An efficient approach to quantile capital allocation and sensitivity analysis

Vali Asimit
City University London

Liang Peng
Georgia State University

Ruodu Wang
University of Waterloo

Alex Yu
Georgia State University

This version: December 2018*

Abstract

In various fields of applications such as capital allocation, sensitivity analysis and systemic risk evaluation, one often needs to compute or estimate the expectation of a random variable given that another random variable is equal to its quantile at some pre-specified probability level. A primary example of such an application is the Euler capital allocation formula for the quantile (often called the Value-at-Risk), which is of crucial importance in financial risk management. It is well known that classic nonparametric estimation for the above quantile allocation problem has a slower rate of convergence than the standard rate. In this paper, we propose an alternative approach to the quantile allocation problem via adjusting the probability level in connection with an expected shortfall. The asymptotic distribution of the proposed nonparametric estimator of the new capital allocation is derived for dependent data under the setup of a mixing sequence. In order to assess the performance of the proposed nonparametric estimator, AR-GARCH models are proposed to fit each risk variable and further, a bootstrap method based on residuals is employed to quantify the estimation uncertainty. A simulation study is conducted to examine the finite sample performance of the proposed inference. Finally, the proposed methodology of quantile capital allocation is illustrated for a financial data set.

Key-words: Bootstrap; Capital allocation; Expected Shortfall; Nonparametric estimation; Sensitivity analysis; Value-at-Risk.

*A few typos corrected in May 2019 and some corrections made in August 2022.

1 Introduction

1.1 VaR, ES, and capital allocation

In the banking regulatory frameworks of Basel II and III, as well as the insurance regulatory regimes such as Solvency II and Swiss Solvency Test (see BCBS (2016), Sandström (2010), EIOPA (2011) and IAIS (2014)), an institution is required to hold a certain capital according to a pre-specified regulatory risk measure. These regulatory environments set the capital via two standard risk measures: *Value-at-Risk (VaR)* and *Expected Shortfall (ES)*. For a random variable Y representing the future loss of a financial institution, the VaR at probability level $p \in (0,1)$ is defined as

$$\mathrm{VaR}_p(Y) := q_p(Y) := \inf\big\{x \in \mathbb{R} : \mathbb{P}(Y \leqslant x) \geqslant p\big\}.$$

Further, the ES at probability level $p \in (0,1)$ is defined as

$$\mathrm{ES}_p(Y) := c_p(Y) := \frac{1}{1-p}\int_p^1 \mathrm{VaR}_q(Y)\mathrm{d}q.$$

Moreover, if Y has a finite mean, let $\mathrm{ES}_0(Y) = \mathbb{E}[Y]$.

Risk measures such as VaR and ES are not only used externally for calculating regulatory capital, but also internally for risk management and performance measurement. In the internal use of risk measures, a crucial problem is the allocation of total capital to individual business lines in order to assess the performance of each business line (*performance analysis*), for instance via their *Return on Risk-adjusted Capital (RORAC)*.

Suppose that a financial institution has d different lines of business, each with an individual risk X_j, $j = 1,\dots,d$. Let S be the total risk of this firm, that is, $S := X_1 + \cdots + X_d$. The allocation of the total regulatory capital, set via either $\mathrm{ES}_p(S)$ or $\mathrm{VaR}_p(S)$, is typically carried out in practice via the *Euler capital allocation rule* (see Section 8.5 of McNeil et al. (2015)). Roughly speaking, the Euler allocation rule calculates how much extra capital would be needed if one business line increases its size by a small portion, or in other words, how much extra capital it consumes to generate one unit of its return. As such, it is shown to be the only RORAC compatible capital allocation rule (see Section 8.5.3 of McNeil et al. (2015)).

If the total regulatory capital for the aggregate risk S is calculated via $\mathrm{ES}_p(S)$, then the Euler allocation rule for the risk vector $(X_1,\dots,X_d)$ is given by

$$C_p(X_i|S) := \mathbb{E}[X_i|S \geqslant q_p(S)], \quad i = 1,\dots,d.$$

This allocation satisfies that $\mathrm{ES}_p(S) = \sum_{i=1}^d C_p(X_i|S)$, which is a natural and required condition for any capital allocation rules. On the other hand, if the total regulatory capital is calculated via $\mathrm{VaR}_p(S)$, then the Euler allocation rule for $(X_1, \ldots, X_d)$ is given by

$$Q_p(X_i|S) := \mathbb{E}[X_i|S = q_p(S)], \quad i = 1, \ldots, d.$$

Once again, we have $\mathrm{VaR}_p(S) = \sum_{i=1}^d Q_p(X_i|S)$.

A crucial observation is that, for a classic nonparametric estimator of $Q_p(X_i|S)$, the convergence rate is slower than the standard convergence rate $n^{-1/2}$ of a nonparametric estimator for $C_p(X_i|S)$, where n is the sample size. This is due to the fact that $Q_p(X_i|S)$ is a conditional expectation on a slice $\{S = q_p(S)\}$ of the probability space, as opposed to $C_p(X_i|S)$, which is a conditional expectation on $\{S \geqslant q_p(S)\}$, an event of non-zero probability. Therefore, a practical question is whether one could find an allocation $(C_1, \cdots, C_d)$ such that $\sum_{i=1}^d C_i = \mathrm{VaR}_p(S)$ and each of $C_1, \ldots, C_d$ could be estimated nonparametrically at the standard rate of convergence $n^{-1/2}$. Ideally, this new allocation should also be close to $Q_p(X_i|S)$, so that it roughly gives the Euler allocation.

1.2 Sensitivity analysis

Another major interpretation of $C_p(X_i|S)$ and $Q_p(X_i|S)$ is in the context of *sensitivity analysis*, which refers to the evaluation of the impact of some model inputs over the model outputs. Reasonable quantifications of the uncertainty with the input variables and model parameters represent standard ways to understand how sensible the model predictions are and the limitations of a model as part of the model validation. These are of particular interest when the performance measures are associated with the model outputs, which assist with achieving effectiveness, one of the primary aims of any business model. Parameter sensitivity analysis appears to be quite popular amongst academics and practitioners as it is easier to interpret and communicate the outcomes of such analysis. If a parameter is estimated or an expert opinion-based evaluation is in place, then sensitivity analysis illustrates sufficient information about acceptable modeling errors for such parameters. On the other hand, if a parameter is controllable by the decision-maker, then sensitivity analysis provides valuable information on how likely the outputs meet some desired standards, i.e., achieve effectiveness.

A performance measurement is usually chosen to match the common practice across a specific sector and quite often, it is due to external pressures, i.e., government regulations or independent rating agencies. One of the most common performance measurements is based on quantiles and therefore, quantile sensitivity becomes a topical area of research.

Let $h(\mathbf{X},\boldsymbol{\theta})$ be a random outcome that depends on an observable random input $\mathbf{X}$ and a given parameter vector $\boldsymbol{\theta}$, where $\mathbf{X} := (X_1, \cdots, X_d) \in \mathbb{R}^d$ and $\boldsymbol{\theta} := (\theta_1, \cdots, \theta_k) \in \mathbb{R}^k$. The sensitivity of $q_p(h(\mathbf{X},\boldsymbol{\theta}))$ and that of $c_p(h(\mathbf{X},\boldsymbol{\theta}))$ with respect to each component of the parameter $\boldsymbol{\theta}$ are defined, respectively, as

$$q_p^{(i)}(\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}\theta_i} q_p(h(\mathbf{X},\boldsymbol{\theta})) \quad \text{and} \quad c_p^{(i)}(\boldsymbol{\theta}) = \frac{\mathrm{d}}{\mathrm{d}\theta_i} c_p(h(\mathbf{X},\boldsymbol{\theta})), \quad i = 1, \dots, k.$$

Under sufficient continuity and differentiability assumptions on h and $\mathbf{X}$, one has

$$q_p^{(i)}(\boldsymbol{\theta}) = \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}\theta_i} h(\mathbf{X},\boldsymbol{\theta}) \middle| h(\mathbf{X},\boldsymbol{\theta}) = q_p(h(\mathbf{X},\boldsymbol{\theta}))\right], \tag{1}$$

and

$$c_p^{(i)}(\boldsymbol{\theta}) = \mathbb{E}\left[\frac{\mathrm{d}}{\mathrm{d}\theta_i} h(\mathbf{X},\boldsymbol{\theta}) \middle| h(\mathbf{X},\boldsymbol{\theta}) \geqslant q_p(h(\mathbf{X},\boldsymbol{\theta}))\right]. \tag{2}$$

See Theorem 2 of Hong (2009) and Theorem 3.1 of Hong and Liu (2009) for details.

