathbf{u}_2, \\ & \quad s_3 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_3 \leq r_3, \quad \mathbf{0} \leq \mathbf{u}_3, \quad \mathbf{z}_1 - s_3 \mathbf{1} \leq \mathbf{u}_3, \\ & \quad s_4 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_4 \leq r_4, \quad \mathbf{0} \leq \mathbf{u}_4, \quad \mathbf{y}_2 - s_4 \mathbf{1} \leq \mathbf{u}_4, \\ & \quad \mathbf{y}_k \leq \mathbf{x}_k, \quad \mathbf{z}_k \leq \mathbf{x}_k, \quad \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k \in \{1, 2\}, \end{aligned}$$

where  $M$  is a very large number.

One would wonder how large  $M$  should be. If  $X_1$  and  $X_2$  are non-negative liabilities, as we have assumed by now, then if one aims to solve (5.3) accurate to  $n_1$  decimal places, then

$$M > 10^{n_1} \max \left\{ \frac{C_1}{\lambda_1}, \frac{C_2}{\lambda_2} \right\}$$

would be a good choice. For details, see the proof of Theorem 5.1 that is given in Section 7.

We are now ready to elaborate a numerical illustration for a fictitious IG such the future liabilities  $X_1$  and  $X_2$  are LogNormal distributed and their means and squared coefficients of variation are

$$EX_1 = 1,000, \quad EX_2 = 1,500, \quad CV_1^2 = 3, \quad CV_2^2 = 1.5$$

Plausible choices for the cost-of-capital rates are  $\lambda_1 = 10\%$  and  $\lambda_2 = 12\%$ . The choice of  $a = 99\%$  is made as it is the case within the Swiss regulatory environment. We set  $C_k = 40\% \times \rho_k(X_k)$  for  $k \in \{1, 2\}$ . Three dependence models for the liabilities are assumed

- (A) Independence;
- (B) Gaussian copula with parameter  $\rho$ ;
- (C) Pareto copula.

Recall that the Pareto copula is given by

$$C(u, v; \delta) := \left( (1-u)^{-1/\delta} + (1-v)^{-1/\delta} - 1 \right)^{-\delta} + u + v - 1, \quad 0 \leq u, v \leq 1, \delta > 0.$$

(for details, see Nelsen, 2006). The strength of dependence in the upper tail is known as the *upper tail dependence coefficient*. This measure of dependence equals  $2^{-\delta}$  for the Pareto copula and 0 for (A) and (B). Recall that this measure of dependence is an asymptotic value and one should differentiate between (A) and (B). When  $\rho > 0$ , as we assume a positive dependence among the risks, the Gaussian copula allows a stronger dependence for concomitant extreme events than the independence case, and it is even stronger as  $\rho$  increases. The Pareto copula allows a strong dependence, the strongest amongst the three dependence models, which decreases as  $\delta$  increases.

The numerical illustrations are based on a sample size  $n = 9,000$  and we have experienced computational difficulties for samples larger than 10,000, but we believe that a sample of 9,000 is sufficiently informative for our purposes. The optimal solutions of  $\mathbf{y}_1^*$  and  $\mathbf{z}_2^*$  are presented by several scatter plots. Recall that  $\mathbf{y}_1^*$  and  $\mathbf{z}_2^*$  represent the optimal level of risks retained by LE1 and LE2, respectively. The scatter plots for the three dependence models are displayed in Figures 5.3 and 5.4

