Finance & Stochastics manuscript No.
(will be inserted by the editor)

Risk bounds for factor models

**Carole Bernard · Ludger Rüschendorf ·
Steven Vanduffel · Ruodu Wang**

Received: date / Accepted: date

Carole Bernard acknowledges support from the Humboldt Foundation and from the Project on Systemic Risk funded by the GRI in financial services and by the Louis Bachelier Institute. Ludger Rüschendorf acknowledges support from DFG grant RU 704/11-1. Steven Vanduffel acknowledges support from the Chair Stewardship of Finance and FWO. Ruodu Wang acknowledges support from NSERC (RGPIN-435844-2013). We thank Edgars Jakobsons from ETH Zurich for his interesting comments on an earlier draft. The authors thank the Editor, an Associate Editor and the two reviewers for their careful reading of the paper and for their many valuable comments and suggestions, which helped to improve the paper.

C. Bernard
Grenoble Ecole de Management, 12 Rue Pierre Sémard, 38000 Grenoble, France.
Vrije Universiteit Brussel, Pleinlaan 2, 1050 Elsene, Belgium.
E-mail: carole.bernard@grenoble-em.com

L. Rüschendorf
University of Freiburg, Eckerstraße 1, 79104 Freiburg, Germany.
E-mail: ruschen@stochastik.uni-freiburg.de

S. Vanduffel
Corresponding author. Vrije Universiteit Brussel, Pleinlaan 2, 1050 Bruxelles, Belgium.
E-mail: steven.vanduffel@vub.ac.be

R. Wang
University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L3G1, Canada.
E-mail: wang@uwaterloo.ca.

Abstract Recent literature has investigated the risk aggregation of a portfolio $X = (X_i)_{1 \le i \le n}$ under the sole assumption that the marginal distributions of the risks X_i are specified but not their dependence structure. There exists a range of possible values for any risk measure of $S = \sum_{i=1}^n X_i$ and the dependence uncertainty spread, as measured by the difference between the upper bound and the lower bound on these values, is typically very wide. Obtaining bounds that are more practically useful requires additional information on dependence.

Here, we study a partially specified factor model in which each risk X_i has a known joint distribution with the common risk factor Z, but we dispense with the conditional independence assumption that is typically made in fully specified factor models. We derive easy-to-compute bounds on risk measures such as Value-at-Risk (VaR) and law-invariant convex risk measures (e.g., Tail Value-at-Risk (TVaR)) and demonstrate their asymptotic sharpness. We show that the dependence uncertainty spread is typically reduced substantially and that, contrary to the case in which only marginal information is used, it is *not* necessarily larger for VaR than for TVaR.

Keywords factor models · risk aggregation · dependence uncertainty · Value-at-Risk

Mathematics Subject Classification (2010) MSC 97K50 · MSC 60E05 · MSC 60E15 · MSC 62P05 · JEL C02 · JEL C10 · JEL G11

1 Introduction

The primary objective of this paper is to study the range of possible values of a risk measure of an aggregate risk $S = \sum_{i=1}^n X_i$ under model uncertainty, i.e., in a context in which the joint distribution of the vector $X = (X_i)_{1 \le i \le n}$ is not perfectly known. Throughout, we refer to the bounds on such ranges of values as **risk bounds.** The difference between the upper bound and the lower bound is known as **the dependence uncertainty spread** and serves as a measure of model risk.

In a number of recent papers, risk bounds for S have been derived under the assumption that only marginal information is available, i.e., the distributions $F_i,\ i = 1, \dots, n$ of the $X_i,\ i = 1, \dots, n$ are known, but not their dependence. Due to the missing information regarding the dependence among the $X_1, \dots, X_n$, there is a wide range of possible values for a risk measure of S. Typically, one observes a huge dependence uncertainty spread, showing that models that are based on dependence assumptions that cannot be justified (by data) are not reliable for portfolio risk assessment. For a recent discussion and numerical illustrations of these risk bounds and their consequences in risk management, we refer to Embrechts et al. (2014) and the references therein. In this regard, we point out that the study of risk bounds relates to the important question of choice of a risk measure, which involves issues such as impact on risk taking, robustness and backtesting properties. For a select sampling of the relevant literature we point to the work of Föllmer and Schied (2004), Cont et al. (2010), Daníelsson et al. (2005), Emmer et al. (2014), Gneiting (2011), Jorion (2006), Krätschmer et al. (2012, 2014) and Embrechts et al. (2015).

Recent work on risk bounds concentrates on taking into account additional information on the dependence among the risks $X_1, \ldots, X_n$. Perspectives that researchers have introduced in this regard include the consideration of higher dimensional marginals (see e.g., Puccetti and Rüschendorf (2012)), the inclusion of positive or negative dependence information (Bignozzi et al. (2015)), the influence of the correlations among the X_i (Bernard et al. (2015)), the knowledge of the moments of the sum (Bernard et al. (2016, 2017)) and the influence of precise information about the joint distribution in some part of the space (Bernard and Vanduffel (2015)). These studies provide evidence that adding dependence information makes it possible to reduce the dependence uncertainty spread significantly. In this paper, we consider dependence information through the structural assumption of a factor model and study risk bounds in this context.

Factor models appear to offer a useful device for modeling multivariate distributions in various disciplines, including statistics, econometrics and finance. In particular, they play a central role in asset pricing (Fama and French (1993), Engle et al. (1990)) and are used in monitoring mutual fund performance (Carhart (1997)) and in portfolio optimization (Santos et al. (2013)). In risk management, where they also drive regulatory capital requirements (Gordy (2003)), they constitute the industry standard for the evaluation of credit risk (Gordy (2000)). Specifically, the multivariate normal mean-variance mixture model can be seen as a factor model and generates many of the standard and well-established distributions in quantitative finance, such as the Variance Gamma, Hyperbolic and Normal Inverse Gaussian distributions. Importantly, economic theories such as as the Arbitrage Pricing Theory (APT) (Ross (1976)), the CAPM (Sharpe (1964)) and the rank theory of consumer demands (Lewbel (1991)) are based explicitly on factor models.

In a factor model, the risks X_i are expressed in a functional form as

$$X_i = f_i(Z, \varepsilon_i), \quad 1 \le i \le n, \tag{1.1}$$

where ε_i are idiosyncratic risk components and Z is a common risk factor taking value in a set $D \subset \mathbb{R}^d$. The typical assumptions for factor models, as in (1.1), are that the factor Z has known distribution and that, conditionally on $Z = z \in D$, the risks X_i are independent (i.e., the ε_i are independent) with known (conditional) distribution $F_{i|z}$. However, the assumption that the risks are conditionally independent given the factor is challenging and often appears to be made in an ad-hoc fashion and not grounded in data or statistics (Connor and Korajczyk (1993)). The results of our paper make it possible to assess the sensitivity of any factor model with respect to this assumption. In the context of asset pricing, Chamberlain and Rothschild (1982) and Ingersoll (1984) relax the conditional independence assumption slightly and develop so-called approximate factor models. In this paper we dispense with the conditional independence assumption and investigate the consequences of doing so on risk bounds, leaving other possible applications for future work. Clearly, there might be further possible model risk due to misspecification of the law of the factor or of the conditional laws of the risks given the factors or as a result of further assumptions such as the number of factors. In this paper, however, we concentrate on the risk contribution arising from possible departures from the conditional independence assumption.

When we dispense with the assumption of conditional independence among the individual risks X_i, $i = 1, \ldots, n$, their joint distribution is no longer specified. However, the joint distributions H_i of (X_i, Z), and thus also the marginal distribution F_i of X_i, $i = 1, \ldots, n$, are known. We label this setting **a partially specified factor model.** As compared to considering only the information on the marginal distributions of $X_1, \ldots, X_n$, using the additional information on the common risk factor Z leads to improved risk bounds (smaller dependence uncertainty spread) when assessing the risk of the aggregated portfolio $S = \sum_{i=1}^n X_i$.

In Section 2, following this introduction, we formally introduce the partially specified factor model. By representing X as a mixture, we obtain sharp upper and lower bounds for the tail probabilities and thus, by inversion, for the VaR as well. Unfortunately, the evaluation of these bounds typically poses a considerable challenge. Hence, we derive a more explicit mixture representation for the sharp VaR bound that will be the basis for obtaining VaR bounds in Section 4 that are asymptotically sharp and that can be practically evaluated.

First, however, in Section 3, we study sharp upper and lower bounds for law-invariant convex risk measures (including the Tail-Value-at-Risk (TVaR) as a special case). These bounds follow from the availability of the largest and smallest elements with respect to convex order for the distribution of S. The largest elements are attained by a dependence structure of comonotonicity conditionally on Z (see Dhaene et al. (2006) for an overview on comonotonicity) and, when the distributions satisfy suitable assumptions, the smallest elements are attained by a dependence structure of joint mixes (Wang and Wang (2016)) conditionally on Z. We obtain explicit convex lower bounds for S in the context of a mean-variance mixture model.

In Section 4, based on the mixing type formula for the sharp VaR bounds, we obtain approximations of the VaR bounds by means of TVaR-based estimates. This procedure leads to greatly simplified formulas that are well suited to numerical evaluation. We demonstrate that the TVaR-based approximations are asymptotically sharp. These results extend those of Puccetti and Rüschendorf (2014) and Embrechts et al. (2015), where only the marginal information was used. Furthermore, Embrechts et al. (2015) show that in this setting the VaR of large portfolios (asymptotically) exhibits a larger dependence uncertainty spread than the TVaR. These authors use this feature as an argument in support of the use of TVaR in risk management. By contrast, it cannot be expected that such a result will hold for a partially specified factor model, and we provide an example to illustrate this point. By supplementing the partially specified factor model with (conditional) variance information, we derive further improved bounds, and we discuss their asymptotic sharpness.

Finally, in Section 5 we assess the model uncertainty of a credit risk portfolio that is modeled using a Bernoulli mixture model (KMV model). This application and other examples that we provide in the text illustrate the results and establish a clear impression of the range of reduction of dependence uncertainty that one can obtain by using factor information.

2 Risk factor models and VaR bounds

2.1 Partially specified risk factor models

Let $(\Omega, \mathcal{A}, P)$ be an atomless probability space and $\mathcal{X}$ be a set of real-valued random variables (rvs) on $(\Omega, \mathcal{A}, P)$. We take $\mathcal{X} = L^0$, the set of all rvs in this section and $\mathcal{X} = L^1$ in the sections that follow. A risk measure ϱ is a mapping from $\mathcal{X}$ to $(-\infty, \infty]$. In this paper we consider only law-invariant risk measures. Let $Z \in \mathcal{X}^d$ be a random vector with essential support $D \subset \mathbb{R}^d$. We refer to Z as a **risk factor** and denote its distribution by G.

Let $H = (H_i)_{1 \le i \le n}$ be a vector of $(1 + d)$-variate distributions and define the **partially specified factor model**

$$A(H) = \{X \in \mathcal{X}^n \,:\, (X_i, Z) \sim H_i,\ 1 \le i \le n\}$$

as the set of random vectors $X = (X_i)_{1 \le i \le n}$ such that for each $i = 1, \dots, n$, (X_i, Z) has joint distribution (function) H_i. For each $i = 1, \dots, n$ X_i has distribution F_i and conditional distribution $F_{i|z}$ given $Z = z$, $z \in D$. However, the distribution of X is not completely specified. In this paper, we aim at determining (sharp) upper and lower bounds on $\varrho(S)$ where ϱ is some risk measure, $S = \sum_{i=1}^n X_i$ and $X \in A(H)$. Specifically, we consider the problems

$$\overline{\varrho}^f = \sup\{\varrho(S) \,:\, X \in A(H)\} \text{ and } \underline{\varrho}^f = \inf\{\varrho(S) \,:\, X \in A(H)\}. \tag{2.1}$$

In the remainder of the paper we often refer to the partially specified risk factor model as *the constrained setting.*

Write $F = (F_i)_{1 \le i \le n}$. In comparison to the partially specified factor model, the **model with marginal information** (only) is defined as

$$A_1(F) = \{X \in \mathcal{X}^n \,:\, X_i \sim F_i,\ 1 \le i \le n\},$$

and in this setting one considers the problems

$$\overline{\varrho} = \sup\{\varrho(S) \,:\, X \in A_1(F)\} \text{ and } \underline{\varrho} = \inf\{\varrho(S) \,:\, X \in A_1(F)\}.$$

We refer to this setting as the *unconstrained setting.* This setting has been studied extensively in recent literature; we refer to Embrechts et al. (2014) for a fairly up-to-date account.