By taking $Y_i = \frac{\mathrm{d}}{\mathrm{d}\theta_i} h(\mathbf{X},\boldsymbol{\theta})$ and $Y = h(\mathbf{X},\boldsymbol{\theta})$, (1) and (2) read as $q_p^{(i)}(\boldsymbol{\theta}) = Q_p(Y_i|Y)$ and $c_p^{(i)}(\boldsymbol{\theta}) = C_p(Y_i|Y)$, the key quantities studied in this paper. Moreover, if $d = k$ and $h(\mathbf{X},\boldsymbol{\theta}) = \sum_{i=1}^d \theta_i X_i$, then $q_p^{(i)}(\mathbf{1}) = Q_p(X_i|S)$ and $c_p^{(i)}(\mathbf{1}) = C_p(X_i|S)$, where $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^d$ and $S := h(\mathbf{X}, \mathbf{1})$. Thus, the Euler allocation rules for VaR and ES are given as special cases of (1) and (2).

As sensitivity analysis and capital allocation problems both boil down to similar mathematical formulations, i.e., $Q_p(X|Y)$ and $C_p(X|Y)$ for suitably chosen random variables X and Y, we simply focus on the interpretation of the capital allocation in the rest of the paper.

1.3 Related literature

Euler allocation rules are extensively studied in Denault (2001) and Kalkbrener (2005); we refer to Section 8.5 of McNeil et al. (2015) for a general treatment. Indeed, the allocation rule proposed in this paper is briefly mentioned in Kalkbrener (2005) without either formal theoretical development or statistical analysis; it is our intention to give a theoretical treatment of this idea and to provide an efficient inference procedure.

If $(X_1, \ldots, X_d)$ is modeled by a parametric distribution family and independent observations are available, Glasserman (2005) and Glasserman and Li (2005) employed importance sampling technique to estimate $C_p(X_i|S)$ and $Q_p(X_i|S)$ with applications to portfolio credit risk, but without addressing the uncertainty issue in fitting the parametric family. Nonparametric estimation for $C_p(X_i|S)$ has been studied in Scaillet (2004, 2005). For estimating $Q_p(X_i|S)$ nonparametrically, we refer to Gourieroux et al. (2000), Liu and Hong (2009), Hong (2009), where, as mentioned previously, these methods suffer from a slower convergence rate than the standard convergence rate $n^{-1/2}$. Under some special cases, $Q_p(X_i|S)$ could be written as a ratio of two quantities and each quantity is estimated nonparametrically at the standard rate of convergence, i.e., $Q_p(X_i|S)$ is estimated at the standard rate of convergence; see Fu et al. (2009) and Jiang and Fu (2015).

Sensitivity analysis regarding input variables requires the concept of directional derivatives and it has been applied to capital modeling problems in Tasche (2009). A much wider scope within risk analysis is discussed in Tsanakas and Millossovich (2016). Quantile sensitivity has appeared in the literature in various forms. For example, the linear risk portfolio is discussed in Gourieroux et al. (2000); kernel estimation and importance sampling techniques are combined by Tasche (2009) for a portfolio credit risk model; other nonparametric estimations are employed via infinitesimal perturbation analysis (see Hong (2009)); importance sampling techniques for a parametric portfolio credit risk model are detailed in Glasserman (2005) and Glasserman and Li (2005) for both quantiles and average tail risk. Finally, sensitivity analysis for the VaR-based and ES-based measures of performance are investigated by using the special features of Fourier Transform Monte Carlo in Siller (2013) within the portfolio credit risk setting.

The main reason for us to tackle the question of capital allocation is simply that we would like both VaR-based allocation and ES-based allocation to be estimated nonparametrically with the standard rate of convergence. This paper does not intend to discuss which risk measure is more appropriate, although we are well aware of debates/discussions on these two risk measures, briefly summarized below.

In the recent Basel documents (see e.g., BCBS (2016)), a move from VaR to ES as the standard risk measure for market risk has been initiated and confirmed. But VaR and ES co-exist in insurance regulation in different parts of the world (e.g., Solvency II vs the Swiss

Solvency Test). This situation has stimulated extensive academic and industry debates on the desirable properties of VaR and ES; see Embrechts et al. (2014) and Emmer et al. (2015) for comprehensive discussions. As a result, various quantitative concepts enter into the discussion and model uncertainty becomes the central focus. For instance, robustness of risk measures are addressed in Cont et al. (2010), Kou et al. (2013), Krätschmer et al. (2014) and Embrechts et al. (2015); for recent advances on elicitability and forecasting, see Ziegel (2016), Fissler and Ziegel (2016), and Kou and Peng (2016); for development on model uncertainty in risk aggregation, see Embrechts et al. (2013), Bernard and Vanduffel (2015) and Cai et al. (2018). In summary, ES and VaR both have various advantages and disadvantages. Finally, the Euler allocation formula $C_p(X_i|S)$ finds mathematical similarity to the systemic risk measure *CoES* studied in Adrian and Brunnermeier (2016) and Acharya et. al. (2012), where S represent the overall economy; see also Acharya (2009), Chen et al. (2013) and Rogers and Veraart (2013) for more recent results on measures of the systemic risk.

2 An ES-based approach to quantile capital allocation

We work with an atomless probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and let L^q be the set of random variables with finite q-th moment. For $p \in (0,1)$ and random variables X and Y, the key quantities $Q_p(X|Y)$ and $C_p(X|Y)$ are defined before. Recall that $q_p(X) = Q_p(X|X)$ and $c_p(X) = C_p(X|X)$.

For $Y \in L^1$ and $p \in (0,1)$, let $p^* \in (0,1)$ be given by the following equation

$$p^* := \inf\{t \in [0,1] : c_t(Y) \geqslant q_p(Y)\}. \tag{3}$$

Note that in the above notation, we omit the dependence of p^* on Y and p, which is clear from the context. First, let us verify the existence of p^*.

Proposition 2.1. *For $Y \in L^1$ and $p \in (0,1)$, assume $\mathbb{E}[Y] \leqslant q_p(Y)$. Then $p^* \in [0,p]$. Moreover, if Y is continuously distributed, then $c_{p^*}(Y) = q_p(Y)$.*

All proofs are relegated in Section 7. Note that in insurance and financial applications, p is typically close to 1 (e.g., $p = 0.99, 0.995, 0.999, \ldots$), which makes the condition from Proposition 2.1, $\mathbb{E}[Y] \leqslant q_p(Y)$, a realistic assumption.

The primary application of the above proposition is the capital allocation outlined in Section 1.1, where, for a portfolio of risks $\mathbf{X} := (X_1, \ldots, X_d)$, we take $X = X_i$ and $Y = X_1 + \cdots + X_d$. In the sequel, we shall refer to $C_{p^*}(X|Y)$ as the *ES-based allocation for the p-quantile*. If Y is continuously distributed, an immediate feature of the allocation $(C_{p^*}(X_1|Y), \ldots, C_{p^*}(X_d|Y))$ is

$$\sum_{i=1}^{d} C_{p^*}(X_i|Y) = C_{p^*}\left(\sum_{i=1}^{d} X_i \middle| Y\right) = c_{p^*}(Y) = q_p(Y). \tag{4}$$

Therefore, $(C_{p^*}(X_1|Y), \ldots, C_{p^*}(X_d|Y))$ indeed gives an allocation of the VaR-based total risk capital $q_p(Y)$.