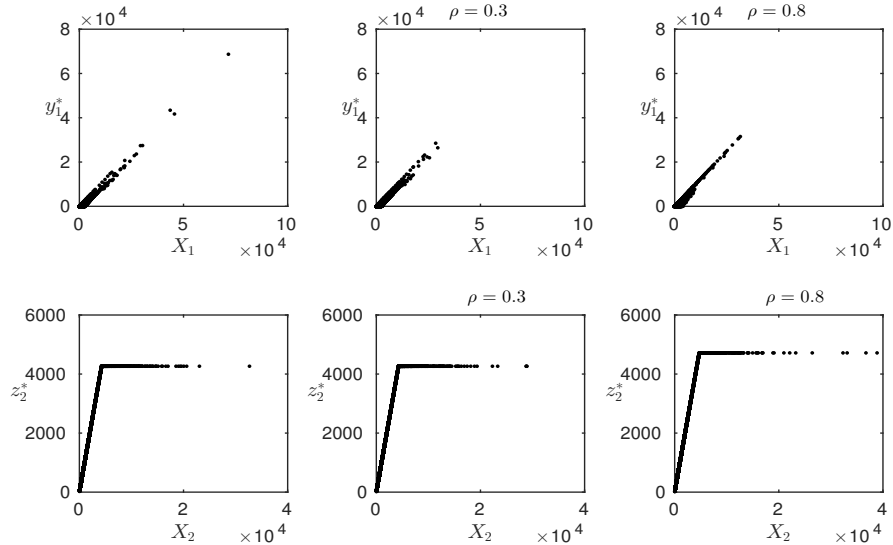


FIGURE 5.3. Empirical solutions of  $\mathbf{y}_1^*$  and  $\mathbf{z}_2^*$  for Model (A) (left column) and Model (B) with  $\rho \in \{0.3, 0.8\}$  (middle and right column).

Figures 5.3 and 5.4 show that  $\mathbf{y}_1^*$  mimics a straight line, which suggests that a proportional risk transfer is optimal when sharing  $X_1$ . When analysing  $\mathbf{z}_2^*$ , a proportional share for  $X_2$  is observed with a limit

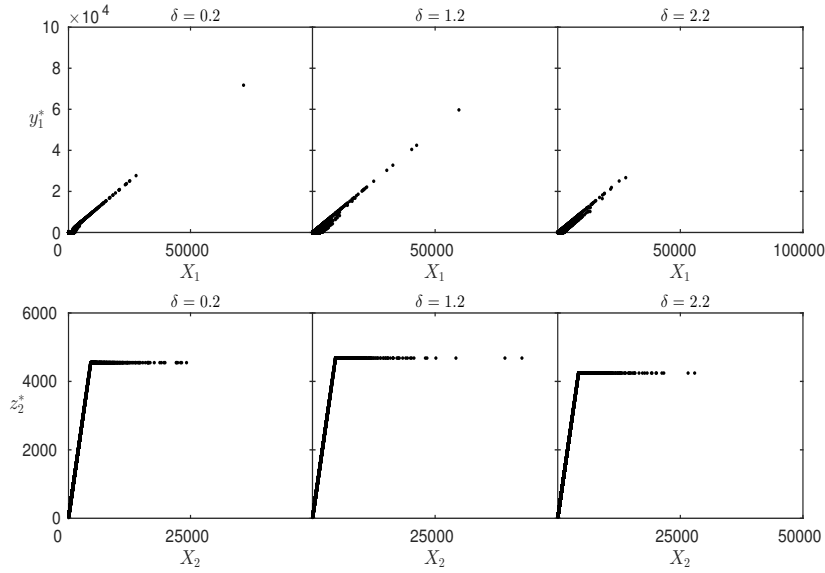


FIGURE 5.4. Empirical solutions of  $\mathbf{y}_1^*$  and  $\mathbf{z}_2^*$  for Model (C) with  $\delta \in \{0.2, 1.2, 2.2\}$ .

imposed by LE2. Therefore, two estimations via OLS (as detailed earlier) are further performed:

$$y_{1i}^* = \theta x_{1i} \quad \text{and} \quad z_{1i}^* = c \min(x_{1i}, d) \quad \text{for all } i \in \{1, \dots, n\}$$

The results are summarised in Table 5.3.

Parameter	Model (A)	Model (B)		Model (C)		
		$\rho = 0.3$	$\rho = 0.8$	$\delta = 0.2$	$\delta = 1.2$	$\delta = 2.2$
$\theta$	0.6635	0.6551	0.7670	0.8437	0.7734	0.6872
$c$	1.000	1.000	1.000	1.000	1.000	1.000
$d$	4,264.8	4,264.2	4,716.4	4,548.0	4,681.9	4,243.7

TABLE 5.3. Estimates for  $\theta$ ,  $c$  and  $d$  for various dependence models.