By definition, the admissible class $A(H)$ of risk vectors X with information on the risk factor Z is contained in $A_1(F)$, i.e., $A(H) \subset A_1(F)$. Hence, $\underline{\varrho} \le \underline{\varrho}^f$ and $\overline{\varrho}^f \le \overline{\varrho}$. Note that when the risk factor Z is independent of $X_1, \dots, X_n$, i.e., $F_{i|z} = F_i$, $1 \le i \le n$, only the information on the marginal distributions is useful and the study of the constrained bounds $\overline{\varrho}^f$ and $\underline{\varrho}^f$ reduces to the well-studied cases of the unconstrained bounds $\overline{\varrho}$ and $\underline{\varrho}$.

Following Embrechts et al. (2015), in the unconstrained setting we define the dependence uncertainty spread of a risk measure ϱ as the difference $\overline{\varrho} - \underline{\varrho}$ and, in the constrained setting, we define it as $\overline{\varrho}^f - \underline{\varrho}^f$. To measure the improvement we obtain through the factor information, we propose the measure of improvement Δ_ϱ, defined as

$$\Delta_\varrho = 1 - \frac{\overline{\varrho}^f - \underline{\varrho}^f}{\overline{\varrho} - \underline{\varrho}},$$

in which we assume by convention that $\Delta_\varrho = 1$ when $\overline{\varrho} = \underline{\varrho}$. Specifically, we study the problem (2.1) using as risk measure ϱ the tail probability. That is, we consider

$$M(t) := P(S \geq t) \quad \text{for} \quad t \in \mathbb{R},$$

but we also use risk measures such as VaR and TVaR. Here, the VaR at α-confidence level, $0 < \alpha < 1$, is defined by

$$\mathrm{VaR}_\alpha(Y) = \inf\{x \in \mathbb{R} \,:\, F_Y(x) \geq \alpha\}, \quad Y \in \mathcal{X},$$

where $F_Y(x)$ is the distribution function of Y. The VaR is thus defined as the (left) *generalized inverse* of the distribution function, i.e., $\mathrm{VaR}_\alpha(Y) = F_Y^{-1}(\alpha)$. The TVaR at confidence level $\alpha \in (0,1)$ is defined as

$$\mathrm{TVaR}_\alpha(Y) = \frac{1}{1-\alpha}\int_\alpha^1 \mathrm{VaR}_u(Y)\,\mathrm{d}u, \quad Y \in \mathcal{X}.$$

When the risk measure ϱ is the tail probability, the VaR or the TVaR we denote in the partially specified risk factor model the upper bounds $\overline{\varrho}^f$ in (2.1) by $\overline{M}^f(t)$ $(t \in \mathbb{R})$, $\overline{\mathrm{VaR}}^f_\alpha$ or $\overline{\mathrm{TVaR}}^f_\alpha$, $\alpha \in (0,1)$, respectively. The corresponding risk infima are denoted by $\underline{M}^f(t)$, $\underline{\mathrm{VaR}}^f_\alpha$ and $\underline{\mathrm{TVaR}}^f_\alpha$, respectively. In the model with marginal information only, $\overline{\varrho}$ is specifically denoted as $\overline{M}(t)$ $(t \in \mathbb{R})$, $\overline{\mathrm{VaR}}_\alpha$ and $\overline{\mathrm{TVaR}}_\alpha$, $\alpha \in (0,1)$, and similarly for other quantities.

It turns out to be useful to describe the risk vector $X = (X_i)_{1\leq i\leq n} \in A(H)$ through a **mixture representation**: $X \overset{d}{=} X_Z$ with $X_z = (X_{i,z})_{1\leq i\leq n} \in A_1(F_z)$, $z \in D$. Here, $F_z = (F_{i|z})_{1\leq i\leq n}$ is the vector of conditional distributions of X_i given $Z = z$, $i = 1, \dots, n$. By conditioning, the distribution F_S of S satisfies

$$F_S = \int F_{S_z}\,\mathrm{d}G(z), \tag{2.2}$$

where (and throughout) $S_z = \sum_{i=1}^n X_{i,z}$ is the sum of the conditional variables $(X_{i,z})_{1\leq i\leq n}$, and the above integral, without further specification, is taken over its natural region D. The random variables $X_{i,z}$, $i = 1, 2, \dots, n$, $z \in D$ can be constructed as

$$X_{i,z} = F^{-1}_{i|z}(U_{i,z}), \quad 1 \leq i \leq n, \tag{2.3}$$

where $U_z = (U_{1,z}, \dots, U_{n,z})$ is some random vector with $U(0,1)$ marginal distributions, and $(U_z)_{z\in D}$ is independent of Z. Of course, in a similar way as for $A_1(F)$ and $A(H)$, risk bounds can also be defined for the admissible class $A_1(F_z)$. In this paper, the notation $\overline{M}_z(t)$ is used to denote the sharp tail probability bound for the class $A_1(F_z)$.

2.2 VaR bounds

The mixture representation in (2.2) implies the following sharp tail probability bounds. The proof is provided in the appendix. In this section we do not impose any assumptions on G and F_z, $z \in D$.

Proposition 2.1 (Sharp tail probability bounds) *The sharp upper and lower tail probability bounds for the partially specified risk factor model are given by*

$$\overline{M}^f(t) = \int \overline{M_z}(t)\,\mathrm{d}G(z), \textit{ and } \underline{M}^f(t) = \int \underline{M}_z(t)\,\mathrm{d}G(z), \quad t \in \mathbb{R}. \tag{2.4}$$

As a corollary to Proposition 2.1, we obtain the following sharp VaR bounds.

Corollary 2.2 (Sharp VaR bounds) *The sharp upper and lower* VaR *bounds in the partially specified risk factor model are given by*

$$\overline{\mathrm{VaR}}^f_\alpha = (\overline{M}^f)^{-1}(1-\alpha), \textit{ and } \underline{\mathrm{VaR}}^f_\alpha = (\underline{M}^f)^{-1}(1-\alpha), \ \alpha \in (0,1), \tag{2.5}$$

where, for $\alpha \in (0,1)$*, respectively,*

$$(\overline{M}^f)^{-1}(1-\alpha) := \sup\left\{t \in \mathbb{R} \,:\, \overline{M}^f(t) > 1-\alpha\right\},$$

$$(\underline{M}^f)^{-1}(1-\alpha) := \sup\left\{t \in \mathbb{R} \,:\, \underline{M}^f(t) > 1-\alpha\right\}.$$

The representation result in (2.4) shows that when the risk measure at hand is the tail probability or the VaR, the problem of determining sharp bounds in the constrained setting essentially reduces to the aggregation of bounds that are derived using information on conditional distributions $(F_{i|z})$ only. Hence, we can build on the results that have been derived in this unconstrained setting; see Embrechts et al. (2014) and the references therein for a summary of existing results. We apply some of these results to the following two-dimensional example with normally distributed risks and compare the dependence uncertainty spread in the unconstrained setting with the one in the constrained setting.

*Example 2.3 (*VaR *bounds for normally distributed risks)* Assume that X_1 and X_2 have $N(0,1)$ distributed marginals with distribution denoted by Φ and that Z is a risk factor such that (X_i, Z) has a bivariate normal distribution with correlation parameter $r_i \in (-1,1)$, $i = 1,2$. A stochastic representation is given by $X_i = r_i Z + \sqrt{1-r_i^2}\varepsilon_i, \quad i = 1,2$, where ε_1 and ε_2 have $N(0,1)$ distributed marginals and are independent of Z (but have an unknown joint dependence).

As for the *unconstrained bounds* (with information on marginal distributions only), we obtain from results known in the literature (see Rüschendorf (1982) and Bernard et al. (2015))

$$\overline{\mathrm{VaR}}_\alpha = \mathrm{VaR}_0(\Phi^{-1}(U)+\Phi^{-1}(1+\alpha-U)) = 2\Phi^{-1}\left(\frac{1+\alpha}{2}\right),\ \alpha \in (0,1), \quad (2.6)$$

and

$$\underline{\mathrm{VaR}}_\alpha = \mathrm{VaR}_\alpha(\Phi^{-1}(V)+\Phi^{-1}(\alpha-V)) = 2\Phi^{-1}\left(\frac{\alpha}{2}\right),\ \alpha \in (0,1), \quad (2.7)$$

where $U \sim U[\alpha, 1]$ and $V \sim U(0, \alpha)$.

As for the *constrained bounds*, we first consider $r_1 = r_2$. Observe that $X_{i|z}$ has $N(r_i z, 1-r_i^2)$ distributed marginals, $i = 1, 2$. Hence, from (2.6)–(2.7) we obtain sharp upper bounds and lower bounds on $\overline{M}_z(t)$ and $\underline{M}_z(t)$, $t \in \mathbb{R}$. Using formula (2.4), we find, after a numerical inversion, the values of $\overline{\mathrm{VaR}}_\alpha^f$ and $\underline{\mathrm{VaR}}_\alpha^f$. Here, the values inside the integrals in (2.4) are known explicitly and the integral is evaluated numerically. Next, we consider $r_1 = -r_2$. We obtain that $\overline{\mathrm{VaR}}^f = \sqrt{1-r_1^2}\,\overline{\mathrm{VaR}}_\alpha$ and $\underline{\mathrm{VaR}}^f = \sqrt{1-r_1^2}\,\underline{\mathrm{VaR}}_\alpha$.

Table 2.1 displays the bounds for different values of r_i and α as well as the measure of improvement Δ_{VaR} obtained by using factor information. In Panel A, where $r_1 = r_2$, we observe that the upper bounds do not improve, whereas the lower bounds show essential improvements. In Panel B, where $r_1 = -r_2$, we find the opposite picture: the upper bounds improve significantly whereas the lower bounds remain essentially the same.

2.3 A mixture representation of VaR bounds

In general, there are two main challenges when evaluating the bounds $\overline{\mathrm{VaR}}_\alpha^f$ and $\underline{\mathrm{VaR}}_\alpha^f$. First, the representation result (2.5) requires, for a given probability level $\alpha \in (0,1)$, to establish the function $t \to \overline{M}^f(t)$, $t \in \mathbb{R}$, in which each $\overline{M}^f(t)$ requires aggregation of the (marginal) tail probability bounds $\overline{M}_z(t)$, $z \in D$; see also Example 2.3. Second, even in the unconstrained setting, obtaining sharp VaR bounds is an open problem in general, and analytical results are available only for small portfolios ($n = 2$), some classes of homogeneous portfolios and asymptotically large portfolios ($n \to \infty$) (see Embrechts et al. (2014)). Puccetti and Rüschendorf (2012) and Embrechts et al. (2013) propose the Rearrangement Algorithm (RA) as a practical approach to approximating the unconstrained bounds. In response to these issues, we proceed, by expressing in the remainder of this section, the VaR bounds $\overline{\mathrm{VaR}}_\alpha^f$ and $\underline{\mathrm{VaR}}_\alpha^f$ directly in terms of (marginal) VaR bounds. These expressions, combined with the use of some results and ideas that are valid in the unconstrained setting, provide the basis for obtaining, in Section 4, bounds that can be practically evaluated and that are asymptotically sharp.

To evaluate

$$\overline{\mathrm{VaR}}_\alpha^f = \sup\{\mathrm{VaR}_\alpha(S_Z)\ :\ X_Z \in A(H)\},$$

Panel A	r_1	VaR_α	$(\underline{\mathrm{VaR}}_\alpha, \overline{\mathrm{VaR}}_\alpha)$	$(\underline{\mathrm{VaR}}^f_\alpha, \overline{\mathrm{VaR}}^f_\alpha)$	Δ_{VaR}
	0	2.326	$(-0.125, 3.920)$	$(-0.125, 3.920)$	0%
$\alpha = 0.95$	0.5	2.849	$(-0.125, 3.920)$	$(0.822, 3.920)$	23.44%
	0.8	3.121	$(-0.125, 3.920)$	$(1.894, 3.880)$	50.92%
	1	3.290	$(-0.125, 3.920)$	$(3.290, 3.290)$	100%
	0	3.643	$(-0.0125, 5.614)$	$(-0.013, 5.614)$	0%
$\alpha = 0.995$	0.5	4.461	$(-0.0125, 5.614)$	$(1.893, 5.614)$	33.87%
	0.8	4.887	$(-0.0125, 5.614)$	$(3.464, 5.606)$	61.93%
	1	5.152	$(-0.0125, 5.614)$	$(5.152, 5.152)$	100%
Panel B	r_1	VaR_α	$(\underline{\mathrm{VaR}}_\alpha, \overline{\mathrm{VaR}}_\alpha)$	$(\underline{\mathrm{VaR}}^f_\alpha, \overline{\mathrm{VaR}}^f_\alpha)$	Δ_{VaR}
	0	2.326	$(-0.125, 3.920)$	$(-0.125, 3.920)$	0%
$\alpha = 0.95$	0.5	2.849	$(-0.125, 3.920)$	$(-0.109, 3.395)$	13.4%
	0.8	3.121	$(-0.125, 3.920)$	$(-0.075, 2.352)$	40%
	1	3.290	$(-0.125, 3.920)$	$(0.000, 0.000)$	100%
	0	3.643	$(-0.0125, 5.614)$	$(-0.0125, 5.614)$	0%
$\alpha = 0.995$	0.5	4.461	$(-0.0125, 5.614)$	$(-0.011, 4.862)$	13.4%
	0.8	4.887	$(-0.0125, 5.614)$	$(-0.007, 3.368)$	40%
	1	5.152	$(-0.0125, 5.614)$	$(0.000, 0.000)$	100%

Table 2.1 VaR bounds in the normal case. Panel A: $r_1 = r_2$. Panel B: $r_1 = -r_2$. VaR_α corresponds to the case in which (X_1, X_2) is bivariate normally distributed with correlation r_1^2.

we first provide two explicit representations of $\mathrm{VaR}_\alpha(S_Z)$. Hence, for $\gamma \in \mathbb{R}$ and $z \in D$, define

$$p_z^\gamma = F_{S_z}(\gamma) \text{ and } b^\gamma = \operatorname{esssup}_{z\in D} \mathrm{VaR}_{p_z^\gamma}(S_z), \tag{2.8}$$

where $\operatorname{esssup}_{z\in D}$ in (2.8) is taken with respect to G (the distribution of Z). Note that for distributions with positive densities on its support, b^γ is equal to γ.