Statistical inference for the new allocation $C_{p^*}(X_i|Y)$ is given in Section 3. Before approaching the estimation for $C_{p^*}(X_i|Y)$, we first discuss some useful properties. As mentioned in Section 1.1, we would like the new allocation $C_{p^*}(X_i|Y)$ to be close to $Q_p(X_i|Y)$. Indeed, for elliptically distributed (X, Y), we have the equality $C_{p^*}(X|Y) = Q_p(X|Y)$, as illustrated by the following example. The elliptical family includes the multivariate normal and t-distributions. Due to its ease of implementation, this family of distributions is commonly used in risk management, see McNeil et al. (2015), and credit risk analysis, see Glasserman (2005) and Glasserman and Li (2005). Further results on the relation between $C_{p^*}(X|Y)$ and $Q_p(X|Y)$ are presented in Section 5.

Example 2.1 (Elliptical distributions)**.** For a vector $\boldsymbol{\mu} = (\mu_1, \mu_2)$ and a positive semi-definite matrix $\Sigma = (\sigma_{ij}) \in \mathbb{R}^{2\times 2}$, let (X, Y) follow from an elliptical distribution $\mathcal{E}_2(\boldsymbol{\mu}, \Sigma, \phi)$, where ϕ is the generator of $\mathcal{E}_2(\boldsymbol{\mu}, \Sigma, \phi)$ (see Cambanis et al. (1981)). Assume $\mathbb{E}[|Y|] < \infty$. By Corollary 5 of Cambanis et al. (1981), for $p \in (0, 1)$,

$$Q_p(X|Y) = \mathbb{E}[X|Y = q_p(Y)] = \mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(q_p(Y) - \mu_2).$$

Analogously, (see, e.g., Theorems 2-3 of Landsman and Valdez (2003)), for $p \in (0, 1)$,

$$C_p(X|Y) = \mathbb{E}[X|Y > q_p(Y)] = \mu_1 + \frac{\sigma_{12}}{\sigma_{22}}(c_p(Y) - \mu_2).$$

If $p \in [1/2, 1)$, then $p^* \in [0, 1)$. Therefore, we have $C_{p^*}(X|Y) = Q_p(X|Y)$ for $p \geqslant 1/2$. In the capital allocation setting, if $(X_1, \ldots, X_d) \sim \mathcal{E}_d(\boldsymbol{\mu}, \Sigma, \phi)$ and $Y = X_1 + \cdots + X_d$ is non-degenerate and integrable, then (X_i, Y) is elliptically distributed for all $i = 1, \ldots, d$, and in turn, for $p \geqslant 1/2$ we have that

$$(C_{p^*}(X_1|Y),\dots,C_{p^*}(X_d|Y)) = (Q_p(X_1|Y),\dots,Q_p(X_d|Y)),$$

which can also be derived from Corollary 8.43 of McNeil et al. (2015).

Next, we build up a simple example to show that $Q_p(X|Y) \neq C_{p^*}(X|Y)$ is possible, although the two values are typically quite close. This example is recalled again later when we discuss the asymptotic equivalence of $Q_p(X|Y)$ and $C_{p^*}(X|Y)$.

Example 2.2 (Pareto risks)**.** For $\alpha > 1$, take $X \sim \text{Pareto}(2\alpha)$, i.e., $\mathbb{P}(X > x) = x^{-2\alpha}$, $x \geqslant 1$. Let $Y = X^2$ and therefore, $Y \sim \text{Pareto}(\alpha)$. One may calculate, for $t > 1$,

$$y(t) = \mathbb{E}[Y|Y > t] = \frac{\int_t^\infty \alpha y^{-\alpha}\mathrm{d}y}{t^{-\alpha}} = t\frac{\alpha}{\alpha-1},$$

and hence, $\mathbb{E}[X|Y = y(t)] = \sqrt{t\frac{\alpha}{\alpha-1}}$. On the other hand,

$$\mathbb{E}[X|Y > t] = \mathbb{E}[X|X > \sqrt{t}] = \sqrt{t}\frac{2\alpha}{2\alpha-1}.$$

Therefore,

$$\frac{\mathbb{E}[X|Y = y(t)]}{\mathbb{E}[X|Y > t]} = \frac{\sqrt{\frac{\alpha}{\alpha-1}}}{\frac{2\alpha}{2\alpha-1}} = \frac{2\alpha-1}{2\sqrt{\alpha(\alpha-1)}} \neq 1. \tag{5}$$

For some p with $q_p(Y) > \mathbb{E}[Y]$, we can replace t by $q_{p^*}(Y)$. Then $y(t) = c_{p^*}(Y) = q_p(Y)$, and (5) gives that $Q_p(X|Y) \neq C_{p^*}(X|Y)$. Moreover we note that the value in (5) is quite close to 1 if $\alpha \geqslant 2$ (e.g., it is 1.06066 for $\alpha = 2$ and 1.00416 for $\alpha = 6$), which motivates our study on the asymptotic equivalence in Section 5.

Remark 2.1. It might be worth noting that the applicability of the proposed ES-based allocation for a quantile does not rely on the fact that the total capital is calculated via $q_p(Y)$. Indeed, if the total capital is C, then one may always write $p^* = \inf\{t \in [0,1] : c_t(Y) \geqslant C\}$ and calculate $\big(C_{p*}(X_1|Y),\dots,C_{p^*}(X_d|Y)\big)$ to arrive at an allocation rule satisfying $C_{p*}(X_1|Y) + \cdots + C_{p^*}(X_d|Y) = C$. Certainly, the nice interpretation of the Euler allocation (e.g., RORAC compatibility) cannot be discussed in this case, because we are not given the formula of the total capital. Note that in practice, there might be many adjustments to the total capital allocation (e.g., stress scenarios, moving averages and liquidity adjustments; see BCBS (2016)). Our method could be easily adapted to such cases, as the total capital C is always known in practical capital allocation problems.

3 Nonparametric inferences

Throughout this section, for the purpose of illustration, we again adopt a capital allocation setting. Let $p \in (0,1)$ and $(X_1, \ldots, X_d)$ be a random vector such that $Y := X_1 + \cdots + X_d$ is continuously distributed. For simplicity, write $C_j = C_{p^*}(X_j|Y)$ for all $j = 1, \ldots, d$, which is our target to estimate. Let $\{\mathbf{X}_t = (X_{1,t}, \cdots, X_{d,t})\}_{t=1}^n$ be a stationary sequence from the distribution of $(X_1, \ldots, X_d)$ and write $Y_t := X_{1,t} + \cdots + X_{d,t}$ for all $t = 1, \ldots, n$.

A nonparametric estimator $\hat{p}^*$ for p^* in (3) is obtained by solving the following equation for $t \in [0,1)$,

$$\frac{1}{1-t}\int_t^1 G_n^-(s)\,ds = G_n^-(p),$$

where $G_n(s) := \frac{1}{n}\sum_{t=1}^n I(Y_t \leqslant s)$ and G_n^- denotes the generalized inverse function of G_n. It is not difficult to check that the above equation has a unique solution when $\int_0^1 G_n^-(s)\,ds < G_n^-(p)$. Therefore, a nonparametric estimator for $C_j = C_{p^*}(X_j|Y)$ is given by

$$\hat{C}_j := \frac{1}{1-\hat{p}^*}\frac{1}{n}\sum_{t=1}^n X_{j,t} I(Y_t > G_n^-(\hat{p}^*)). \tag{6}$$

To derive the asymptotic distribution of the above estimator, $\left\{\mathbf{X}_t := (X_{1,t}, \cdots, X_{d,t})\right\}_{t=-\infty}^{\infty}$ is assumed to be a strictly stationary α-mixing sequence, i.e.,

$$\alpha_{\mathbf{X}}(k) := \sup\left\{\left|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)\right| : A \in \mathcal{F}_{-\infty}^i, B \in \mathcal{F}_{i+k}^{\infty}, -\infty < i < \infty\right\} \to 0$$

as $k \to \infty$, where $\mathcal{F}_a^b$ denotes the σ-field generated by $\{\mathbf{X}_t\}_{t=a}^b$.