It becomes clear from Table 5.3 that non-trivial proportional risk sharing is optimal for LE1, while a stop-loss risk sharing is optimal for LE2. The reason that LE2 prefers to release a high layer is due to a higher cost of capital faced in its jurisdiction. LE1 and LE2 retain more risk for themselves (estimates of  $\hat{\theta}$  and  $\hat{d}$  are increasing) if stronger dependence in the upper tail is observed. This is not surprising to observe a reduction of the diversification effect as a result of an increased dependence of concomitant extreme events. The asset fungibility constraints have an important effect over the optimal risk sharing and the immediate effect is that LE1 prefers a proportional “reinsurance” for its risk, which would not have happened otherwise. It is well-known that an unconstrained risk sharing problem does not allow proportional transfers, since only layer “reinsurance” is possible (for details, see Jouini *et al.*, 2008).

## 6. CONCLUSIONS

It is well-known that many actuarial applications involve optimisation and we have tried to provide some of them related to risk sharing that may arise in many forms. Further, we have shown how one may reduce those optimisation problems to efficient and reliable computations. These methods may be extended to many other applications via similar ideas. We have shown that the results are credible by performing stability and consistency checking. The final numerical example is meant to fully explain how a practical problem may be implemented via existing optimisation techniques.

## 7. PROOFS

The proof of Theorems 3.2-3.4, 4.1 and 4.2 are very similar to the proof of Theorem 3.1, and for this reason, only the latter theorem is proved.

**PROOF OF THEOREM 3.1** It is not difficult to find that (3.4) is solved by the following surrogate problem:

$$\begin{aligned} \min_{\substack{(\mathbf{y}_1, \mathbf{z}_1, \mathbf{u}_1, P, t) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}}} & \left\{ (\delta_1 - \delta_2)P + \delta_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right) \right\} \\ \text{s.t. } & \mathbf{0} \leq \mathbf{u}_1, \mathbf{y}_1 - s_1 \mathbf{1} \leq \mathbf{u}_1, \|\mathbf{Qz}_1\| \leq t \\ & \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \leq P, \underline{P} \leq P \leq \bar{P}, \end{aligned} \quad (7.1)$$

$$\begin{aligned} & \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 \right\} + P \leq \text{CVaR}_a(X_1), \\ & b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathbb{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 - c_i \mathbf{1})_+ \right\} \leq P, \\ & \mathbf{y}_1 \leq \mathbf{x}_1, \mathbf{z}_1 \leq \mathbf{x}_1, \mathbf{y}_1 + \mathbf{z}_1 = \mathbf{x}_1. \end{aligned} \quad (7.2)$$

Let  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  be the optimal solution to the surrogate problem. Note that the objective function in the above problem is increasing in  $\mathbf{u}_1$  and  $t$ , and therefore, constraints  $\mathbf{0} \leq \mathbf{u}_1$  and  $\mathbf{y}_1 - s_1 \mathbf{1} \leq \mathbf{u}_1$  ensure that  $\mathbf{u}_1^* = (\mathbf{y}_1^* - s_1 \mathbf{1})_+$ , while  $\|\mathbf{Qz}_1\| \leq t$  implies that  $t^* = \|\mathbf{Qz}_1^*\|$ . As a result,  $(\mathbf{y}_1^*, \mathbf{z}_1^*, P^*)$  is also feasible to the optimisation problem (3.4). If  $(\mathbf{y}_1^*, \mathbf{z}_1^*, P^*)$  is not an optimal solution to (3.4), there must exist another solution  $(\mathbf{y}'_1, \mathbf{z}'_1, P')$  such that

$$\begin{aligned} & (\delta_1 - \delta_2)P' + \delta_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{y}'_1 - s_1 \mathbf{1})_+ \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}'_1 + b \frac{\|\mathbf{Qz}'_1\|}{\sqrt{n-1}} \right) \\ & < (\delta_1 - \delta_2)P^* + \delta_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T (\mathbf{y}_1^* - s_1 \mathbf{1})_+ \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1^* + b \frac{\|\mathbf{Qz}_1^*\|}{\sqrt{n-1}} \right) \\ & = (\delta_1 - \delta_2)P^* + \delta_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1^* \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1^* + \frac{bt^*}{\sqrt{n-1}} \right), \end{aligned}$$

which implies that  $(\mathbf{y}'_1, \mathbf{z}'_1, \mathbf{u}'_1, P', t')$ , with  $\mathbf{u}'_1 = (\mathbf{y}'_1 - s_1 \mathbf{1})_+$  and  $t' = \|\mathbf{Qz}'_1\|$ , is feasible to the surrogate problem and also reduces the objective function value comparing to  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$ . This contradicts the assumption that  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  is the optimal solution to the surrogate problem. Therefore,  $(\mathbf{y}_1^*, \mathbf{z}_1^*, P^*)$  must solve the optimisation problem (3.4) if it solves the surrogate problem.