Proposition 2.4 (VaR **representation of mixtures)** *For $\alpha \in (0,1)$, the* VaR *at level α of the mixture S_Z has the following representations:*

$$\mathrm{VaR}_\alpha(S_Z) = \inf\left\{\gamma \in \mathbb{R} : \int p_z^\gamma \,\mathrm{d}G(z) \geq \alpha\right\}, \tag{2.9}$$

$$\mathrm{VaR}_\alpha(S_Z) = \inf\left\{b^\gamma : \gamma \in \mathbb{R}, \int p_z^\gamma \,\mathrm{d}G(z) \geq \alpha\right\}. \tag{2.10}$$

Proof We first prove (2.9). For $\alpha \in (0,1)$, we have, by definition,

$$\begin{aligned}\mathrm{VaR}_\alpha(S_Z) &= \inf\{\gamma \in \mathbb{R} : F_{S_Z}(\gamma) \geq \alpha\} = \inf\left\{\gamma \in \mathbb{R} : \int F_{S_z}(\gamma)\,\mathrm{d}G(z) \geq \alpha\right\} \\ &= \inf\left\{\gamma \in \mathbb{R} : \int p_z^\gamma \,\mathrm{d}G(z) \geq \alpha\right\}. \end{aligned}\tag{2.11}$$

To prove (2.10), observe that for all z, $\mathrm{VaR}_{p_z^\gamma}(S_z) = F_{S_z}^{-1}(p_z^\gamma) \le \gamma$. Hence,

$$b^\gamma = \mathrm{esssup}_{z\in D}\, \mathrm{VaR}_{p_z^\gamma}(S_z) \le \gamma,$$

and, therefore, $\mathrm{VaR}_\alpha(S_Z) \ge \inf\left\{ b^\gamma \,:\, \gamma \in \mathbb{R},\ \int p_z^\gamma \,\mathrm{d}G(z) \ge \alpha \right\}$.

Conversely, for any $\gamma \in \mathbb{R}$ with $\int p_z^\gamma \,\mathrm{d}G(z) \ge \alpha$ it holds that $F_{S_z}(b^\gamma) \ge F_{S_z} \circ F_{S_z}^{-1}(p_z^\gamma) \ge p_z^\gamma$. This implies that $\int F_{S_z}(b^\gamma)\,\mathrm{d}G(z) \ge \int p_z^\gamma \,\mathrm{d}G(z) \ge \alpha$, i.e., b^γ is also an admissible constant in (2.11) and, therefore, $\mathrm{VaR}_\alpha(S_Z) \le \inf\{b^\gamma \,:\, \gamma \in \mathbb{R},\ \int p_z^\gamma \,\mathrm{d}G(z) \ge \alpha\}$ and we obtain equality.

The second representation for $\mathrm{VaR}_\alpha(S_Z)$ in Proposition 2.4 is of some independent interest and may provide the intuition to develop a convenient expression for $\overline{\mathrm{VaR}}_\alpha^f$ in terms of (marginal) VaR bounds. In the formal derivation of this expression, however, we solely build on the first (basic) representation for $\mathrm{VaR}_\alpha(S_Z)$. Using the shorthand notation $q_z(\beta)$ for $\mathrm{VaR}_\beta(S_z)$ $z \in D,\ \beta \in (0,1)$ with right-continuous generalized inverse

$$q_z^{-1}(\gamma) = \sup\{x \in [0,1] \,:\, q_z(x) \le \gamma\}$$

it holds that

$$p_z^\gamma = q_z^{-1}(\gamma).$$

Hence, we obtain that

$$\mathrm{VaR}_\alpha(S_Z) = \inf\left\{ \gamma \in \mathbb{R} \,:\, \int q_z^{-1}(\gamma)\,\mathrm{d}G(z) \ge \alpha \right\} =: b^*(\alpha). \tag{2.12}$$

In order to obtain a corresponding representation of the sharp VaR bound $\overline{\mathrm{VaR}}_\alpha^f$, we define the worst (conditional) VaRs for the conditional sum S_z by

$$\overline{q}_z(\beta) := \overline{\mathrm{VaR}}_\beta(S_z) = \sup\{\mathrm{VaR}_\beta(S_z) \,:\, X_z \in A_1(F_z)\}, \quad z \in D,\ \beta \in (0,1),$$

with the right-continuous generalized inverse denoted by $(\overline{q}_z)^{-1}(\gamma)$ and, finally,

$$\overline{b}^*(\alpha) := \inf\left\{ \gamma \in \mathbb{R} \,:\, \int (\overline{q}_z)^{-1}(\gamma) dG(z) \ge \alpha \right\}. \tag{2.13}$$

The next proposition shows that $\overline{b}^*(\alpha)$ is the sharp upper bound of the VaR in the presence of information on a risk factor.

Proposition 2.5 (Mixture representation for the sharp VaR **upper bound)** *For $\alpha \in (0,1)$, the sharp upper bound of* VaR *in the partially specified risk factor model is given by*

$$\overline{\mathrm{VaR}}_\alpha^f = \overline{b}^*(\alpha) = \mathrm{VaR}_\alpha(\overline{q}_Z(V)),$$

where V is a $U(0,1)$ distributed random variable independent of Z.

Proof We have by definition $\overline{\mathrm{VaR}}^f_\alpha = \sup\{\mathrm{VaR}_\alpha(S_Z),\ X_Z \in A(H)\}$. Since $q_z^{-1}(\gamma) \geq \overline{q}_z^{-1}(\gamma)$ for $z \in D$, (2.12) implies that

$$\overline{\mathrm{VaR}}^f_\alpha = \sup_{X_Z \in A(H)} \inf\left\{\gamma \in \mathbb{R} : \int q_z^{-1}(\gamma) dG(z) \geq \alpha\right\} \tag{2.14}$$

$$\leq \inf\left\{\gamma \in \mathbb{R};\ \int \overline{q}_z^{-1}(\gamma) dG(z) \geq \alpha\right\} \tag{2.15}$$

$$= \overline{b}^*(\alpha).$$

On the other hand, for $z \in D$, let $\overline{X}_z \sim F_z$ be a solution to $\overline{q}_z^{-1}(\alpha) = \mathrm{VaR}_\alpha(\overline{S}_z)$, where $\overline{S}_z = \sum_{i=1}^n \overline{X}_{i,z}$. Then $\mathrm{VaR}_\alpha(\overline{S}_Z) = \overline{b}^*(\alpha)$, and thus equality in (2.14) holds. The equality $\overline{b}^*(\alpha) = \mathrm{VaR}_\alpha(\overline{q}_Z(V))$ follows from

$$\mathrm{VaR}_\alpha(\overline{q}_Z(V)) = \inf\{\gamma \in \mathbb{R} : P(\overline{q}_Z(V) \leq \gamma) \geq \alpha\}$$
$$= \inf\left\{\gamma \in \mathbb{R} : \int \overline{q}_z^{-1}(\gamma) dG(z) \geq \alpha\right\} = \overline{b}^*(\alpha).$$

While formula (2.13) for the VaR bound $\overline{b}^*(\cdot)$ is explicit, in general it is still not straightforward to evaluate it. Indeed, we need to obtain the (conditional) VaR bounds $\overline{q}_z(v)$ for $z \in D$, $v \in (0,1)$. Few explicit results exist, and a practical evaluation of $\overline{b}^*(\cdot)$ thus appears to require a repeated use of the RA (for approximating all $\overline{q}_z(v)$). In Section 4, however, we show that the sharp upper bound $\overline{b}^*(\alpha)$ can be approximated (from above) by easy-to-compute upper bounds that are defined in terms of the TVaR. Furthermore, these approximations are asymptotically sharp.

Remark 2.6 A mixture representation of $\underline{\mathrm{VaR}}^f_\alpha$ can be obtained in a similar way by replacing the upper bound quantities $\overline{q}_z(\alpha) = \overline{\mathrm{VaR}}_\alpha(S_z)$ with the corresponding lower bound quantities $\underline{q}_z(\alpha) = \underline{\mathrm{VaR}}_\alpha(S_z)$, in which $\underline{\mathrm{VaR}}_\alpha(S_z) = \inf\{\mathrm{VaR}_\beta(S_z) : X_z \in A_1(F_z)\},\quad z \in D,\ \beta \in (0,1)$. □

3 Bounds for convex risk measures

We first recall the definition of convex order.

Definition 3.1 (Convex order) Let X and Y be two random variables with finite means. X is smaller than Y *in convex order*, denoted by $X \leq_{\mathrm{cx}} Y$, if for all convex functions f,

$$E[f(X)] \leq E[f(Y)], \tag{3.1}$$

whenever both sides of (3.1) are well defined.

It is well-known that a law-invariant convex risk measure ϱ (e.g., TVaR) is consistent w.r.t. convex order on proper probability spaces such as L^1 (integrable rvs) and L^∞ (bounded rvs); see Chapter 4 of Föllmer and Schied (2004), Jouini et al. (2006), Bäuerle and Müller (2006) and Burgert and Rüschendorf

(2006). From this section on, we let $\mathcal{X} = L^1$, and all marginal distributions of F and F_z, $z \in D$ are assumed to have finite first moment.

In this instance, the study of $\overline{\varrho}^f$ (resp. $\underline{\varrho}^f$) is closely connected to finding $X \in A(H)$ such that $S = X_1 + \cdots + X_n$ becomes the largest (resp. smallest) element w.r.t. convex order. We define the admissible class of sums in the partially specified risk factor model

$$\mathcal{S}(H) = \{X_1 + \cdots + X_n \,:\, X \in A(H)\}$$

and note that the upper and lower bounds $\overline{\varrho}^f$ and $\underline{\varrho}^f$ can be equivalently defined in terms of $\mathcal{S}(H)$ rather than $A(H)$, i.e.,

$$\overline{\varrho}^f = \sup\{\varrho(S) \,:\, S \in \mathcal{S}(H)\} = \sup\{\varrho(S) \,:\, X \in A(H)\}$$

and

$$\underline{\varrho}^f = \inf\{\varrho(S) \,:\, S \in \mathcal{S}(H)\} = \inf\{\varrho(S) \,:\, X \in A(H)\}.$$

3.1 Upper bound

We first focus on $\overline{\varrho}^f$ and thus aim at finding an element in $\mathcal{S}(H)$ that is largest w.r.t. convex order. To this end, we recall that a classical result of Meilijson and Nadas (1979) established that the comonotonic sum $S^c = \sum_{i=1}^n F_i^{-1}(U)$, $U \sim U(0,1)$ is larger in the sense of convex order than any other sum $X_1 + \cdots + X_n$, $X \in A_1(F)$. This result suggests that on $\mathcal{S}(H)$ the conditionally comonotone sum

$$S_Z^c = \sum_{i=1}^n F_{i|Z}^{-1}(U). \tag{3.2}$$

is a largest element w.r.t. convex order and thus leads to sharp upper bounds for TVaR and for other law-invariant convex risk measures. In the following, $U \sim U(0,1)$ and is independent of Z. The notation for (conditionally) comonotonic sums, $S_z^c = \sum_{i=1}^n F_{i|z}^{-1}(U)$, $S_Z^c = \sum_{i=1}^n F_{i|Z}^{-1}(U)$ and $S^c = \sum_{i=1}^n F_i^{-1}(U)$, will be used repeatedly.

Proposition 3.2 (Sharp upper bounds for convex risk measures) *The following statements hold.*

a) For all $S \in \mathcal{S}(H)$ it holds that $S \leq_{cx} S_Z^c \in \mathcal{S}(H)$.
b) For a law-invariant convex risk measure ϱ, we have $\overline{\varrho}^f = \varrho(S_Z^c)$.
c) $S_Z^c \leq_{cx} S^c$.

The proof of this proposition is given in the appendix.