Let F_j and G denote the distribution function of $X_{j,t}$ and Y_t, respectively. Further, define

$$C(x,y;j) := \mathbb{P}\big(F_j(X_{j,t}) \leqslant x, 1-G(Y_t) \leqslant y\big),\ C_n(x,y;j) := \frac{1}{n}\sum_{t=1}^n I\big(F_j\big(X_{j,t}\big) \leqslant x, 1-G(Y_t) \leqslant y\big),$$

$$C(x_1, \cdots, x_{d-1}, y) := \mathbb{P}\big(F_1(X_{1,t}) \leqslant x_1, \cdots, F_{d-1}(X_{d-1,t}) \leqslant x_{d-1}, 1-G(Y_t) \leqslant y\big),$$

and

$$C_n(x_1, \cdots, x_{d-1}, y) := \frac{1}{n}\sum_{t=1}^n I\big(F_1(X_{1,t}) \leqslant x_1, \cdots, F_{d-1}(X_{d-1,t}) \leqslant x_{d-1}, 1-G(Y_t) \leqslant y\big).$$

A key technique in deriving our theoretical results is the following weighted approximation of empirical copula process. Under the regularity condition C1) given in the next paragraph, it follows from Proposition 4.4 of Berghaus et al. (2017) that

$$\sup_{0\leqslant x_1,\cdots,x_{d-1},y\leqslant 1} |\sqrt{n}\big(C_n(x_1,\cdots,x_{d-1},y)-C(x_1,\cdots,x_{d-1},y)\big)-W(x_1,\cdots,x_{d-1},y)|=o_p(1) \tag{7}$$

and

$$\sup_{1/n\leqslant u_1,u_2\leqslant 1-1/n} \frac{\big|\sqrt{n}\big(C_n(u_1,u_2;j)-C(u_1,u_2;j)\big)-W(u_1,u_2;j)\big|}{\{\min(u_1,u_2,1-u_1,1-u_2)\}^{\delta}}=o_p(1) \tag{8}$$

for any $\delta\in(0,1/2)$, where $W(u_1,u_2;j)$ is $W(x_1,\cdots,x_{d-1},y)$ with $x_j=u_1, y=u_2$ and the rest being one, $W_d(y)=W(1,\cdots,1,y)$, and $W(x_1,\cdots,x_{d-1},y)$ is a multivariate centered Gaussian process with covariance

$$\begin{aligned}
&\operatorname{Cov}\big(W(x_1,\cdots,x_{d-1},y),W(\tilde{x}_1,\cdots,\tilde{x}_{d-1},\tilde{y})\big)\\
=\ &\textstyle\sum_{i=-\infty}^{\infty}\operatorname{Cov}\Big(I\big(F_1(X_{1,1})\leqslant x_1,\cdots,F_{d-1}(X_{d-1,1})\leqslant x_{d-1},1-G(Y_1)\leqslant y\big),\\
&I\big(F_1(X_{1,1+i})\leqslant \tilde{x}_1,\cdots,F_{d-1}(X_{d-1,1+i})\leqslant \tilde{x}_{d-1},1-G(Y_{1+i})\leqslant \tilde{y}\big)\Big).
\end{aligned}$$

Throughout we assume the following regularity conditions:

C1) $\alpha_{\mathbf{X}}(k)=O(a^k)$ for some $a\in(0,1)$;

C2) $\mathbb{E}[Y]<q_p(Y)$;

C3) G is differentiable, $G'(x)$ is continuous at $G^-(p)$ and $C^{(2)}(x,y;j)=\frac{\partial}{\partial y}C(x,y;j)$ is continuous at $1-p^*$ for all $x\in[0,1]$;

C4) There exist $\eta_0>0$ and $\delta_0>0$ such that $\mathbb{E}\left[|Y_t|^{2+\delta_0}I\big(Y_t>G^-(p^*)-\eta_0\big)\right]<\infty$, $\mathbb{E}\left[|X_{j,t}|^{2+\delta_0}I\big(S_t>G^-(p^*)-\eta_0\big)\right]<\infty$ and $\sup_{|s-p^*|\leqslant\eta_0}\int_0^1 C^{(2)}(x,1-s;j)\mathrm{d}F_j^-(x)<\infty$ for $j=1,\cdots,d-1$.

Theorem 3.1. *Under conditions C1)–C4), for a given $p\in(0,1)$, we have*

$$\sqrt{n}(\hat{p}^*-p^*)=-\frac{1}{G^-(p)-G^-(p^*)}\int_{p^*}^1\frac{W_d(1-s)}{G'(G^-(s))}\,ds+\frac{(1-p^*)W_d(1-p)}{G'\big(G^-(p)\big)\big(G^-(p)-G^-(p^*)\big)}+o_p(1), \tag{9}$$

$$\begin{aligned}
&\sqrt{n}(\hat{C}_j-C_j)\\
=\ &\Big(-\tfrac{1}{G^-(p)-G^-(p^*)}\textstyle\int_{p^*}^1\frac{W_d(1-s)}{G'(G^-(s))}\mathrm{d}s+\frac{(1-p^*)W_d(1-p)}{G'(G^-(p))(G^-(p)-G^-(p^*))}\Big)\times\\
&\Big(\tfrac{1}{(1-p^*)^2}\textstyle\int_0^1 F_j^-(x)\mathrm{d}C(x,1-p^*;j)+\frac{1}{1-p^*}\int_0^1 C^{(2)}(x,1-p^*;j)\mathrm{d}F_j^-(x)\Big)\\
&+\tfrac{1}{1-p^*}\Big(-\textstyle\int_0^1 W(x,1-p^*;j)\mathrm{d}F_j^-(x)+W_d(1-p^*)\int_0^1 C^{(2)}(x,1-p^*;j)\mathrm{d}F_j^-(x)\Big)\\
&+o_p(1)\quad \textit{for}\quad j=1,\cdots,d-1,
\end{aligned}\tag{10}$$

and

$$\begin{aligned} & \sqrt{n}\sum_{j=1}^d(\hat{C}_j-C_j) \\ = & \left(-\frac{1}{G^-(p)-G^-(p^*)}\int_{p^*}^1\frac{W_d(1-s)}{G'(G^-(s))}\mathrm{d}s+\frac{(1-p^*)W_d(1-p)}{G'(G^-(p))(G^-(p)-G^-(p^*))}\right)\times \\ & \left(\frac{1}{(1-p^*)^2}\int_{p^*}^1 G^-(y)\,dy+\frac{G^-(p^*)}{1-p^*}\right)-\frac{1}{1-p^*}\int_{p^*}^1 W_d(1-x)\mathrm{d}G^-(x)+o_p(1). \end{aligned} \tag{11}$$

Erratum. The convergence statements in the above theorem should be in distribution, because the approximation (7) does not hold in probability.[1]

Remark 3.1. Similarly to Fermanian and Scaillet (2003), a smoothed version of $\hat{C}_j$ can be employed. We do not study the smoothed estimation here as it is known that a smoothed distribution estimation only improves the empirical distribution in the sense of the second order error (see Cheng and Peng (2002)).