The final step is to show that the surrogate optimisation problem is equivalent to (3.5). Clearly, the inequality constraints from (7.1) and (7.2) can be rewritten as follows:

$$\max \left\{ P, \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}}, b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathfrak{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 - c_i \mathbf{1})_+ \right\} \right\} \leq P \leq \bar{P}. \quad (7.3)$$

The ordering of the three terms involved in the max operator is crucial for the coming proof.

**Case I:** Assume that the maximum in (7.3) is attained by the third term. We then show that  $S^* := (\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, \mathbf{w}^*, P^*, s_1^*, r_1^*, t^*, \mathbf{e}^*, \mathbf{c}^*)$  solves (3.5) if and only if  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  solves the surrogate optimisation problem. This means that we show the equivalence between the two optimisation problems when the following constraint

$$\max \left\{ P, \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right\} \leq b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathfrak{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 - c_i \mathbf{1})_+ \right\}$$

is added to both problems.

Suppose first that  $S^*$  solves (3.5). The objective function in (3.5) is increasing in  $P$  and  $r_1$  because  $\delta_1 - \delta_2 > 0$  and  $\delta_1 > 0$ . Thus, the inequality constraint  $s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1 \leq r_1$  ensures that

$$s_1^* = \arg \min_{s_1 \in \mathfrak{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1^* \right\} \quad \text{and} \quad s_1^* + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1^* = r_1^*,$$

while the constraint

$$b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i e_i \leq P \quad \text{for all } i \in \{1, \dots, n-1\}$$

becomes an identity at the optimum, forcing  $\mathbf{e}$  to be as small as possible. Further, all constraints

$$c_i + \frac{1}{n-i} \mathbf{1}^T \mathbf{w}_i \leq e_i, \quad \mathbf{0} \leq \mathbf{w}_i, \quad \mathbf{z}_1 - c_i \mathbf{1} \leq \mathbf{w}_i$$

became identities (due to similar arguments) implying that for all  $i \in \{1, \dots, n-1\}$  we have that

$$c_i^* = \arg \min_{c_i \in \mathfrak{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1^* - c_i^* \mathbf{1})_+ \right\}, \quad (\mathbf{z}_1^* - c_i^* \mathbf{1})_+ = \mathbf{w}_i^*,$$

$$e_i^* = c_i^* + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1^* - c_i^* \mathbf{1})_+ \quad \text{and} \quad b_0 \mathbf{1}^T \mathbf{z}_1^* + \sum_{i=1}^{n-1} b_i e_i^* = P^*.$$

Consequently,  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  is also feasible to the surrogate problem. Now, if  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  does not solve the surrogate problem, then there must exist a feasible  $(\mathbf{y}'_1, \mathbf{z}'_1, \mathbf{u}'_1, P', t')$  such that

$$\begin{aligned} & (\delta_1 - \delta_2) P' + \delta_1 \min_{s_1 \in \mathfrak{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}'_1 \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}'_1 + \frac{bt'}{\sqrt{(n-1)}} \right) \\ & < (\delta_1 - \delta_2) P^* + \delta_1 \min_{s_1 \in \mathfrak{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1^* \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1^* + \frac{bt^*}{\sqrt{(n-1)}} \right) \\ & = (\delta_1 - \delta_2) P^* + \delta_1 r_1^* + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1^* + \frac{bt^*}{\sqrt{(n-1)}} \right), \end{aligned}$$

where the last step is a consequence of the fact that  $S^*$  solves (3.5) and it was shown earlier. As a result, the tuple  $(\mathbf{z}'_1, P', r'_1, t')$  reduces, as compared to  $(\mathbf{z}_1^*, P^*, r_1^*, t^*)$ , the objective value in (3.5), where  $r'_1 = \min_{s_1 \in \mathfrak{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}'_1 \right\}$ , which contradicts the assumption that  $S^*$  is the optimal solution to (3.5). In summary, if  $S^*$  solves (3.5), then it must also solve the surrogate problem.