Remark 3.3 The statement $S \leq_{cx} S_Z^c$ can (essentially) also be found in Kaas et al. (2000), who showed that for any $X_i \sim F_i$ and for any random variable Z that is a function of the X_i, $i = 1, 2, \ldots n$,

$$S = \sum_{i=1}^n X_i \leq_{cx} \sum_i F_{i|Z}^{-1}(U). \tag{3.3}$$

Formula (3.3) has been applied in several examples to obtain improved upper risk bounds for basket options and Asian options - mainly, however, in a log-normal context with conditional distributions that are easy to evaluate (see e.g., Vanmaele et al. (2006), Deelstra et al. (2008), Vanduffel et al. (2008)). □

3.2 Lower bound

As for the study of $\underline{\varrho}^f$, we notice that obtaining a lower bound w.r.t. convex order in $\mathcal{S}(H)$ is a more difficult task than obtaining an upper bound. In fact, even lower bounds w.r.t. convex order for sums $S = \sum_{i=1}^n X_i,\ X \in A_1(F)$ are generally not available; some analytical cases, however, can be found in Wang and Wang (2011) and in Bernard et al. (2014). In this regard, Wang et al. (2013) introduce the notion of *joint mixability* and show its relevance for obtaining such lower bounds. We will see that joint mixability is also relevant to finding lower bounds w.r.t. convex order in $\mathcal{S}(H)$.

Definition 3.4 (Joint mixability) Suppose n is a positive integer. An n-tuple $(F_1, \dots, F_n)$ of probability distributions on $\mathbb{R}$ is *jointly mixable* (JM) if there exist n random variables $X_1 \sim F_1, \dots, X_n \sim F_n$ such that $X_1 + \cdots + X_n$ is a constant.

For any $X = X_Z \in A(H)$ and $z \in D$, let $\mu_z = E[S_z]$, which is the sum of the means of $F_{i|z}$, $i = 1, \dots, n$ and hence it is independent of the choice of $X_Z \in A(H)$. It is easy to observe that $E(S_Z|Z) = \mu_Z$. We show that μ_Z serves as a natural candidate for the smallest element in $\mathcal{S}(H)$ w.r.t. convex order. Recall that $F_z = (F_{i|z})_{1\le i\le n}$, $z \in D$ and consider U as a $U(0,1)$ distributed random variable that is independent of Z.

Proposition 3.5 (Lower bounds for convex risk measures) *The following statements hold.*

a) For all $S \in \mathcal{S}(H)$ it holds that $\mu_Z \le_{cx} S$.
b) For a law-invariant convex risk measure ϱ, we have $\varrho(\mu_Z) \le \underline{\varrho}^f$.
c) $\mu_Z \in \mathcal{S}(H)$ if and only if F_z is jointly mixable for G-almost surely $z \in D$.
d) For $n = 2$ and a law-invariant convex risk measure ϱ, we have $\underline{\varrho}^f = \varrho(S_Z^a)$ where $S_z^a = F_{1|z}^{-1}(U) + F_{2|z}^{-1}(1-U)$, $z \in D$.

Proof
a) For any $S \in \mathcal{S}(H)$, write $S = S_Z$, and conditional Jensen's inequality implies that $E(S_Z|Z) = \mu_Z \le_{cx} S_Z$.
b) From a), $\varrho(\mu_Z) \le \varrho(S)$ for all $S \in \mathcal{S}(H)$, which implies the result.
c) Suppose that F_z is jointly mixable for $z \in D_0$. By the definition of joint mixability and (2.3), there exist $U_z = (U_{1,z}, \dots, U_{n,z})$, $z \in D_0$ such that $X_{i,z} = F_{i|z}^{-1}(U_{i,z})$, $1 \le i \le n$, and $S_z = \sum_{i=1}^n X_{i,z} = \mu_z$, for $z \in D_0$. This shows $\mu_Z = S_Z$, almost surely, and hence $\mu_Z \in \mathcal{S}(H)$.

For the other direction, take $X_Z \in A(H)$ such that $S_Z = \mu_Z$ almost surely. Then, since $\sum_{i=1}^n X_{i,Z} = \mu_Z$ almost surely, there exists D_0, $P(Z \in D_0) = 1$ such that $\sum_{i=1}^n X_{i,z} = S_z = \mu_z$ for each $z \in D_0$. That is, F^z is jointly mixable for each $z \in D_0$.

d) Note that for any $U(0,1)$ random variables U_1 and U_2, we have $F_{1|z}^{-1}(U) + F_{2|z}^{-1}(1-U) \leq_{cx} F_{1|z}^{-1}(U_1) + F_{2|z}^{-1}(U_2)$ since counter-monotonicity yields a sum that is minimum w.r.t. convex order. Thus, for any $S_Z \in \mathcal{S}(H)$, we have $S_z^a \leq_{cx} S_z$. By definition of convex order, for any convex function f, such that $E[f(S_Z)]$ and $E[f(S_Z^a)]$ are well defined, we have

$$E[f(S_Z^a)] = \int_D E[f(S_z^a)]dG(z) \leq \int_D E[f(S_z)]dG(z) = E[f(S_Z)]$$

and hence $S_Z^a \leq_{cx} S_Z$, implying $\varrho(S_Z^a) \leq \varrho(S_Z)$. □

As a consequence of Propositions 3.2 and 3.5, we obtain that, for any law-invariant convex risk measure ϱ and any $S \in \mathcal{S}(H)$,

$$\varrho(\mu) \leq \varrho(\mu_Z) \leq \varrho(S) \leq \varrho(S_Z^c) \leq \varrho(S^c),$$

where $\mu = E[\mu_Z]$. In particular, Proposition 3.5 suggests that w.r.t. convex order the best-case risk $S \in \mathcal{S}(H)$ is the one whose randomness derives entirely from the factor Z. However, to prove that $\mu_Z \in \mathcal{S}(H)$ one needs to establish joint mixability, which is difficult and not valid in general. Some analytical results for joint mixability are given in Wang and Wang (2016). An example of the sharp lower bound $\mu_Z \in \mathcal{S}(H)$ in a location-scale family is provided next.

Example 3.6 (Convex order bounds in location-scale families) Let $Z = (Z_1, Z_2)$ have an arbitrary distribution on $D = \mathbb{R}^2$. For some real numbers $a_i, b_i \in \mathbb{R}$, $i = 1, \ldots, n$, and positive numbers $\sigma_1, \ldots, \sigma_n$ satisfying $2\max_{i=1,\ldots,n} \sigma_i \leq \sum_{i=1}^n \sigma_i$, let

$$X_i = a_i + Z_1 b_i + Z_2(\sigma_i \varepsilon_i), \quad i = 1, \ldots, n,$$

where the $\varepsilon_1, \ldots, \varepsilon_n$ are identically distributed and are independent of Z but the joint distribution of $(\varepsilon_1, \ldots, \varepsilon_n)$ is not known. That is, Z_1 is a common location factor and Z_2 is a common scale factor for the n risks $X_1, \ldots, X_n$. Assume that the distribution F_0 of the ε_i has a unimodal and symmetric density; this includes the normal and t-distribution. Let ε be an F_0-distributed random variable independent of Z and write

$$a = \sum_{i=1}^n a_i, \quad b = \sum_{i=1}^n b_i, \quad \sigma = \sum_{i=1}^n \sigma_i.$$

From Proposition 3.2, a largest element in $\mathcal{S}(H)$ w.r.t. convex order is given by

$$S_Z^c = a + bZ_1 + \sigma\varepsilon Z_2.$$

From Corollary 3.6 of Wang and Wang (2016), we know that, for $z \in \mathbb{R}^2$, the tuple of distributions of $X_{1,z}, \ldots, X_{n,z}$ is jointly mixable. This allows us to apply Proposition 3.5 to obtain a smallest element in $\mathcal{S}(H)$ w.r.t. convex order, which is given by

$$\mu_Z = a + bZ_1,$$

where we used the fact that $E[\varepsilon] = 0$. Thus, we have sharp bounds on a convex risk measure ϱ, i.e., for $S \in \mathcal{S}(H)$,

$$a + \varrho bZ_1 \leq \varrho(S) \leq a + \varrho\left(bZ_1 + \sigma\varepsilon Z_2\right), \quad \alpha \in (0,1),$$

In the following example we illustrate the TVaR bounds for normally distributed risks.

*Example 3.7 (*TVaR *bounds for normally distributed risks)* The set up is as in Example 2.3. Recall that, for a standard normally distributed risk X,

$$\mathrm{TVaR}_\alpha(X) = \frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}, \quad \alpha \in (0,1),$$

where ϕ is the standard normal density. As for the *unconstrained bounds,* S^c has a $N(0,4)$ distribution and $E(S) = 0$. We obtain that

$$\overline{\mathrm{TVaR}}_\alpha = 2\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} \text{ and } \underline{\mathrm{TVaR}}_\alpha = 0, \quad \alpha \in (0,1).$$

As for the *constrained bounds,* S_Z^c has a $N(0,\sigma_1^2)$ distribution with $\sigma_1^2 = 2(1 + r_1r_2 + \sqrt{(1-r_1^2)(1-r_2^2)})$ and S_Z^a has a $N(0,\sigma_2^2)$ distribution with $\sigma_2^2 = 2(1 + r_1r_2 - \sqrt{(1-r_1^2)(1-r_2^2)})$. Hence,

$$\overline{\mathrm{TVaR}}_\alpha^f = \sigma_1\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha} \text{ and } \underline{\mathrm{TVaR}}_\alpha = \sigma_2\frac{\phi(\Phi^{-1}(\alpha))}{1-\alpha}, \quad \alpha \in (0,1).$$

Table 3.1, Panel A, shows these TVaR bounds for different values of $r_1 = r_2$ and α. We observe that there is no difference between the unconstrained upper bound and the constrained one, whereas there is an improvement of the lower bound. In Table 3.1, Panel B we show the TVaR bounds for different values of $r_1 = -r_2$ and α (note that in this case, $S_z^a = \mu_Z = 0$). In this case, we find the opposite picture; the upper bounds may improve significantly, whereas the lower bounds remain unchanged.

To assess the impact of heterogeneity, we fix $r_1 = -0.5$ where r_2 varies between -1 and 1. We represent the bounds and the improvements Δ_{TVaR} in Figure 3.1. When $|r_1| \neq |r_2|$, both the upper and the lower bounds improve.

□

From the following example we obtain further insight into the influence of dependence information of a factor model on risk bounds.

Panel A	r_1	TVaR_α	$(\underline{\text{TVaR}}_\alpha, \overline{\text{TVaR}}_\alpha)$	$(\underline{\text{TVaR}}^f_\alpha, \overline{\text{TVaR}}^f_\alpha)$	Δ_{TVaR}
	0	2.917	$(0.000, 4.125)$	$(0.000, 4.125)$	0%
$\alpha = 0.95$	0.5	3.573	$(0.000, 4.125)$	$(2.063, 4.125)$	50%
	0.8	3.914	$(0.000, 4.125)$	$(3.300, 4.125)$	80%
	1	4.125	$(0.000, 4.125)$	$(4.125, 4.125)$	100%
	0	4.090	$(0.000, 5.784)$	$(0.000, 5.784)$	0%
$\alpha = 0.995$	0.5	5.009	$(0.000, 5.784)$	$(2.892, 5.784)$	50%
	0.8	5.487	$(0.000, 5.784)$	$(4.627, 5.784)$	80%
	1	5.784	$(0.000, 5.784)$	$(5.784, 5.784)$	100%
Panel B	r_1	TVaR_α	$(\underline{\text{TVaR}}_\alpha, \overline{\text{TVaR}}_\alpha)$	$(\underline{\text{TVaR}}^f_\alpha, \overline{\text{TVaR}}^f_\alpha)$	Δ_{TVaR}
	0	2.917	$(0.000, 4.125)$	$(0.000, 4.125)$	0%
$\alpha = 0.95$	0.5	3.573	$(0.000, 4.125)$	$(0.000, 3.573)$	13.4%
	0.8	3.914	$(0.000, 4.125)$	$(0.000, 2.475)$	40%
	1	4.125	$(0.000, 4.125)$	$(0.000, 0.000)$	100%
	0	4.090	$(0.000, 5.784)$	$(0.000, 5.784)$	0%
$\alpha = 0.995$	0.5	5.009	$(0.000, 5.784)$	$(0.000, 5.009)$	13.4%
	0.8	5.487	$(0.000, 5.784)$	$(0.000, 3.470)$	40%
	1	5.784	$(0.000, 5.784)$	$(0.000, 0.000)$	100%

Table 3.1 TVaR bounds for the normal case. Panel A: $r_1 = r_2$. Panel B: $r_1 = -r_2$. The column TVaR_α provides the TVaR in the case in which (X_1, X_2) is bivariate normally distributed with correlation r_1^2.