Remark 3.2. The asymptotic distribution of $\sqrt{n}(\hat{C}_d-C_d)$ follows from (10) and (11). Further we can estimate the percentage $C_j/\sum_{i=1}^d C_i = C_j/q_p(Y)$ by $\hat{C}_j/\sum_{i=1}^d\hat{C}_i$ and its asymptotic distribution readily follows from (10) and (11) by noting that

$$\sqrt{n}\left(\frac{\hat{C}_j}{\sum_{i=1}^d\hat{C}_i}-\frac{C_j}{\sum_{i=1}^d C_i}\right)=\frac{\sqrt{n}(\hat{C}_j-C_j)}{\sum_{i=1}^d C_i}-\frac{C_j}{(\sum_{i=1}^d C_i)^2}\sqrt{n}\sum_{i=1}^d(\hat{C}_i-C_i)+o_p(1).$$

Obviously, Theorem 3.1 shows that C_j can be estimated nonparametrically at the standard rate of convergence $n^{-1/2}$ and the asymptotic distribution of the proposed nonparametric estimator $\hat{C}_j$ is normal, but with a complicated asymptotic variance. To construct a confidence interval for C_j, one may simply employ the blockwise bootstrap method for a mixing sequence. However, because we are dealing with quantiles, nonparametric inferences become very inefficient due to a small number of blocks. A simple remedy is to model the dependence of each $\{X_{i,t}\}_{t=1}^n$ by a time series model and then to employ a bootstrap method based on residuals to construct confident intervals. Specifically, we assume that each $\{X_{i,t}\}_{t=1}^n$ follows an AR-GARCH model (for details, see Chen and Fan (2006)), i.e.,

$$X_{j,t}=\mu_j+\sum_{i=1}^{P_j}a_{j,i}X_{j,t-i}+e_{j,t},e_{j,t}=h_{j,t}^{1/2}\eta_{j,t},\quad h_{j,t}=w_j+\sum_{i=1}^{q_j}\alpha_{j,i}e_{j,t-i}^2+\sum_{k=1}^{p_j}\beta_{j,k}h_{j,t-k} \tag{12}$$

for all $j=1,\ldots,d$, where $\left\{\boldsymbol{\eta}_t := (\eta_{1,t},\ldots,\eta_{d,t})\right\}$ is a sequence of independent and identically distributed random vectors with zero means and variances of one.

[1]Added August 2022.

In order to estimate the asymptotic variance of $\hat{C}_j$ via a bootstrap method, one has to estimate $e_{j,t}$, which requires estimation of the unknown parameters in (12). An obvious estimator is the so-called quasi maximum likelihood estimator; its asymptotic normality is available in Francq and Zakoïan (2004), which requires finite fourth moments of both $e_{j,t}$ and $\eta_{j,t}$. However, it is quite often in practice that $\sum_{i=1}^{q_j} \alpha_{j,i} + \sum_{k=1}^{p_j} \beta_{j,k}$ are close to one, and thus, assuming $\mathbb{E}e_{j,t}^4 < \infty$ may be questionable. Here, we propose to employ the self-weighted estimator from Ling (2007) with the following weights

$$\delta_{j,t} = \left\{ \max\left(1, \frac{1}{d_j} \sum_{k=1}^{t-1} \frac{|X_{t-k}| I\big(|X_{t-k}| > d_j\big)}{k^9}\right)\right\}^{-4}$$

to estimate

$$\boldsymbol{\theta}_j = (\mu_j, a_{j,1}, \ldots, a_{j,P_j}, \alpha_{j,1}, \ldots, \alpha_{j,q_j}, \beta_{j,1}, \ldots, \beta_{j,p_j}),$$

say $\hat{\boldsymbol{\theta}}_j$, where the asymptotic normality only requires $\mathbb{E}[|e_{j,t}|] < 1$ and $\mathbb{E}[\eta_{j,t}^2] < \infty$. Here, d_j is chosen as the 90% sample quantile of $\{X_{j,t}\}_{t=1}^n$, as suggested in Zhu and Ling (2011).

After obtaining $\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_d$, we get our estimators for $\boldsymbol{\eta}_t$, $t = 1, \ldots, n$, say $\hat{\boldsymbol{\eta}}_t$. Therefore, we resample from $\{\hat{\boldsymbol{\eta}}_t\}_{t=1}^n$ with sample size n and then refit models (12) to obtain bootstrap samples $X_{j,t}^*$ for $j = 1, \ldots, d$ and $t = 1, \ldots, n$. Based on this bootstrap sample, one may calculate the bootstrapped estimator of $\hat{C}_j$. By repeating this procedure, a bootstrap confidence interval for C_j via $\hat{C}_j$ could be constructed.

Remark 3.3. If $\boldsymbol{X}_t$ can not be modeled by (12) directly such as stock prices, but one can write $\boldsymbol{X}_t = k(\boldsymbol{X}_{t-1}, \boldsymbol{Z}_t)$ with a known function k and $\boldsymbol{Z}_t$ being modeled by (12), then a so-called one-step-ahead conditional capital allocation $C_{p^*}(X_{j,n+1}|Y_{n+1})$ given $\boldsymbol{X}_1, \cdots, \boldsymbol{X}_n$ become more meaningful due to nonstationarity of $\{\boldsymbol{X}_t\}$. In this case, the above proposed nonparametric inference is still valid.

4 Numerical results

4.1 Real data analysis

The proposed nonparametric estimator is now investigated for financial data. The bootstrap method is applied to the 100 times log-returns of IBM stock price $(X_{1,t})$ and S&P 500 index $(X_{2,t})$ from December 1, 2005 to December 31, 2015. Hence, $Y_t := X_{1,t} + X_{2,t}$.

First, we compute the proposed nonparametric estimates $\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)$ and $\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)$ for levels $p = 0.95$ and 0.99. To examine the closeness of these new allocations to the VaR-based allocations $Q_p(X_{1,1}|Y_1)$ and $Q_p(X_{2,1}|Y_1)$, we estimate them nonparametrically by

$$\hat{Q}_p(X_{1,1}|Y_1) = \frac{\sum_{t=1}^n X_{1,t} k(\frac{Y_t - \hat{\theta}}{h})}{\sum_{t=1}^n k(\frac{Y_t - \hat{\theta}}{h})} \text{ and } \hat{Q}_p(X_{2,1}|Y_1) = \frac{\sum_{t=1}^n X_{2,t} k(\frac{Y_t - \hat{\theta}}{h})}{\sum_{t=1}^n k(\frac{Y_t - \hat{\theta}}{h})},$$

where $\hat{\theta}$ is the p-th sample quantile of Y_t's. We use $k(x) = \frac{3}{4}(1 - x^2)$ for $x \in (-1, 1)$ and $h = 0.5n^{-1/5}, n^{-1/5}, 1.5n^{-1/5}$, which have the same rate of convergence as the optimal bandwidth in terms of minimizing the mean squared error. We also estimate the percentages $\frac{C_{p^*}(X_{1,1}|Y_1)}{C_{p^*}(X_{1,1}|Y_1)+C_{p^*}(X_{2,1}|Y_1)}$ and $\frac{C_{p^*}(X_{2,1}|Y_1)}{C_{p^*}(X_{1,1}|Y_1)+C_{p^*}(X_{2,1}|Y_1)}$ by $\frac{\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)}{\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)+\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)}$ and $\frac{\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)}{\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)+\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)}$, respectively. Because the sum of these two estimators is one, it is easy to check that these two estimators have the same standard deviation. Hence we only report results for the first estimator.

To compute the standard deviations of the proposed estimators by the bootstrap method given after Theorem 3.1, which uses the self-weighted quasi maximum likelihood estimation in Ling (2007) to fit AR(1)-GARCH(1,1) models to our data. The estimates are

$$\begin{cases} \mu_1 = 0.0389, \quad a_{1,1} = -0.0264, \quad w_1 = 0.1201, \quad \alpha_{1,1} = 0.1187, \quad \beta_{1,1} = 0.8161, \\ \mu_2 = 0.0770, \quad a_{2,1} = -0.063, \quad w_2 = 0.0366, \quad \alpha_{2,1} = 0.1386, \quad \beta_{2,1} = 0.8393. \end{cases} \tag{13}$$

Based on the above estimates and drawing $1,000$ bootstrap samples, we compute the bootstrapped standard deviations for $\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)$, $\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)$, $\hat{Q}_p(X_{1,1}|Y_1)$ and $\hat{Q}_p(X_{2,1}|Y_1)$. We also compute the bootstrapped correlation coefficients between $\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)$ and $\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)$, and between $\hat{Q}_p(X_{1,1}|Y_1)$ and $\hat{Q}_p(X_{2,1}|Y_1)$. These results are reported in Table 1. We remark the bootstrapped standard deviations of $\hat{Q}_p(X_{1,1}|Y_1)$ and $\hat{Q}_p(X_{2,1}|Y_1)$ are incorrect as Hall (1990) showed that bootstrap method can not catch the asymptotic bias in kernel density estimation.