We now prove the converse statement. Suppose that  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  solves the surrogate optimisation problem, but its counterpart  $S^*$  does not solve (3.5). Therefore, there exists a feasible  $S' := (\mathbf{y}'_1, \mathbf{z}'_1, \mathbf{u}'_1, \mathbf{w}', P', s'_1, r'_1, t', \mathbf{e}', \mathbf{c}')$  that solves problem (3.5). It follows directly that  $(\mathbf{y}'_1, \mathbf{z}'_1, \mathbf{u}'_1, P', t')$  is also feasible to the surrogate problem. Since  $S'$  solves (3.5), but  $S^*$  is not the optimal solution in (3.5), we have that

$$(\delta_1 - \delta_2)P' + \delta_1 r'_1 + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}'_1 + \frac{bt'}{\sqrt{n-1}} \right) < (\delta_1 - \delta_2)P^* + \delta_1 r_1^* + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1^* + \frac{bt^*}{\sqrt{n-1}} \right) \quad (7.4)$$

where  $r_1^* = \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1^* \right\}$ . The identity could be proved via similar arguments given earlier, since the objective function from (3.5) is increasing in  $r_1$  as  $\delta_1 > 0$ . Since  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  solves the surrogate optimisation problem, we may write

$$\begin{aligned} & (\delta_1 - \delta_2)P^* + \delta_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_1^* \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt^*}{\sqrt{n-1}} \right) \\ & \leq (\delta_1 - \delta_2)P' + \delta_1 \min_{s_1 \in \mathbb{R}} \left\{ s_1 + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}'_1 \right\} + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}'_1 + \frac{bt'}{\sqrt{n-1}} \right) \\ & \leq (\delta_1 - \delta_2)P' + \delta_1 r'_1 + \delta_2 \left( \frac{1}{n} \mathbf{1}^T \mathbf{z}'_1 + \frac{bt'}{\sqrt{n-1}} \right), \end{aligned}$$

where the last step is due to the fact that  $S'$  is feasible in (3.5). This contradicts (7.4), which justifies the converse statement. Thus, the proof of Case I is complete.

**Case II:** Assume that the maximum in (7.3) is not attained by the third term. That is,

$$\max \left\{ \frac{P}{n}, \frac{1}{n} \mathbf{1}^T \mathbf{z}_1 + \frac{bt}{\sqrt{n-1}} \right\} > b_0 \mathbf{1}^T \mathbf{z}_1 + \sum_{i=1}^{n-1} b_i \min_{c_i \in \mathbb{R}} \left\{ c_i + \frac{1}{n-i} \mathbf{1}^T (\mathbf{z}_1 - c_i \mathbf{1})_+ \right\}.$$

We then show that  $S^* := (\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, \mathbf{w}^*, P^*, s_1^*, r_1^*, t^*, \mathbf{e}^*, \mathbf{c}^*)$  solves (3.5) if and only if  $(\mathbf{y}_1^*, \mathbf{z}_1^*, \mathbf{u}_1^*, P^*, t^*)$  solves the surrogate optimisation problem, by including the above constraint. The proof is a tad simpler, since the constraint (7.2) becomes redundant, and it can be achieved by applying the same reasoning as used in Case I. This concludes the proof.

**PROOF OF THEOREM 5.1** First of all, we show that the optimisation problem (5.2) is solved by the following surrogate problem

$$\begin{aligned} & \min_{\substack{(\mathbf{y}_1, \mathbf{y}_2, \mathbf{z}_1, \mathbf{z}_2, \mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4) \in \\ \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n}} \left\{ M \left( \lambda_1 D_1 + \lambda_2 D_2 - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2) - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2) \right) + D_3 + D_4 \right\} \\ & \text{s.t. } \frac{1 - \lambda_1}{n} \mathbf{1}^T \mathbf{z}_1 + \lambda_1 D_3 \leq C_1, \quad \frac{1 - \lambda_2}{n} \mathbf{1}^T \mathbf{y}_2 + \lambda_2 D_4 \leq C_2, \\ & \mathbf{0} \leq \mathbf{u}_1, \mathbf{y}_1 + \mathbf{y}_2 - s_1 \mathbf{1} \leq \mathbf{u}_1, \mathbf{0} \leq \mathbf{u}_2, \mathbf{z}_1 + \mathbf{z}_2 - s_2 \mathbf{1} \leq \mathbf{u}_2, \quad (7.5) \\ & \mathbf{0} \leq \mathbf{u}_3, \mathbf{z}_1 - s_3 \mathbf{1} \leq \mathbf{u}_3, \mathbf{0} \leq \mathbf{u}_4, \mathbf{y}_2 - s_4 \mathbf{1} \leq \mathbf{u}_4, \quad (7.6) \\ & \mathbf{y}_k \leq \mathbf{x}_k, \mathbf{z}_k \leq \mathbf{x}_k, \mathbf{y}_k + \mathbf{z}_k = \mathbf{x}_k, \quad k = 1, 2, \end{aligned}$$