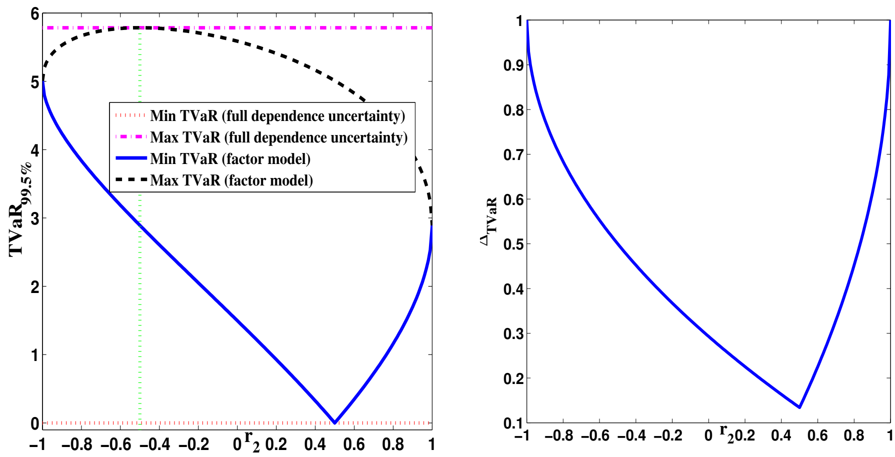

Fig. 3.1 Bounds on TVaR $_{99.5\%}$ are displayed (left panel), along with the degree of improvement (right panel). We consider $r_1 = -0.5$, and r_2 varies between -1 and 1.

Example 3.8 (Pareto risks) We consider a risk factor model for the case $n = 2$ given by

$$X_1 = (1 - Z)^{-1/3} - 1 + \varepsilon_1,$$
$$X_2 = p((1 - Z)^{-1/3} - 1) + (1 - p)(Z^{-1/3} - 1) + \varepsilon_2,$$

where $Z \sim U(0,1)$, ε_1 and ε_2 are Pareto(4) distributed and independent of Z, and $p \in (0,1)$. We allow any dependence between the variables ε_1 and ε_2. In this example the common component $(1-Z)^{-1/3}-1$ is Pareto(3)-distributed and thus dominates the idiosyncratic risk components ε_i. Based on the risk bounds established in Section 3 (Propositions 3.2 and 3.5), we obtain, for TVaR at level $\alpha = 0.95$, the dependence uncertainty spread as in Figure 3.2. For $p \approx 0$ the common risk factor Z creates strong negative dependence between X_1 and X_2 and, as a consequence, we obtain a strong reduction in the upper risk bounds. For $p \approx 1$ the risk factor Z induces strong positive dependence between X_1 and X_2 and we obtain, as a consequence, a strong improvement in the lower bounds (but not in the upper bounds). For all intermediate p we have a total reduction of a similar order.

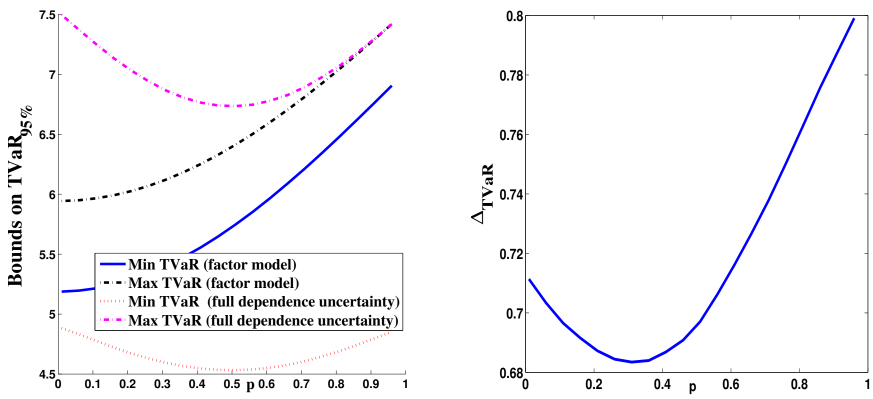

Fig. 3.2 Bounds for TVaR at 95 % are displayed (left panel), along with the degree of improvement (right panel). derive the TVaR bounds are given in Theorem 3.2.

□

4 The relation between VaR and TVaR bounds

4.1 Dependence uncertainty spread

In Example 3.7, we observed that adding factor information does not always yield improved bounds. In particular, the upper bound on TVaR does not improve when $F_{1|z} = F_{2|z}$, $z \in D$ (i.e., when $r_1 = r_2$), and the lower bound remains unchanged when $\mu_z = \mu$ for $z \in D$. The following proposition generalizes these observations and provides conditions under which upper and lower bounds on a law-invariant convex risk measure do not improve when using factor information.

Proposition 4.1 (No improvement for convex risk measures) *Let ϱ be a law-invariant convex risk measure.*

a) If $F_{1|z} = \cdots = F_{n|z}$ for all $z \in D$, then $\overline{\varrho}^f = \overline{\varrho} = \varrho(S^c)$.
b) If $\mu_z = \mu$ and F_z is jointly mixable for all $z \in D$, then $\underline{\varrho}^f = \underline{\varrho} = \varrho(\mu)$.
c) If the conditions of both a) and b) hold, then $\Delta_\varrho = 0$.

Proof a) Note that $\varrho(S_Z^c) = \overline{\varrho}^f \leq \overline{\varrho} = \varrho(S^c)$. It suffices to show $\varrho(S_Z^c) = \varrho(S^c)$. Note that $F_1 = \cdots = F_n$, and we have $S_z^c \stackrel{\mathrm{d}}{=} nF_{1|z}^{-1}(U)$ and $S^c \stackrel{\mathrm{d}}{=} nF_1^{-1}(U)$ for some $U \sim U(0,1)$, $z \in D$. We can check for $x \in \mathbb{R}$, that

$$
\begin{aligned}
P(S_Z^c \leq x) &= \int P(nF_{1|z}^{-1}(U) \leq x)\mathrm{d}G(z) \\
&= \int F_{1|z}\left(\frac{x}{n}\right)\mathrm{d}G(z) \\
&= F_1\left(\frac{x}{n}\right) \\
&= P(nF_1^{-1}(U) \leq x) = P(S^c \leq x).
\end{aligned}
$$

Thus, $S_Z^c \stackrel{\mathrm{d}}{=} S^c$ and $\overline{\varrho}^f = \varrho(S_Z^c) = \varrho(S^c) = \overline{\varrho}$.

b) Note that $\overline{\varrho}^f \geq \overline{\varrho} \geq \varrho(\mu)$. From Proposition 3.5, we have $\mu_Z \in \mathcal{S}(H)$, and thus $\mu \in \mathcal{S}(H)$ since $\mu_Z = \mu$. Therefore, $\overline{\varrho}^f \geq \overline{\varrho} \geq \varrho(\mu) \geq \overline{\varrho}^f$; that is, $\underline{\varrho}^f = \underline{\varrho} = \varrho(\mu)$.

c) This is a direct consequence of a) and b).

The above statements are applicable to TVaR and do not hold when VaR is used as a risk measure. Counterexamples can be easily constructed (see also Example 4.2 hereafter). In particular, there are situations in which using factor information yields improved bounds on VaR but not on TVaR. This observation raises the issue of whether the result of Embrechts et al. (2015) that, in the unconstrained setting, (asymptotically) large portfolios exhibit a larger dependence uncertainty spread for VaR than for TVaR carries over to the constrained setting. The answer to this question is negative, as the following example shows.

Example 4.2 (Dependence uncertainty spread of VaR *is not necessarily larger than that of* TVaR*)* Assume that the factor Z takes two values: $Z = 0$ with probability 0.95 and $Z = 1$ with probability 0.05. When $Z = 0$, we assume that all X_i $(i = 1, \ldots, n)$ are degenerated and take value 0.5. When $Z = 1$, X_i $(i = 1, \ldots, n)$ are Bernoulli distributed with parameter 0.5. Let n be even and take $\alpha = 0.9$. As for the unconstrained bounds, observe that S^c satisfies $P(S^c = 0) = 0.025$, $P(S^c = 0.5n) = 0.95$ and $P(S^c = n) = 0.025$. From results regarding VaR bounds (see e.g. Bernard et al. (2014)) it is easy to see that $\overline{\mathrm{VaR}}_\alpha = \overline{\mathrm{TVaR}}_\alpha = \frac{n}{4} + \frac{3\times 0.5n}{4} = 0.625n$ and that $\underline{\mathrm{VaR}}_\alpha < \underline{\mathrm{TVaR}}_\alpha = 0.5n$. Hence, the dependence uncertainty spread of VaR is larger than that of TVaR. In the constrained setting, since for any $S_Z \in \mathcal{S}(H)$, $P(S_Z = 0.5n) > 0.95$ and $\mathrm{VaR}_{0.9}(S_Z) = 0.5n$, there is no longer dependence uncertainty spread on

$\mathrm{VaR}_{0.9}(S_Z)$. However, the dependence uncertainty spread for TVaR remains unchanged. In particular, it is higher than in the case in which the VaR is used and this holds also as $n \to \infty$. □

Embrechts et al. (2015) show that in the unconstrained case the VaR of large portfolios is more sensitive to misspecification of the model (i.e., has a higher dependence uncertainty spread) than TVaR. Example 4.2 shows that, whether TVaR can be seen as less sensitive in this regard (i.e., has less dependence uncertainty spread), depends merely on the available set of information. In particular, when structural factor information is available as a source of dependence information, TVaR has, in general, no such advantage over VaR.

4.2 Approximation of VaR bounds based on TVaR bounds

Recall that the formula (2.13) for $\overline{\mathrm{VaR}}_\alpha^f$ remains difficult to evaluate practically. However, the TVaR bounds that we developed in the previous section can be used to determine an easy-to-evaluate upper bound for $\overline{\mathrm{VaR}}_\alpha^f$.

Indeed, note that under the condition $Z = z \in D$ the (conditional) VaR is bounded above by the conditional comonotonic TVaR, i.e., for all conditional sums S_z and all $\beta \in (0,1)$, it holds that

$$q_z(\beta) = \mathrm{VaR}_\beta(S_z) \le \mathrm{TVaR}_\beta(S_z) \le t_z(\beta) := \mathrm{TVaR}_\beta(S_z^c).$$

The above inequality implies that $(\overline{q}_z)^{-1}(\gamma) \ge t_z^{-1}(\gamma),\ \ \gamma \in \mathbb{R}$ and, therefore,

$$\begin{aligned}\overline{\mathrm{VaR}}_\alpha^f = \overline{b}^*(\alpha) = &\ \inf\left\{\gamma \in \mathbb{R} \,:\, \int (\overline{q}_z)^{-1}(\gamma)dG(z) \ge \alpha\right\} \\ \le &\ \inf\left\{\gamma \in \mathbb{R} \,:\, \int t_z^{-1}(\gamma)dG(z) \ge \alpha\right\} \\ =: &\ b_t^*(\alpha).\end{aligned}$$

So, for each set $\{Z = z\}$, we replace the VaR upper bound with the TVaR upper bound.

Proposition 4.3 (TVaR-**based bounds for** $\overline{\mathrm{VaR}}^f$)
For $\alpha \in (0,1)$ the sharp upper bound $\overline{\mathrm{VaR}}_\alpha^f$ in the partially specified risk factor model is bounded above by the TVaR-*based bound $b_t^*(\alpha)$ in* (4.2)*, i.e.,*

$$\overline{\mathrm{VaR}}_\alpha^f \le b_t^*(\alpha). \tag{4.1}$$

Note that, conditional on $Z = z$, one has a simple expression for $t_z(\beta)$, i.e.,

$$t_z(\beta) = \mathrm{TVaR}_\beta(S_z^c) = \sum_{i=1}^n \mathrm{TVaR}_\beta(X_{i|z}), \tag{4.2}$$

and thus $t_z(\beta)$ is easy to calculate. As a result, the calculation of the upper TVaR-based bound $b_t^*(\alpha)$ of $\overline{\mathrm{VaR}}_\alpha^f$ is much simpler than the calculation of $\overline{\mathrm{VaR}}_\alpha^f$. In particular, we avoid the iterated application of the RA algorithm.

There is also an alternative way to establish the simplified VaR bounds in (4.1). This method leads to a stochastic representation that is useful in evaluating the bounds by simulation. Define, for any $z \in D$, a random variable T_z^+ as

$$T_z^+ = \mathrm{TVaR}_V(S_z^c),$$

where $V \sim U(0,1)$ is a random variable that is uniformly distributed on $(0,1)$ and independent of Z and of $(S_z^c)_{z\in D}$. It is easy to simulate V and Z, and hence also the random variable T_Z^+, as well as to approximate its VaR. The following proposition therefore yields an interesting connection between T_Z^+ and the upper bound $b_t^*(\alpha)$ in Proposition 4.3.