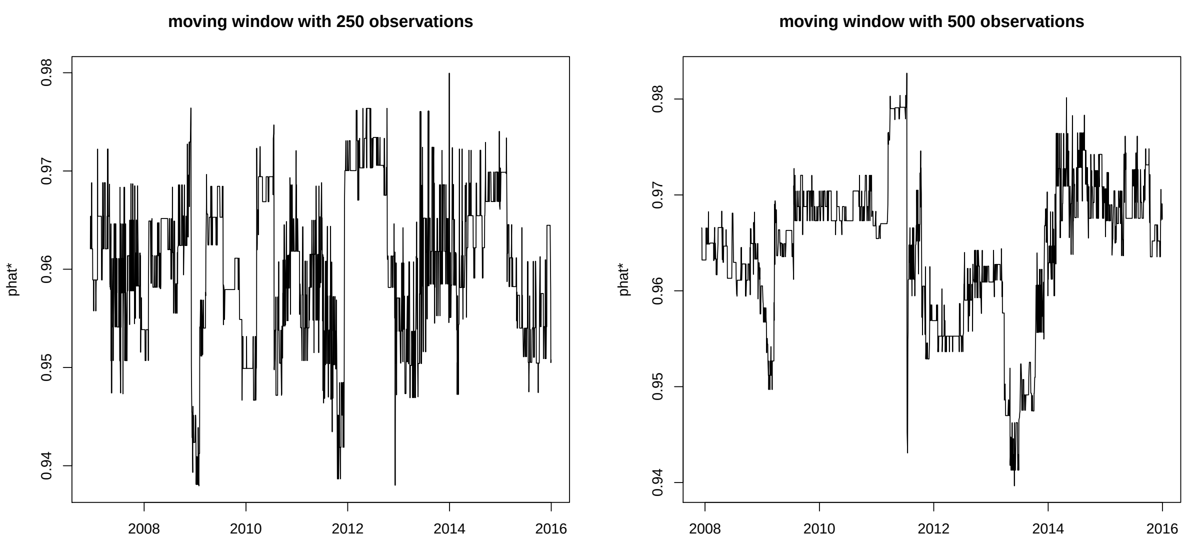
From Table 1, we observe that i) the new allocations are close to their corresponding VaR-based allocations; ii) the capital allocation estimate for S&P 500 index has a larger variance than that for the IBM stock price, which may be due to the fact that $\beta_{2,1} + \alpha_{2,1}$ is closer to one than $\beta_{1,1} + \alpha_{1,1}$; iii) the variance for each estimate increases as p becomes larger; iv) the variances for the VaR-based allocation estimates are quite sensitive to the choice of bandwidth, which may be explained by the inconsistency of the employed bootstrap method as argued in Hall (1990); v) the variances of the VaR-based allocation estimates are larger than those for the new allocation estimates when $p = 0.99$, which may be explained by the faster rate of convergence of the new allocation estimates.

Table 1: *Estimates, Bootstrapped Standard Deviations and Bootstrapped Correlation Coefficient.* Estimates and bootstrapped standard deviations are reported for $\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)$, $\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)$, $\hat{Q}_p(X_{1,1}|Y_1)$ and $\hat{Q}_p(X_{2,1}|Y_1)$ for levels $p = 0.95$ and 0.99. The bootstrapped correlation coefficients between $\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)$ and $\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)$ are -0.0891 for $p = 0.95$ and -0.2432 for $p = 0.99$. The bootstrapped correlation coefficients between $\hat{Q}_p(X_{1,1}|Y_1)$ and $\hat{Q}_p(X_{2,1}|Y_1)$ with $h = 0.5n^{-1/5}, n^{-1/5}, 1.5n^{-1/5}$ are, respectively, $-0.4493, -0.2008, 0.0427$ for $p = 0.95$ and $-0.5996, -0.6117, -0.4389$ for $p = 0.99$.

	$p = 0.95$		$p = 0.99$	
	Estimate	Bootstrapped SD	Estimate	Bootstrapped SD
p^*	0.8484	0.0084	0.9687	0.0035
$C_{p^*}(X_{1,1}\|Y_1)$	1.9441	0.1318	3.3539	0.3921
$C_{p^*}(X_{2,1}\|Y_1)$	1.7000	0.1817	3.3379	0.5869
$Q_p(X_{1,1}\|Y_1),\ h = 0.5n^{-1/5}$	1.8889	0.2000	3.4250	0.7938
$Q_p(X_{2,1}\|Y_1),\ h = 0.5n^{-1/5}$	1.7562	0.2097	3.2555	0.8989
$Q_p(X_{1,1}\|Y_1),\ h = n^{-1/5}$	1.9545	0.1632	3.3627	0.6044
$Q_p(X_{2,1}\|Y_1),\ h = n^{-1/5}$	1.6771	0.1789	3.3111	0.7781
$Q_p(X_{1,1}\|Y_1),\ h = 1.5n^{-1/5}$	2.0329	0.1395	3.3840	0.4862
$Q_p(X_{2,1}\|Y_1),\ h = 1.5n^{-1/5}$	1.5974	0.1643	3.2869	0.6773
$\frac{C_{p^*}(X_{1,1}\|Y_1)}{C_{p^*}(X_{1,1}\|Y_1)+C_{p^*}(X_{2,1}\|Y_1)}$	0.5335	0.0344	0.5012	0.0645

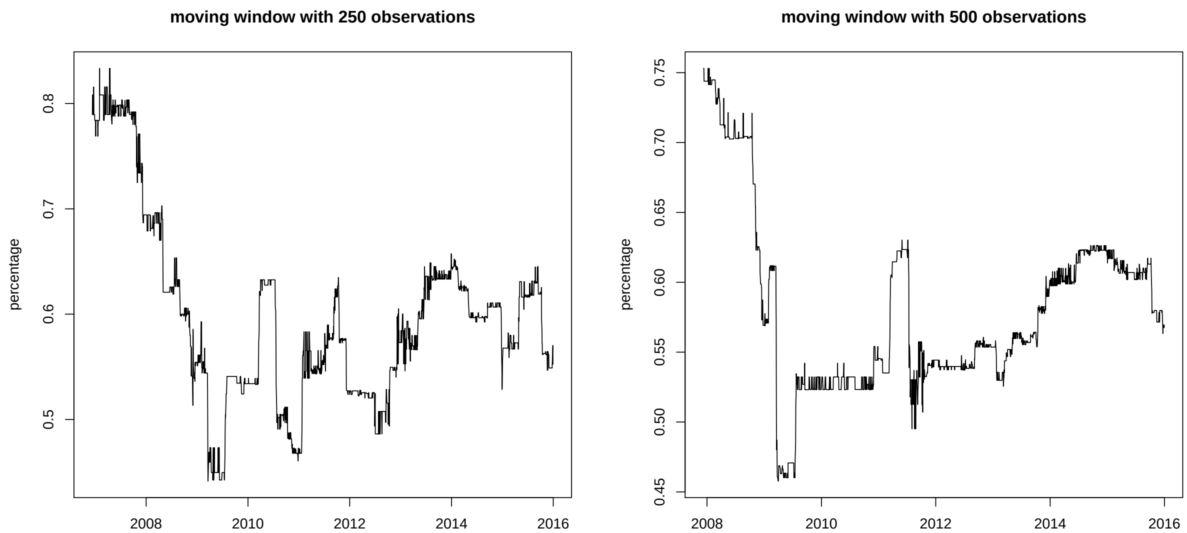
To distinguish different time periods, we use 10-day return data and a moving window of 250 or 500 observations to compute $\hat{p}^*$ and the percentage of capital allocated to the S&P 500 index with respect to the level $p = 0.99$. These estimates are reported in Figures 1-2. During a majority of time spots reported in Figures 1, the value of p^* is slightly less than 0.975, indicating two conclusions for this particular data set. First, the Basel Committee on Banking Supervision recently replaced $\mathrm{VaR}_{0.99}$ by $\mathrm{ES}_{0.975}$ as the risk measure for 10-day market risk (BCBS (2016)). Figure 1 shows that the transition from $\text{VaR}_{0.99}$ to $\text{ES}_{0.975}$ will slightly increase the capital requirement for market risk. Second, as $p^* \approx 0.974$ for normal risks, the plots in Figure 1 show that the 10-day return data have a heavier tail than normal distribution; see McNeil et al. (2015) for more general discussions on this phenomenon. In Figure 2, we observe that the percentage of the capital allocated to S&P takes the smallest value during the financial crisis. This may be explained by the fact that the IBM stock volatility were high during the period of 2007-2009, and at the same time the value of S&P decreased drastically (hence, a less portion of the portfolio is invested in S&P).