where

$$D_i = \min_{s_i \in \mathbb{R}} \left\{ s_i + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_i \right\}, \quad \text{for all } i \in \{1, 2, 3, 4\}.$$

Note that the objective function in our surrogate problem is increasing in  $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$  and  $\mathbf{u}_4$ , and therefore, constraints in (7.5) ensure that  $\mathbf{u}_1 = (\mathbf{y}_1 + \mathbf{y}_2 - s_1 \mathbf{1})_+$  and  $\mathbf{u}_2 = (\mathbf{z}_1 + \mathbf{z}_2 - s_2 \mathbf{1})_+$  at optimum. Similarly,

constraints in (7.6) imply that  $\mathbf{u}_3 = (\mathbf{z}_1 - s_3 \mathbf{1})_+$  and  $\mathbf{u}_4 = (\mathbf{y}_2 - s_4 \mathbf{1})_+$ . Also,  $M$  is a very large number and acts as a penalty in the objective function of the surrogate problem, which forces to minimise the term

$$\lambda_1 D_1 + \lambda_2 D_2 - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}_1 + \mathbf{y}_2) - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}_1 + \mathbf{z}_2),$$

but maintaining to minimise  $D_3 + D_4$  as well.<sup>2</sup> Thus, any optimal solution of the surrogate optimisation problem solves (5.2).

The final step is to show that the surrogate optimisation problem is equivalent to the problem displayed in (5.3). In particular, we justify that  $S^* := (\mathbf{y}_1^*, \mathbf{y}_2^*, \mathbf{z}_1^*, \mathbf{z}_2^*, \mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_3^*, \mathbf{u}_4^*, \mathbf{s}^*, \mathbf{r}^*)$  solves the optimisation problem (5.3) if and only if  $T^* := (\mathbf{y}_1^*, \mathbf{y}_2^*, \mathbf{z}_1^*, \mathbf{z}_2^*, \mathbf{u}_1^*, \mathbf{u}_2^*, \mathbf{u}_3^*, \mathbf{u}_4^*)$  solves the surrogate problem.

Suppose now that  $S^*$  solves (5.3), but  $T^*$  is not an optimal solution of the surrogate problem. The objective function in (5.3) is increasing in  $r_1, r_2, r_3$  and  $r_4$  as  $\lambda_1$  and  $\lambda_2$  are positive. Thus, the inequality constraint  $s_i + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_i \leq r_i$  ensures that ,

$$s_i^* = \arg \min_{s_i \in \mathbb{R}} \left\{ s_i + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_i^* \right\} \quad \text{and} \quad D_i^* = s_i^* + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}_i^* \quad \text{for all } i \in \{1, 2, 3, 4\}. \quad (7.7)$$

Thus,  $T^*$  is also feasible to the surrogate problem. Recall that  $T^*$  does not solve the surrogate problem and therefore, there must exist another feasible solution  $(\mathbf{y}'_1, \mathbf{y}'_2, \mathbf{z}'_1, \mathbf{z}'_2, \mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3, \mathbf{u}'_4)$  such that

$$\begin{aligned} & M \left( \lambda_1 D'_1 + \lambda_2 D'_2 - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}'_1 + \mathbf{y}'_2) - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}'_1 + \mathbf{z}'_2) \right) + D'_3 + D'_4 \\ & < M \left( \lambda_1 D_1^* + \lambda_2 D_2^* - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}_1^* + \mathbf{y}_2^*) - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}_1^* + \mathbf{z}_2^*) \right) + D_3^* + D_4^* \\ & = M \left( \lambda_1 r_1^* + \lambda_2 r_2^* - \frac{\lambda_1}{n} \mathbf{1}^T (\mathbf{y}_1^* + \mathbf{y}_2^*) - \frac{\lambda_2}{n} \mathbf{1}^T (\mathbf{z}_1^* + \mathbf{z}_2^*) \right) + r_3^* + r_4^*, \end{aligned} \quad (7.8)$$