Proposition 4.4 (Representation of TVaR-based bounds) *For $\alpha \in (0,1)$, we have*

$$\overline{\mathrm{VaR}}_\alpha^f \le \mathrm{VaR}_\alpha(T_Z^+) = b_t^*(\alpha). \tag{4.3}$$

Proof We only need to show that $\mathrm{VaR}_\alpha(T_Z^+) = b_t^*(\alpha)$. Using comonotone additivity of TVaR, we obtain

$$T_z^+ = \mathrm{TVaR}_V(S_z^c) = \sum_{i=1}^n \mathrm{TVaR}_V(X_{i|z}) =: \sum_{i=1}^n T_{i|z}^+.$$

Note that $\mathrm{VaR}_\alpha(T_{i|z}^+) = \mathrm{TVaR}_\alpha(X_{i|z})$. From the first representation for VaR in Proposition 2.4 we obtain $\mathrm{VaR}_\alpha(T_Z^+) = b_t^*(\alpha)$. □

Remark 4.5 In a similar way, we also obtain approximations of the lower sharp VaR-bound $\underline{\mathrm{VaR}}_\alpha^f$. Define $T_z^- := \mathrm{LTVaR}_V(S_z^c)$; then

$$\mathrm{VaR}_\alpha(T_Z^-) \le \underline{\mathrm{VaR}}_\alpha^f, \tag{4.4}$$

i.e., $\mathrm{VaR}_\alpha(T_Z^-)$ is a lower bound for $\underline{\mathrm{VaR}}_\alpha^f$. Here, $\mathrm{LTVaR}_\alpha(S_z^c)$ is the left TVaR, we have that $\mathrm{LTVaR}_\alpha(S_z^c) = \frac{1}{\alpha}\int_0^\alpha \mathrm{VaR}_u(S_z^c)du$. □

4.3 Asymptotic sharpness

In the unconstrained setting, Puccetti and Rüschendorf (2014) showed that $\mathrm{TVaR}_\alpha(S^c)$ is an asymptotically sharp bound on $\mathrm{VaR}_\alpha(S)$, i.e.,

$$\frac{\overline{\mathrm{VaR}}_\alpha}{\overline{\mathrm{TVaR}}_\alpha} \to 1, \text{ as } n \to \infty \tag{4.5}$$

holds under some moment conditions; weaker conditions that also ensure asymptotic equivalence can be found in Puccetti and Rüschendorf (2014), Puccetti et al. (2013), Wang and Wang (2015) and Embrechts et al. (2015). In this section, we extend these results to the constrained setting by showing that the TVaR-based bounds $\mathrm{VaR}_\beta(T_Z^+)$ and $\mathrm{VaR}_\beta(T_Z^-)$ that we developed in Section 4.2 are asymptotically sharp bounds for VaR, which includes the equivalence (4.5) as a special case (in which Z is independent of X).

In the constrained setting, it is clear that $\mathrm{TVaR}_\alpha(S_Z^c)$ is a bound on $\mathrm{VaR}_\alpha(S_Z)$ but an asymptotic equivalence for the ratio $\overline{\mathrm{VaR}}_\alpha^f/\overline{\mathrm{TVaR}}_\alpha^f$ fails to hold in general. A simple counterexample would be to choose $F_{i|z}$, $z \in D$, $i = 1, 2, \ldots$, to be all degenerate distributions. In this case, dependence uncertainty is no longer relevant and all elements in $\mathcal{S}(H)$ are distributed as μ_Z (which depends on n). There is no hope that $\mathrm{VaR}_\alpha(\mu_Z)/\mathrm{TVaR}_\alpha(\mu_Z) \to 1$ would hold generally. In fact, in order to ensure asymptotic equivalence for the ratio $\overline{\mathrm{VaR}}_\alpha^f/\overline{\mathrm{TVaR}}_\alpha^f$, one would need asymptotic mixability, which can intuitively be seen as a (strengthened) version of the law of large numbers; see the related discussions in Bernard et al. (2015), Wang (2014) and Embrechts et al. (2014). Although an asymptotic equivalence fails to hold for $\overline{\mathrm{VaR}}_\alpha(S_Z)/\overline{\mathrm{TVaR}}_\alpha(S_Z)$, we have the following asymptotic equivalence theorem for risk aggregation with a risk factor.

Proposition 4.6 (Asymptotic sharpness of TVaR-based bounds) *Suppose that for any $z \in D$ and some $k > 1$,*

a) $E|X_{i,z} - EX_{i,z}|^k < M < \infty$ *for* $i = 1, \ldots, n$;
b) $\liminf_{n\to\infty} n^{-1/k} \sum_{i=1}^n E(X_{i,z}) = \infty$.

Then, for $\alpha \in (0,1)$, as $n \to \infty$,

$$\frac{\overline{\mathrm{VaR}}_\alpha^f}{\mathrm{VaR}_\alpha\left(T_Z^+\right)} \to 1 \text{ and } \frac{\underline{\mathrm{VaR}}_\alpha^f}{\mathrm{VaR}_\alpha\left(T_Z^-\right)} \to 1.$$

Proof From conditions a), b) for the sequence $\{F_{i,z};\ i \in \mathbb{N}\}$ it follows that the conditions from Theorem 3.3 of Embrechts et al. (2015) are satisfied for all $\beta \in (0,1)$ and $z \in D$. Therefore, we have

$$\frac{\overline{\mathrm{VaR}}_\beta(S_z)}{\overline{\mathrm{TVaR}}_\beta(S_z)} = \frac{\overline{q}_z(\beta)}{t_z(\beta)} \to 1$$

as $n \to \infty$ for each $\beta \in (0,1)$ and $z \in D$. Thus, we have, for a $U(0,1)$-distributed random variable V independent of Z, $\overline{q}_Z(V)/t_Z(V) \to 1$ as $n \to \infty$, and this convergence holds for all $\omega \in \Omega$. Then, from (2.13), we can conclude that

$$\frac{\overline{\mathrm{VaR}}_\alpha^f}{\mathrm{VaR}_\alpha\left(T_Z^+\right)} = \frac{\mathrm{VaR}_\alpha(\overline{q}_Z(V))}{\mathrm{VaR}_\alpha(t_Z(V))} \to 1.$$

The case for lower bounds is similar.

In the following example we consider a low-dimensional portfolio ($n = 2$) and compare the sharp VaR bounds with the ones that are asymptotically sharp.

Example 4.7 (Approximations using TVaR*-based bounds for normal distributions)* We consider the setting as in Example 2.3. In Table 4.1 we compare sharp VaR bounds with the TVaR-based bounds for VaR (Proposition 4.4) for various parameter values of the correlations r_1 and r_2.

The values of VaR_α, $\overline{\mathrm{VaR}}^f_\alpha$ and $\underline{\mathrm{VaR}}^f_\alpha$ are taken from Example 2.3. To compute $\mathrm{VaR}_\alpha\left(T_Z^+\right)$ and $\mathrm{VaR}_\alpha\left(T_Z^-\right)$ we simulate z from $\mathcal{N}(0,1)$ and v from $\mathcal{U}(0,1)$. Next, we compute, for $r_1 = r_2$, $T_z^+ = 2r_1 z + 2\sqrt{1-r_1^2}\frac{\phi(\Phi^{-1}(v))}{1-v}$ and, for $r_1 = -r_2$, $T_z^+ = 2\sqrt{1-r_1^2}\frac{\phi(\Phi^{-1}(v))}{1-v}$. By generating many values for v and z, we can accurately approximate $\mathrm{VaR}_{95\%}$ of T_Z^+. We proceed similarly for T_Z^-.

$\alpha = 0.95$	VaR_α	$\overline{\mathrm{VaR}}^f_\alpha$	$\mathrm{VaR}_\alpha\left(T_Z^+\right)$	$\underline{\mathrm{VaR}}^f_\alpha$	$\mathrm{VaR}_\alpha\left(T_Z^-\right)$
$r_1 = r_2 = 0$	2.33	3.92	4.12	−0.12	−0.21
$r_1 = r_2 = 0.5$	2.60	3.92	4.11	0.82	0.68
$r_1 = r_2 = 0.8$	2.98	3.88	4.01	1.89	1.78
$r_1 = r_2 = 1$	3.29	3.28	3.28	3.28	3.28
$r_1 = -r_2 = 0$	2.33	3.92	4.13	−0.12	−0.21
$r_1 = -r_2 = -0.5$	2.01	3.39	3.57	−0.11	−0.18
$r_1 = -r_2 = -0.8$	1.40	2.35	2.47	−0.07	−0.13
$r_1 = -r_2 = -1$	0.00	0.00	0.00	0.00	0.00

Table 4.1 Comparison between sharp VaR bounds and their TVaR-based approximations.

For this small portfolio ($n = 2$) the asymptotic sharp bounds perform reasonably well. This is rather expected since for normally distributed risks VaR and TVaR are not very different. □

Example 4.8 (Approximations using TVaR*-based bounds for Pareto distributions)* We consider a Pareto risk model in which $n = 2$, $P(Z = 1) = P(Z = 2) = 1/2$, and X_1, X_2 are Pareto(θ, Z) distributed conditional on $Z = 1, 2$. That is, $P(X_{1,z} > x) = z^\theta x^{-\theta}$, $x \geq z$, where $\theta > 1$. A smaller value of θ indicates a heavier-tailed distribution of X_1, X_2.

For this simple setting, using a result in Rüschendorf (1982) we can analytically obtain the values $\overline{M_z}(t)$, $z = 1, 2$, $t \in \mathbb{R}$, and consequently obtain $\overline{M}^f(t)$, $t \in \mathbb{R}$ by Proposition 2.1. Taking an inverse of $\overline{M}^f$ we can calculate $\overline{\mathrm{VaR}}^f_\alpha$. Omitting all intermediate steps (which are simple exercises), we have

$$\overline{\mathrm{VaR}}^f_\alpha = (2^\theta + 4^\theta)^{1/\theta}(1-\alpha)^{-1/\theta}, \quad \alpha \in (0,1).$$

On the other hand, one can easily calculate $T_z^+ = \mathrm{TVaR}_V(S_z^c)$, $z = 1, 2$, and consequently we obtain the distribution of T_Z^+. Omitting all intermediate steps, we have

$$\mathrm{VaR}_\alpha(T_Z^+) = 2^{-1/\theta}\frac{\theta}{\theta - 1}(2^\theta + 4^\theta)^{1/\theta}(1-\alpha)^{-1/\theta}, \quad \alpha \in (0,1).$$

Therefore, for this model in which $n = 2$, we have

$$\frac{\mathrm{VaR}_\alpha(T_Z^+)}{\overline{\mathrm{VaR}}_\alpha^f} = 2^{-1/\theta}\frac{\theta}{\theta - 1}.$$

It is clear that the above ratio is a decreasing function of θ and is independent of α. We report some numbers for different choices of (α, θ) in Table 4.2.

(α, θ)	$\overline{\mathrm{VaR}}_\alpha^f$	$\mathrm{VaR}_\alpha\left(T_Z^+\right)$	$\mathrm{VaR}_\alpha(T_Z^+)/\overline{\mathrm{VaR}}_\alpha^f$
(0.95,2)	20.000	28.284	1.414
(0.95,5)	7.327	7.973	1.088
(0.95,10)	5.398	5.596	1.037
(0.95,20)	4.646	4.724	1.017
(0.99,2)	44.721	63.246	1.414
(0.99,5)	10.110	11.001	1.088
(0.99,10)	6.340	6.573	1.037
(0.95,20)	5.036	5.120	1.017

Table 4.2 Comparison between sharp VaR bounds and their TVaR-based approximations.

From Table 4.2, it is clear and not surprising that for lighter-tailed Pareto distributions, the TVaR-based approximation performs better than in the case of heavy-tailed distributions. Due to the fact that a reliable calculation of $\overline{\mathrm{VaR}}_\alpha^f$ is not available in general (see discussions in Section 4.5), we only report numbers for $n = 2$ and the asymptotics of Proposition 4.6 have not fully kicked in. Puccetti and Rüschendorf (2014) studied the performance of TVaR-based bounds in the unconstrained case and used the RA to approximate $\overline{\mathrm{VaR}}_\alpha$. For a portfolio containing heavy-tailed Pareto distributions ($\theta = 2$), they conclude that TVaR-based bounds yield good approximations for portfolio size $n \geq 10$. □

4.4 Adding variance information

In this section, we show that the bounds can be further sharpened if (conditional) variance information is also available. The idea of using the variance to sharpen the unconstrained bounds can be found in Bernard et al. (2015), where it is shown that doing so can have a significant impact on the unconstrained VaR bounds. Here, we consider conditional variance information in addition to factor information and discuss how this can be useful in improving the bounds on VaR and TVaR.