Figure 1: $\hat{p}^*$ *based on real data.* $\hat{p}^*$ is calculated for $p = 0.99$ based on ten days' returns with a moving window of observations 250 or 500.



Finally, in order to simulate data close to these two real data sets, we need to fit a parametric family to $\boldsymbol{\eta}_t$. Although we could use the R package 'sn' to fit a skewed bivariate t-distribution, the mean and variance of $\eta_{j,t}$ may not be zero and one, respectively, which are required by the models (12). Instead, we simply use the R function 'stdFit' to fit a t distribution to each marginal and use the R package 'copula' to fit a t copula to the copula of $\boldsymbol{\eta}_t$. This gives marginal distributions $t(4.9504)$ and $t(5.3462)$, and the t copula with $\rho = 0.6976$ and $\nu = 5.3355$. These parameters will be employed to generate samples to examine the finite sample performance of the proposed nonparametric estimator and its bootstrap method in the next subsection.

Figure 2: *Percentage estimate based on real data.* $\frac{\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)}{\hat{C}_{\hat{p}^*}(X_{1,1}|Y_1)+\hat{C}_{\hat{p}^*}(X_{2,1}|Y_1)}$ is calculated for $p = 0.99$ based on ten days' returns with a moving window of observations 250 or 500.



4.2 Simulation study

In this subsection, we examine the finite sample performance of the proposed nonparametric estimation by considering the models described in (12) with $d = 2$, $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ as given in (13), while $\boldsymbol{\eta}_t's$ are independent random vectors with t-copula and t-distributed marginals with all parameters obtained as in the above real data analysis.

Note that our new capital allocation $C_{p^*}(X_j|Y)$ could be different from the studied one $Q_p(X_j|Y)$ in the literature, and our proposed nonparametric estimator for $C_{p^*}(X_j|Y)$ has a faster rate of convergence than the one for $Q_p(X_j|Y)$. Unlike the nonparametric estimator for $Q_p(X_j|Y)$, the proposed nonparametric estimator for the new capital allocation does not need to choose a bandwidth. Moreover, there is no existing results in the literature of quantile capital allocation and sensitivity analysis on quantifying the inference uncertainty based on dependent sequence due to the difficult choice of bandwidth in terms of coverage accuracy. Hence we do not compare the newly proposed quantity and its inference with the Euler allocations based on VaR.

Firstly, in order to obtain the true values of C_1 and C_2 at $p = 0.95$ and 0.99, we draw $100,000$ random samples of size $n = 100,000$ from our models. The next step is to approximate the true values by the averages of these $\hat{C}_1$ and $\hat{C}_2$, which are 2.7294 for $p = 0.95$ and 5.6655 for $p = 0.99$, as reported in Table 2.

Secondly, we draw 400 random samples of size $n = 2000$ and 3000 from the above models and then compute $\hat{C}_1$ and $\hat{C}_2$, whose mean and standard deviation based on these 400 estimates are reported in Table 2 as well. The reason why we do not repeat a larger number of times is that the bootstrap method for a time series model is computationally intensive, because one has to refit the time series model to the resampling from the residuals. Due to the computational constraint, the bootstrap estimates of the standard deviations of the proposed estimates, as described after Theorem 3.1, are based on 400 repetitions, which are reported in Table 2. From Table 2, we observe that the proposed nonparametric estimator becomes more accurate in terms of mean squared errors when the sample size increases. Moreover, the asymptotic variance increases as p becomes larger, while $\hat{C}_1$ has a smaller variance than $\hat{C}_2$ under the considered setting, which is consistent with the observation in the real data analysis.

Table 2: *Nonparametric Estimate:* the true values and proposed nonparametric estimates are given when $p = 0.95$ and 0.99; the numbers in brackets are the standard deviations of $\hat{C}_1$ and $\hat{C}_2$, respectively.

(p, n)	True C_1	$\hat{C}_1$	Bootstrap SD	True C_2	$\hat{C}_2$	Bootstrap SD
(0.95, 2000)	1.9062	1.9137 (0.1451)	0.1545	1.6872	1.6778 (0.2678)	0.2648
(0.99, 2000)	3.2956	3.3177 (0.4242)	0.5265	3.1799	3.1391 (1.0224)	1.1787
(0.95, 3000)	1.9062	1.9055 (0.1216)	0.1203	1.6872	1.6546 (0.2106)	0.2018
(0.99, 3000)	3.2956	3.3138 (0.3835)	0.3741	3.1797	3.0588 (0.6767)	0.9099

5 Properties of the new allocation

Two examples in Section 2 show that the new allocation $C_{p^*}(X|Y)$ may be equal to or different from $Q_p(X|Y)$. As the level p given in risk management is always close to one, we would like $C_{p^*}(X|Y)$ to be close to $Q_p(X|Y)$ for a larger p. That is, we investigate when

$$\frac{Q_p(X|Y)}{C_{p^*}(X|Y)} \to 1 \quad \text{as } p \uparrow 1. \tag{14}$$

Again recall that p^* depends on p.

Assuming that Y is continuously distributed with X and Y being unbounded from above, i.e., $q_p(X) \to \infty$ and $q_p(Y) \to \infty$ as $p \uparrow 1$, it is obvious that (14) is equivalent to

$$\frac{\mathbb{E}[X|Y = y(t)]}{\mathbb{E}[X|Y > t]} \to 1 \quad \text{as } t \to \infty, \tag{15}$$

where $y(t) = \mathbb{E}[Y|Y > t]$. The next result shows that (15) holds (and therefore, (14) holds as well), under some assumptions.

Theorem 5.1. *Let $X, Y \in L^1$, unbounded from above, and Y is continuously distributed.*

(i) If $\lim_{t\to\infty} \mathbb{E}[X|Y = t]/t$ exists, is finite and non-zero, then (15) holds.

(ii) Assume that $X \geqslant 0$, there exists $t_0 > 0$ such that $\mathbb{E}[X|Y = t] < \infty$ for all $t > t_0$, and further,

$$\lim_{t\to\infty} \mathbb{P}(X > tx|Y = t) := g_1(x) \tag{16}$$

holds uniformly on $\mathbb{R}_+ := [0, \infty)$ with g_1 being integrable on $\mathbb{R}_+$. Then, $\mathbb{E}[X|Y = t]/t \to \int_0^\infty g_1(x)\,\mathrm{d}x$ as $t \to \infty$ and in turn, (15) holds as long as $\int_0^\infty g_1(x)\,dx > 0$.

(iii) Assume that $x_0 := \sup\{x \in \mathbb{R} : F(x) = 0\} > -\infty$ and there exists $t_0 > 0$ such that $\mathbb{E}[X|Y = t] < \infty$ for all $t > t_0$. If

$$\lim_{t\to\infty} \mathbb{P}(X > x|Y = t) := g_2(x) \tag{17}$$

holds uniformly on $[x_0, \infty)$, then (15) holds whenever $x_0 + \int_{x_0}^\infty g_2(x)$ is finite and non-zero.