where the last step is due to equation (7.7) and

$$D'_i = s'_i + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}'_i \quad \text{and} \quad s'_i = \arg \min_{s_i \in \mathbb{R}} \left\{ s_i + \frac{1}{n(1-a)} \mathbf{1}^T \mathbf{u}'_i \right\} \quad \text{for all } i \in \{1, 2, 3, 4\}.$$

By taking  $r'_i = D'_i$ , we get that  $(\mathbf{y}'_1, \mathbf{y}'_2, \mathbf{z}'_1, \mathbf{z}'_2, \mathbf{u}'_1, \mathbf{u}'_2, \mathbf{u}'_3, \mathbf{u}'_4, \mathbf{r}', \mathbf{s}')$  is feasible in optimisation problem (5.3), which improves the objective function of (5.3), as pointed out in (7.8). This contradicts our initial assumption that  $S^*$  solves (5.3). Thus, if  $S^*$  solves (5.3), then  $T^*$  solves the surrogate problem.

The converse part can be argued in a similar manner as before, which explains in full the equivalence between the surrogate problem and the optimisation problem displayed in (5.3). This completes our proof.

#### ACKNOWLEDGMENTS

The authors would like to thank the Committee on Knowledge Extension Research (CKER), the Society of Actuaries (SOA) and the Casualty Actuarial Society (CAS) for their financial support.

<sup>2</sup>If one aims to solve (5.2) accurate to  $n_1$  decimal places and we know that  $D_3 + D_4$  is no larger than  $10^{n_2}$ , then any  $M > 10^{n_1 + n_2}$  would be a good choice.

## REFERENCES

- Aase, K.K. 2002. “Perspectives of Risk Sharing, Scandinavian Actuarial”, *Scandinavian Actuarial Journal*, Issue 2, 73–128.
- Acerbi, C. and Tasche, D. 2002. “On the Coherence of Expected Shortfall,” *Journal of Banking & Finance*, 26(7), 1487–1503.
- Adler, I. and Alizadeh, F. 1995. “Primal-Dual Interior Point Algorithms for Convex Quadratically Constrained and Semidefinite Optimization Problems”, *Technical Report RRR-111-95, RUTCOR, Rutgers University*.
- Alizadeh, F. and Goldfarb, D. 2003. “Second-order cone optimization”, *Mathematical Programming Ser. B*, 95, 3–51.
- Alizadeh, F., Haeberly, J.P. and Overton, M.L. 1998. “Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results”, *SIAM Journal on Optimization*, 8(3), 746–768.
- Alizadeh, F. and Schmieta, S.H. 1997. “Optimization with Semidefinite, Quadratic and Linear Constraints”, *Technical Report RRR-23-97, RUTCOR, Rutgers University*.
- Arrow, K.J. 1963. “Uncertainty and the Welfare Economics of Medical Care,” *American Economic Review*, 53(5), 941–973.
- Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. 1999. “Coherent Measures of Risk,” *Mathematical Finance*, 9(3), 203–228.
- Asimit, A.V., Badescu, A.M. and Cheung, K.C. 2013a. “Optimal Reinsurance in the Presence of Counterparty Default Risk”, *Insurance: Mathematics and Economics*, 53(3), 690–697.
- Asimit, V., Badescu, A.M., Haberman, S. and Kim, E.-S. 2016. “Efficient Risk Allocation within a Non-life Insurance Group under Solvency II Regime”, *Insurance: Mathematics and Economics*, 66, 69–76.
- Asimit, A.V., Badescu, A.M. and Tsanakas, A. 2013b. “Optimal Risk Transfers in Insurance Groups,” *European Actuarial Journal*, 3, 159–190.
- Asimit, A.V., Badescu, A.M. and Verdonck, T., 2013c. “Optimal Risk Transfer under Quantile-based Risk Measures,” *Insurance: Mathematics and Economics*, 53(1), 252–265.
- Asimit, A.V., Bignozzi, V., Cheung, K.C., Hu, J. and Kim, E.-S. 2017. “Robust and Pareto Optimality of Insurance Contracts”. Available at <http://ssrn.com/abstract=2834079>
- Balbás, A., Balbás, B., Balbás, R. and Heras, A. 2015. “Optimal Reinsurance under Risk and Uncertainty,” *Insurance: Mathematics and Economics*, 60, 61–74.
- Bernard, C. and Ludkovski, M. 2012. “Impact of Counterparty Risk on the Reinsurance Market,” *North American Actuarial Journal*, 16(1), 87–111.
- 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.
- Cai, J., Lemieux, C. and Liu, F.D. 2014. “Optimal Reinsurance with Regulatory Initial Capital and Default Risk,” *Insurance: Mathematics and Economics*, 57, 13–24.