Consider $v := (v_z)_{z\in D}$, $v_z \geq 0$. We define the partially specified factor model with variance information as

$$A(H, v) = \{X \in A(H) \,:\, \mathrm{var}(S|Z = z) \leq v_z^2,\ z \in D\},$$

where we assume that it contains at least one element. Hence, v_z^2 provides a bound on the conditional variance of S_z, $z \in D$. We study bounds on VaR and TVaR, i.e., we consider the problems

$$\overline{\mathrm{VaR}}_\alpha^{f,v} = \sup\{\mathrm{VaR}_\alpha(S) \,:\, X \in A(H,v)\}$$

and

$$\overline{\mathrm{TVaR}}_\alpha^{f,v} = \sup\{\mathrm{TVaR}_\alpha(S) \,:\, X \in A(H,v)\}.$$

We can consider the lower bound problems in a similar way. We denote the corresponding infima by $\underline{\mathrm{VaR}}_\alpha^{f,v}$ and $\underline{\mathrm{TVaR}}_\alpha^{f,v}$.

Proposition 4.9 (VaR bounds in the factor model with variance information)
For $\alpha \in (0,1)$, we have

$$\overline{\mathrm{VaR}}_\alpha^{f,v} \le \mathrm{VaR}_\alpha\left(\min\left(\mathrm{TVaR}_U(S_Z^c), \mu_Z + v_Z\sqrt{\frac{U}{1-U}}\right)\right)$$

and

$$\underline{\mathrm{VaR}}_\alpha^{f,v} \ge \mathrm{VaR}_\alpha\left(\max\left(\mathrm{LTVaR}_U(S_Z^c), \mu_Z - v_Z\sqrt{\frac{1-U}{U}}\right)\right),$$

where $U \sim U(0,1)$ is independent of Z and $(S_z^c)_{z\in D}$.

The proof of Proposition 4.9 is provided in the appendix.

Proposition 4.10 (TVaR bounds in the factor model with variance information) *For $\alpha \in (0,1)$, we have*

$$\overline{\mathrm{TVaR}}_\alpha^{f,v} \le \mathrm{TVaR}_\alpha\left(\min\left(\mathrm{VaR}_U(S_Z^c), \mu_Z + v_Z\sqrt{\frac{U}{1-U}}\right)\right),$$

where $U \sim U(0,1)$ independent of Z and $(S_z^c)_{z\in D}$.

The proof of Proposition 4.10 is given in the appendix.

Remark 4.11 Bernard et al. (2015) study VaR bounds when, in addition to the marginal information, information on the (unconditional) variance of the sum is also provided; they consider the problem

$$\overline{\mathrm{VaR}}_\alpha^{v} = \sup\{\mathrm{VaR}_\alpha(S) \,:\, X \in A_1(F), \quad \mathrm{var}(S) \le v^2\}$$

where $v \ge 0$ is an admissible variance constraint. In a similar way, they consider $\underline{\mathrm{VaR}}_\alpha^{v}$ and obtain that

$$\overline{\mathrm{VaR}}_\alpha^{v} \le \min\left(\mu + v\sqrt{\frac{\alpha}{1-\alpha}},\ \mathrm{TVaR}_\alpha(S^c)\right) \tag{4.6}$$

.and

$$\underline{\mathrm{VaR}}_\alpha^{v} \ge \max\left(\mu - v\sqrt{\frac{1-\alpha}{\alpha}},\ \mathrm{LTVaR}_\alpha(S^c)\right) \tag{4.7}$$

□

4.5 Practical methods for calculating VaR bounds

Evaluating the sharp bounds $\overline{\mathrm{VaR}}_\alpha^f$ and $\underline{\mathrm{VaR}}_\alpha^f$ is not straightforward. However, the theoretical results developed in Sections 1–4 make it possible to propose practical methods by which to approximate the risk bounds. We explain these approximations for the case of the upper bounds.

Asymptotic bounds: We approximate $\overline{\mathrm{VaR}}_\alpha^f$ from above by $\mathrm{VaR}_\alpha(T_Z^+)$; see Proposition 4.4. Observe that $T_z^+ = \mathrm{TVaR}_V(S_z^c) = \sum_{i=1}^n \mathrm{TVaR}_V(X_{i|z})$ where V is a standard uniformly distributed random variable that is taken independent of Z. Hence, the computation of $\mathrm{VaR}_\alpha(T_Z^+)$ can be performed in a straightforward way using Monte Carlo simulations.

Repeated RA: In the following, we use a discrete approximation for G. We consider the following steps:

a) Define $\beta_j := \frac{j-1/2}{m} \in [0,1]$, $j = 1, 2, ..., m$, where m is a large integer.
b) Use the RA to determine, for $z_i \in D$ and $\beta_j \in [0,1]$, the (approximations for) $\overline{q}_{z_i}(\beta_j) = \overline{\mathrm{VaR}}_{\beta_j}(S_{z_i})$
c) From the $\overline{q}_{z_i}(\beta)$ one obtains an approximation for $\mathrm{VaR}_\alpha(\bar{q}_Z(V)) = \mathrm{VaR}_\alpha^f$ (Proposition 2.5) where V is a standard uniformly distributed random variable that is taken independent of Z.

Analytical method: For some special cases of distributions, $\overline{M}_z^f$ or $\overline{q}_z(\beta)$ may be available analytically (for instance, in Example 4.8). In these cases, analytically calculating $\overline{M}^f$ and $\overline{\mathrm{VaR}}^f$ may be possible via Propositions 2.1 and 2.5. However, as mentioned in Section 2, this procedure involves taking inverses and mixtures of functions repeatedly and there is no guarantee that the resulting functions still have an explicit form.

Discussion: The first method thus computes the asymptotically TVaR-based upper bound. It is fast and does not suffer from the curse of dimensionality. While this method overestimates the true sharp bound, the degree of overestimation is typically small, in particular for large portfolios (Proposition 4.6); see also the numerical evidence provided in Embrechts et al. (2014) and in Bernard et al. (2015). Hence, we recommend it as a standard method. The second method is essentially based on an application of the RA to conditional distributions. It is well-known that the RA is a very suitable method for approximating numerically sharp bounds, and one can expect this method to provide excellent approximations for sharp bounds. However, a drawback of this method is that a repeated application of the RA is needed, which can make its application time consuming, especially when Z can take many values. The third method heavily relies on the distributions in the model. In case $n = 2$, $\overline{M}_z^f$ can be calculated which involves an inverse of the VaR bounds (as in Rüschendorf (1982); see Example 4.8) for each $z \in D$. For $n \geq 3$, analytical formulas are very limited; some results can be found in Wang et al. (2013).

5 Application to Credit Risk Portfolios

Recent financial crises have shown that credit portfolios require careful monitoring. In this regard, many financial institutions and regulatory frameworks, such as Basel III and Solvency II, rely on a Bernoulli mixture model to measure the risk. In the industry, this model is also known as the KMV model (Gordy (2003)), and we refer to this terminology without further ado. Specifically, the risks X_i $(i = 1, \dots, n)$ are modeled as

$$X_i = \begin{cases} 0, & \text{if } \sqrt{r_i}Z + \sqrt{1-r_i}\varepsilon_i > \Phi^{-1}(q_i), \\ \frac{e^{a_i Z}}{1+e^{a_i Z}}, & \text{otherwise,} \end{cases}$$

in which $q_i \in (0,1)$ and Z, ε_i are standard normally distributed and independent. Note that $\text{corr}(X_i, X_j) = r_i r_j$. Under the KMV specifications, it is further assumed that the idiosyncratic risks ε_i are mutually independent. Under these assumptions, risk measures of $S = \sum_{i=1}^n X_i$ such as VaR can be computed using Monte Carlo simulations. We challenge the dependence assumptions among the X_i and compute the bounds on VaR correspondingly.

Assuming that only the marginal distributions are known, we assess the bounds $\underline{\text{VaR}}_\alpha$ and $\overline{\text{VaR}}_\alpha$ using their asymptotic versions $\text{VaR}_\alpha(\text{LTVaR}_U(S^c))$ and $\text{VaR}_\alpha(\text{TVaR}_U(S^c))$. Next, we add dependence information in various ways. First, using the structural factor information, we assess the bounds $\overline{\text{VaR}}_\alpha^f$ and $\underline{\text{VaR}}_\alpha^f$ using their asymptotic versions discussed in Section 4.3; see the expressions (4.3) and (4.4). Second, we add variance information and approximate $\underline{\text{VaR}}_\alpha^{f,v}$ and $\overline{\text{VaR}}_\alpha^{f,v}$ using the bounds established in Proposition 4.9. Finally, we approximate the variance bounds $\underline{\text{VaR}}_\alpha^v$ and $\overline{\text{VaR}}_\alpha^v$ using the expressions (4.6) and (4.7) (assuming that the variance constraint writes as $v_z = \beta\mu_z$).

We study four homogeneous cases (1)-(4) and two heterogeneous cases (5) and (6) and display the results in Table 5.1. Our findings can be summarized as follows. In the unconstrained case, the dependence uncertainty spread is very wide. In particular, the VaR numbers that one obtains by applying the standard KMV model (labeled as VaR_α^{KMV}) lie far away from the unconstrained upper bounds. Adding factor information improves the lower bound significantly but not the upper bounds. This feature is to be expected, as the factor information induces positive dependence among the risks in that the risks are perfectly dependent, conditionally on the ε_i. By contrast, adding variance information improves the upper but not the lower bound. This feature is also to be expected, as the variance of a credit portfolio loss is driven mainly by high outcomes and putting a constraint on variance thus implies that upper VaRs become reduced. All in all, using the factor information supplemented with conditional variance information reduces the unconstrained bounds tremendously.

	α	VaR_α^{KMV}	$(\underline{\text{VaR}}_\alpha, \overline{\text{VaR}}_\alpha)$	$(\underline{\text{VaR}}_\alpha^f, \overline{\text{VaR}}_\alpha^f)$	$(\underline{\text{VaR}}_\alpha^v, \overline{\text{VaR}}_\alpha^v)$	$(\underline{\text{VaR}}_\alpha^{f,v}, \overline{\text{VaR}}_\alpha^{f,v})$
(1)	95 %	16.5	(0 , 125)	(0 , 125)	(4.042 , 47.76)	(6.18 , 45.58)
	99.5 %	29.5	(5.037 , 250)	(8.181 , 250)	(5.586 , 140.6)	(13.89 , 126.7)
(2)	95 %	29.5	(0 , 125)	(0 , 123.8)	(1.775 , 91.27)	(7.034 , 68.21)
	99.5 %	83.5	(5.037 , 250)	(29.35 , 250)	(5.037 , 250)	(32.88 , 249.9)
(3)	95 %	32.96	(0 , 189.8)	(0 , 182.2)	(5.134 , 92.25)	(11.94 , 85.91)
	99.5 %	59	(7.014 , 499.7)	(16.74 , 499.1)	(8.132 , 277.2)	(28.39 , 243.5)
(4)	95 %	58.89	(0 , 235.2)	(0 , 235.6)	(2.778 , 182.4)	(13.99 , 137.4)
	99.5 %	168	(9.366 , 500)	(62.6 , 499.9)	(9.366 , 500)	(68.46 , 498)
(5)	95 %	66.92	(0.143 , 275.3)	(0.4212 , 264.4)	(4.231 , 197.6)	(17.12 , 156.9)
	99.5 %	175	(11.53 , 484.9)	(75.42 , 481.8)	(11.53 , 484.9)	(75.77 , 480)
(6)	95 %	56.88	(0 , 228.9)	(0 , 226.4)	(2.462 , 182.2)	(13.24 , 132.4)
	99.5 %	175	(9.102 , 499.9)	(68.02 , 499.2)	(9.102 , 499.9)	(71.4 , 472.2)

Table 5.1 VaR bounds for credit risk portfolios. Homogeneous cases: (1) $a = 0$, $n = 500$, $q = 2.5\,\%$, $r = 10\,\%$; (2) $a = 0$, $n = 500$, $q = 2.5\,\%$, $r = 40\,\%$; (3) $a = -4$, $n = 500$, $q = 2.5\,\%$, $r = 10\,\%$; (4) $a = -4$, $n = 500$, $q = 2.5\,\%$, $r = 40\,\%$. Heterogeneous cases: (5) $a = -4$, $n = 500$, $q_i \in \{0.1\,\%, ..., 9.9\,\%\}$, $r_i = r = 40\%$; (6) $a = -4$, $n = 500$, $q = 2.5\,\%$, $r_i \in \{10.6\,\%, \ldots, 69.4\,\%\}$.

6 Conclusion

In this paper we study risk bounds for factor models with known marginal distributions of the components but with a dependence structure that is not completely specified. Our structural assumption regarding dependence is prevalent in the statistical and finance literature and can be backtested. We determine upper and lower bounds for VaR and convex risk measures such as TVaR and compare their distance (i.e., the dependence uncertainty spread) with the one obtained when only marginal distributions are assumed to be known. Specifically, we obtain asymptotic sharp bounds for VaR and show that they are straightforward to compute. We also show that in a factor-constrained setting the dependence uncertainty spread of VaR is not necessarily larger than that of TVaR. We obtain further improved bounds by using variance information. The reduction in the dependence uncertainty spread that we observe depends on the magnitude of the common risk factor in comparison to the idiosyncratic factors. All in all, the results of the paper show that the assumption of a partially specified factor model is a flexible tool with a wide range of possible applications and with a promising capability to reduce the risk bounds that are based on knowledge solely of marginal information.

References

N. Bäuerle and A. Müller. Stochastic orders and risk measures: Consistency and bounds. *Insurance Math. Econom.*, 38(1):132–148, 2006.

C. Bernard and S. Vanduffel. A new approach to assessing model risk in high dimensions. *J. Banking and Finance*, 58(1):166–168, 2015.

C. Bernard, X. Jiang, and R. Wang. Risk aggregation with dependence uncertainty. *Insurance Math. Econom.*, 54:93–108, 2014.

C. Bernard, L. Rüschendorf, and S. Vanduffel. Value-at-Risk bounds with variance constraints. *J. of Risk and Insurance*, 2015. doi: 10.1111/jori.12108.

C. Bernard, M. Denuit, and S. Vanduffel. Measuring portfolio risk under partial dependence information. *J. of Risk and Insurance*, 2016. doi: 10.1111/jori.12165.

C. Bernard, L. Rüschendorf, S. Vanduffel, and J. Yao. How robust is the value-at-risk of credit risk portfolios? *European J. of Finance*, 23(6):507–534, 2017.

V. Bignozzi, G. Puccetti, and L. Rüschendorf. Reducing model risk via positive and negative dependence assumptions. *Insurance: Math. Econom.*, 61:17–26, 2015.

C. Burgert and L. Rüschendorf. Consistent risk measures for portfolio vectors. *Insurance Math. Econom.*, 38(2):289–297, 2006.

M. M. Carhart. On persistence in mutual fund performance. *J. Finance*, 52 (1):57–82, 1997.

G. Chamberlain and M. Rothschild. Arbitrage, factor structure, and mean-variance analysis on large asset markets, 1982.

G. Connor and R. A. Korajczyk. A test for the number of factors in an approximate factor model. *J. Finance*, 48(4):1263–1291, 1993.

R. Cont, R. Deguest, and G. Scandolo. Robustness and sensitivity analysis of risk measurement procedures. *Quant. Finance*, 10(6):593–606, 2010.

J. Daníelsson, B. Jorgensen, S. Mandira, G. Samorodnitsky, and C. G. de Vries. Subadditivity re-examined: The case for Value-at-Risk. Discussion paper, Financial Markets Group, London School of Economics and Political Science, 2005.

G. Deelstra, I. Diallo, and M. Vanmaele. Bounds for Asian basket options. *J. Comput. Appl. Math.*, 218(2):215–228, 2008.

J. Dhaene, S. Vanduffel, M. Goovaerts, R. Kaas, Q. Tang, and D. Vyncke. Risk measures and comonotonicity: a review. *Stochastic models*, 22(4):573–606, 2006.

P. Embrechts, G. Puccetti, and L. Rüschendorf. Model uncertainty and VaR aggregation. *J. Banking and Finance*, 37(8):2750–2764, 2013.

P. Embrechts, R. Puccetti, L. Rüschendorf, R. Wang, and A. Beleraj. An academic response to Basel 3.5. *Risks*, 2(1):25–48, 2014.

P. Embrechts, B. Wang, and R. Wang. Aggregation-robustness and model uncertainty of regulatory risk measures. *Finance Stoch.*, 19(4):763–790, 2015.

S. Emmer, M. Kratz, and D. Tasche. What is the best risk measure in practice? a comparison of standard measures. *Preprint ISSN : 1291-9616, ESSEC Business School*, 2014.

R. F. Engle, V. K. Ng, and M. Rothschild. Asset pricing with a factor-arch covariance structure: Empirical estimates for treasury bills. *J. of Econometrics*, 45(1):213–237, 1990.

E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *J. of Financial Economics*, 33(1):3–56, 1993.

H. Föllmer and A. Schied. *Stochastic Finance. An Introduction in Discrete Time.* Berlin: de Gruyter, 2nd revised and extended edition, 2004.

T. Gneiting. Making and evaluating point forecasts. *J. Am. Stat. Assoc.*, 106: 746–762, 2011. doi: 10.1198/jasa.2011.r10138.

M. B. Gordy. A comparative anatomy of credit risk models. *J. Banking and Finance*, 24(1):119–149, 2000.

M. B. Gordy. A risk-factor model foundation for ratings-based bank capital rules. *J. of Financial Intermediation*, 12(3):199–232, 2003.

J. E. Ingersoll. Some results in the theory of arbitrage pricing. *J. Finance*, 39 (4):1021–1039, 1984.

P. Jorion. *Value-at-Risk: The New Benchmark for Managing Financial Risk.* New York: McGraw-Hill, 2006.

E. Jouini, W. Schachermayer, and N. Touzi. Law invariant risk measures have the Fatou property. In S. Kusuoka and A. Yamazaki, editors, *Advances in Mathematical Economics*, pages 49–71. Springer, 2006.

R. Kaas, J. Dhaene, and M. J. Goovaerts. Upper and lower bounds for sums of random variables. *Insurance Math. Econom.*, 27(2):151–168, 2000.

V. Krätschmer, A. Schied, and H. Zähle. Qualitative and infinitesimal robustness of tail-dependent statistical functionals. *J. Multivariate Anal.*, 103(1): 35–47, 2012.

V. Krätschmer, A. Schied, and H. Zähle. Comparative and qualitative robustness for law-invariant risk measures. *Finance Stoch.*, 18(2):271–295, 2014.

A. Lewbel. The rank of demand systems: theory and nonparametric estimation. *Econometrica*, 59:711–730, 1991.

I. Meilijson and A. Nadas. Convex majorization with an application to the length of critical paths. *J. Appl. Probab.*, 16:671–677, 1979.

G. Puccetti and L. Rüschendorf. Computation of sharp bounds on the distribution of a function of dependent risks. *J. Comput. Appl. Math.*, 236(7): 1833–1840, 2012.

G. Puccetti and L. Rüschendorf. Asymptotic equivalence of conservative VaR- and ES-based capital charges. *J. Risk*, 16(3):1–19, 2014.

G. Puccetti, B. Wang, and R. Wang. Complete mixability and asymptotic equivalence of worst-possible VaR and ES estimates. *Insurance Math. Econom.*, 53(3):821–828, 2013.

S. A. Ross. The arbitrage theory of capital asset pricing. *J. Econom. Theory*, 13(3):341–360, 1976.

L. Rüschendorf. Random variables with maximum sums. *Adv. Appl. Prob.*, 14:623–632, 1982.

L. Rüschendorf. The Wasserstein distance and approximation theorems. *Z. Wahrscheinlichkeitstheor. Verw. Geb.*, 70:117–129, 1985.

A. A. Santos, F. J. Nogales, and E. Ruiz. Comparing univariate and multivariate models to forecast portfolio value-at-risk. *J. Financial Econometrics*, 11 (2):400–441, 2013.

W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *J. Finance*, 19(3):425–442, 1964.

S. Vanduffel, Z. Shang, L. Henrard, J. Dhaene, and E. A. Valdez. Analytic bounds and approximations for annuities and asian options. *Insurance: Mathematics and Economics*, 42(3):1109–1117, 2008.

M. Vanmaele, G. Deelstra, J. Liinev, J. Dhaene, and M. Goovaerts. Bounds for the price of discrete arithmetic Asian options. *J. Comput. Appl. Math.*, 185(1):51–90, 2006.

B. Wang and R. Wang. The complete mixability and convex minimization problems with monotone marginal densities. *J. Multivariate Anal.*, 102(10): 1344–1360, 2011.

B. Wang and R. Wang. Extreme negative dependence and risk aggregation. *J. Multivariate Anal.*, 136:12–25, 2015.

B. Wang and R. Wang. Joint mixability. *Math. Oper. Res.*, 41(3):808–826, 2016.

R. Wang. Asymptotic bounds for the distribution of the sum of dependent random variables. *J. Appl. Probab.*, 51(3):780–798, 2014.

R. Wang, L. Peng, and J. Yang. Bounds for the sum of dependent risks and worst Value-at-Risk with monotone marginal densities. *Finance Stoch.*, 17 (2):395–417, 2013.

A Proof of Proposition 2.1

For any admissible risk vector $X \in A(H)$ we have that the conditional distribution of $X_i|Z = z$ is given by $F_{i|z}$. Therefore, conditionally under $Z = z$, the random vector X has marginal distributions $F_{i|z}$, $1 \leq i \leq n$. As a consequence, we obtain, by conditioning,

$$P\left(\sum_{i=1}^{n} X_i \geq t\right) = \int P\left(\sum_{i=1}^{n} X_i \geq t | Z = z\right) dG(z) \leq \int \overline{M}_z(t) dG(z),$$

and thus $\overline{M}^f(t) \leq \int \overline{M_z}(t) dG(t)$.

Conversely, let $X_z = (X_{i,z})$ be random vectors with marginal distributions $F_{i|z}$ such that, for given $\varepsilon > 0$,

$$P\left(\sum_{i=1}^{n} X_{i,z} \geq t\right) \geq \overline{M_z}(t) - \varepsilon. \tag{A.1}$$

The risk vector X has a representation as a mixture model: $X = X_{Z}.$, where Z is a random variable with distribution G, independent of $(X_{i,z})$. Then, by conditioning, we obtain that (X, Z) is admissible, i.e., $X \in A(H)$ and

$$P\left(\sum_{i=1}^{n} X_i \geq t\right) \geq \int \overline{M_z}(t) dG(z) - \varepsilon. \tag{A.2}$$

As a result, (A.1) and (A.2) establish equality in (2.4). The case of the lower bound is proved in a similar way. □

Remark A.1 (Existence of worst case distributions.) By a measurable selection result as in Rüschendorf (1985), a worst case distribution for $\overline{M}^f$ exists, and thus the ε-argument in the proof of Proposition 2.1 could be avoided in the case of the upper bound. However, the lower bounds $\underline{M}^f$ and $\underline{M_z}(t)$ are only attainable when we modify the definition of the Value-at-Risk slightly (see Bernard et al. (2014, 2015)).

B Proof of Proposition 3.2

a) Consider the vector X_Z^c having components $F_{i|Z}^{-1}(U)$ and observe that their conditional dfs are $F_{i|z}$ and that their marginal dfs are F_i. Hence, $X_Z^c \in A(H)$ and $S_Z^c \in \mathcal{S}(H)$. Furthermore, for any $X \in A(H)$ we can use the mixture representation X_Z for X with $X_{i,z} = F_{i|z}^{-1}(U_{i,z})$, as in Section 2. From the convex ordering result in (3.2), it follows that

$$S_z = \sum_i X_{i|z} \leq_{cx} \sum_i F_{i|z}^{-1}(U).$$

This implies, by conditioning, $S_Z \preceq_{cx} \sum_i F_{i|Z}^{-1}(U) = S_Z^c$.

b) Since ϱ is consistent with convex order, the result follows from a).

c) The summands of S_Z^c having dfs F_i, the result follows from (3.2). □

C Proof of Proposition 4.9

For any $X_Z \in A(H)$, it holds that

$$\begin{aligned}\mathrm{VaR}_\alpha(S_Z) &= \mathrm{VaR}_\alpha(\mathrm{VaR}_U(S_Z))\\ &\leq \mathrm{VaR}_\alpha\left(\min\left(\mathrm{TVaR}_U(S_Z^c), \mu_Z + v_Z\sqrt{\frac{U}{1-U}}\right)\right),\end{aligned}$$

where have used that for all $z \in D$, $u \in (0,1)$, $\mathrm{VaR}_u(S_z) \leq \mathrm{TVaR}_u(S_z^c)$ and $\mathrm{VaR}_u(S_z) \leq \mu_z + v_z\sqrt{\frac{u}{1-u}}$ (Cantelli bound). This shows the desired result for $\overline{\mathrm{VaR}}_\alpha^f$. The case of $\underline{\mathrm{VaR}}_\alpha^f$ is similar. □

D Proof of Proposition 4.10

For any $X_Z \in A(H)$, it holds that $\mathrm{TVaR}_\alpha(S_Z) \leq \mathrm{TVaR}_\alpha(S_Z^c) = \mathrm{TVaR}_\alpha(\mathrm{VaR}_U(S_Z^c))$. Furthermore, $\mathrm{TVaR}_\alpha(S_Z) = \mathrm{TVaR}_\alpha(\mathrm{VaR}_U(S_Z))$ and for all $z \in D$, $u \in (0,1)$, $\mathrm{VaR}_u(S_z) \leq \mu_z + v_z\sqrt{\frac{u}{1-u}}$ (Cantelli bound). Hence, by combining we obtain the desired result.