- Cai, J. and Tan, K.S. 2007. "Optimal Retention for a Stop-loss Reinsurance under the VaR and CTE Risk Measures," *Astin Bulletin*, 37(1), 93–112.
- Cai, J., 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.
- Cai, J. and Wei, W. 2012. "Optimal Reinsurance with Positively Dependent Risks," *Insurance: Mathematics and Economics*, 50(1), 57–63.
- Cheung, K.C. 2010. "Optimal Reinsurance Revisited - A Geometric Approach ," *Astin Bulletin*, 40(1), 221–239.
- Cui, W., Yang, J. and Wu, L. 2013. "Optimal Reinsurance Minimising the Distortion Risk Measure under General Reinsurance Premium Principles," *Insurance: Mathematics and Economics*, 53(1), 74–85.
- Fischer, T. (2003). "Risk Capital Allocation by Coherent Risk Measures Based on One-sided Moments," *Insurance: Mathematics and Economics*, 32(1), 135–146.
- Gajek, L. and Zagrodny, D. 2000. "Insurer's Optimal Reinsurance Strategies," *Insurance: Mathematics and Economics*, 27(1), 105–112.
- Jouini, E., Schachermayer, W. and Touzi, N. 2008. "Optimal Risk Sharing for Law-invariant Monetary Utility Functions," *Mathematical Finance*, 18(2), 269–292.
- Kaluszka, M. 2001. "Optimal Reinsurance under Mean-variance Premium Principles," *Insurance: Mathematics and Economics*, 28(1), 61–67.
- Kaluszka, M. 2004. "Mean-Variance Optimal Reinsurance Arrangements," *Scandinavian Actuarial Journal*, Issue 1, 28–41.
- Lobo, M.S., Vandenberghe, L., Boyd, S. and Lebret, H. 1998. "Applications of the second-order cone programming," *Linear Algebra and Applications*, 284, 193–228.
- Monteiro, R.D.C. and Tsuchiya, T. 2000. "Polynomial convergence of primal-dual algorithms for the second-order cone program based on the MZ-family of directions," *Mathematical Programming*, 88, 61–83.
- Nelsen, R. B. 2006. *An Introduction to Copulas*. Springer-Verlag, New York.
- Nesterov, Y.E. and Todd, M.J. 1997. "Self-scaled barriers and interior-point methods for convex programming", *Mathematics of Operations Research*, 22, 1–42.
- Nesterov, Y.E. and Todd, M.J. 1998. "Primal-dual interior-point methods for self-scaled cones", *SIAM Journal on Optimization*, 8, 324–364.
- Rockafeller, R.T. and Uryasev, S. 2000. "Optimization of Conditional Value-at-Risk," *Journal of Risk*, 2, 21–41.
- Tan, K. and Weng, C., 2014. "Empirical Approach for Optimal Reinsurance Design," *North American Actuarial Journal*, 18(2), 315–342.
- Tao, P.D. and An, L.T.H. 1998. "A D.C. Optimization Algorithm for Solving the Trust-Region Subproblem," *SIAM Journal of Optimization*, 8(2), 476–505.
- Wang, S.S., Young, V.R. and Panjer, H.H. 1997. "Axiomatic Characterization of Insurance Prices," *Insurance: Mathematics and Economics*, 21(2), 173–183.
- Wozabal, D., Hochreiter, R. and Pflug, G. 2010. "A D.C. Formulation of Value-at-Risk Constrained Optimization," *Optimization*, 59(3), 377–400.