(iv) Assume that (X, Y) has a density function f such that

$$\lim_{t\to\infty} \frac{f(tx, ty)}{v(t)} := q(x, y) \quad \textit{for } (x, y) \in \mathbb{R}^2, \tag{18}$$

where $v(t)$ is a regularly varying function at infinity with index $-\alpha$ for some $\alpha > 2$ (that is, $\lim_{t\to\infty} v(tx)/v(t) = x^{-\alpha}$ for all $x > 0$). Then, (15) holds as long as $\int_{-\infty}^\infty zq(z, 1)\mathrm{d}z$ and $\int_{-\infty}^\infty q(z, 1)\mathrm{d}z$ are finite and nonzero.

(v) If X and Y are jointly log-normally distributed, i.e., $(\log X, \log Y)$ is bivariate normal, then (15) holds.

Now, let us try to better understand the limits (16) and (17). If F and G are continuous, then there exists a unique survival copula $\hat{C}$ such that

$$\mathbb{P}(X > x, Y > y) = \hat{C}\big(\overline{F}(x), \overline{G}(y)\big) \quad \text{for all } (x, y) \in \mathbb{R}_+^2$$

(see Sklar (1959)). The partial derivatives of $\hat{C}$ with respect to first and second argument are denoted by $\hat{C}_1$ and $\hat{C}_2$, respectively, which exists almost surely (for details, see Nelsen (2006)). Clearly, $\mathbb{P}(X > tx|Y = t) = \hat{C}_2\big(\overline{F}(tx), \overline{G}(t)\big)$ for all $x \geqslant 0$. Now, if the following limits exist

$$\lim_{u \downarrow 0} \hat{C}_2(ux, u) := L_1(x) \quad \text{and} \quad \lim_{t \to \infty} \frac{\overline{F}(tx)}{\overline{G}(t)} := l(x) \quad \text{for all } x > 0,$$

such that the image of l, $l(\mathbb{R}_+)$, is a subset of the continuity set for L_1, $\mathcal{C}_{L_1}$, then (16) holds with $\tilde{g}_1 = L_1 \circ l$ on $\mathbb{R}_+$. This condition is not restrictive as L_1 is monotone and therefore, its discontinuity set, $\mathbb{R} \setminus \mathcal{C}_{L_1}$, is at most countable. A sufficient condition for (16) to hold uniformly is for g_1 to be continuous on $\mathbb{R}_+$.

The strength of upper tail dependence is stronger in condition (16) as compared to (17), which makes the main difference between them. A sufficient condition for (17) to hold uniformly is for g_2 to be continuous on $[x_0, \infty)$. As before, if F and G are continuous with survival copula $\hat{C}$, then

$$\lim_{u \downarrow 0} \hat{C}_2(x, u) := L_2(x), \quad x \in [0, 1] \tag{19}$$

implies that (17) holds with $\tilde{g}_2 = L_2 \circ \overline{F}$ on $[x_0, \infty)$. Many copulas with a weak tail dependence, known as *asymptotic independence*, satisfy this condition, and the trivial independence copula is the simplest example.

Recall that, in Example 2.2, the ratio

$$\frac{Q_p(X|Y)}{C_{p^*}(X|Y)} = \frac{2\alpha - 1}{2\sqrt{\alpha(\alpha - 1)}}$$

which does not converge to 1 as $p \to \infty$. In this example, $\mathbb{E}[X|Y = t]/t \to 0$ as $t \to \infty$ and $\lim_{t \to \infty} \mathbb{P}(X > x|Y = t) = 1$ for all $x \in \mathbb{R}$, suggesting that Theorem 5.1 does not apply.

Robustness issues are essential to various risk management procedures, including capital allocation and sensitivity analysis. We refer to Huber and Ronchetti (2009) for robust statistics and in particular, see Kou et al. (2013), Embrechts et al. (2015) and Kou and Peng (2016) for robustness in risk capital calculation. In the latter context, robustness refers to the continuity of the underlying quantity with respect to a small perturbation in the underlying model, typically modeled by a convergence in distribution.

In the following, we address the robustness issue of $C_{p^*}(X|Y)$. Consider a sequence $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$, let p^* be given by (3) and $p_n^* := \inf\{t \in [0,1] : c_t(Y_n) \geqslant q_p(Y_n)\}$, $n \in \mathbb{N}$. By convention, $\inf \varnothing = 0$. To show the appropriate weak convergence, we do not assume any particular structure for $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$, except for its uniform integrability and weak convergence to the true distribution of (X, Y). For statistical inference, such assumptions are almost always satisfied if (X_n, Y_n) represents the empirical distribution of (X, Y) from n stationary observations.

Theorem 5.2. *Let $p \in (0,1)$ and $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ be a sequence of uniformly integrable random vectors converging to a random vector (X, Y) in distribution. Assume that $q_t(Y)$ is continuous and strictly increasing in $t \in (0,1)$. If $\mathbb{E}[Y] \leqslant q_p(Y)$, then*

(i) $p_n^* \to p^*$ as $n \to \infty$;

(ii) $C_{p_n^*}(X_n|Y_n) \to C_{p^*}(X|Y)$ as $n \to \infty$.

The uniform integrability condition in Theorem 5.2 may be replaced by assuming that the sequence $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ is componentwise uniformly bounded from below and the convergence in expectation holds, i.e., $(\mathbb{E}[X_n], \mathbb{E}[Y_n]) \to (\mathbb{E}[X], \mathbb{E}[Y])$ as $n \to \infty$. This assertion is a direct consequence of Theorem 3.6 of Bilingsley (1999).

Theorem 5.2 suggests that if one estimates $C_{p*}(X|Y)$ using an empirical distribution of observations from (X, Y), then as long as the empirical distribution converges to the true distribution, an empirical version of $C_{p^*}(X|Y)$ (similar to (6)) converges to $C_{p^*}(X|Y)$ and thus it yields a consistent estimator.

The result in Theorem 5.2 directly applies to the capital allocation problem. If the following convergence in distribution

$$\{(X_{1,n}, \dots, X_{d,n})\}_{n\in\mathbb{N}} \to (X_1, \dots, X_d)$$

holds as $n \to \infty$ such that $\{(X_{1,n}, \dots, X_{d,n})\}_{n\in\mathbb{N}}$ is uniformly integrable, $q_t(Y)$ is continuous and strictly increasing in $t \in (0,1)$, then Theorem 5.2 implies that $C_{p_n^*}(X_{j,n}|Y_n) \to C_{p^*}(X_j|Y)$ as $n \to \infty$ for all $j = 1, \dots, n$, where $Y := X_1 + \cdots + X_d$, and $Y_n := X_{1,n} + \cdots + X_{d,n}$, $n \in \mathbb{N}$.

We remark that the robustness property of $C_{p^*}(X|Y)$ in Theorem 5.2 cannot be expected for $Q_p(X|Y)$. Typically, one needs stronger continuity assumptions to ensure the convergence of $Q_p(X_n|Y_n)$ to $Q_p(X|Y)$. We illustrate this by the following simple example.

Example 5.1. Let $Y_1, Z_1, Y_2, Z_2, \ldots$ be an iid standard normal sequence, and let $X_n = Z_n + Z_n^2 \mathrm{I}_{\{-1/n<Y_n<1/n\}}$, $n = 1, 2, \ldots$. It is clear that (X_n, Y_n) converges to a standard bivariate normal random vector (X, Y) in distribution as $n \to \infty$, and all assumptions in Theorem 5.2 are satisfied. However, $Q_{1/2}(X_n|Y_n) = \mathbb{E}[X_n|Y_n = 0] = 1$ and $Q_{1/2}(X|Y) = 0$, suggesting that the robustness property in Theorem 5.2 does not hold for $Q_p(X|Y)$.

Remark 5.1. Theorem 5.2 shows that the ES-based capital allocation is robust, although one might normally expect that an estimation for $\mathbb{E}[X|Y \geqslant q_p(Y)]$ might be more prone to outliers than that for $\mathbb{E}[X|Y = q_p(Y)]$. Note that in Theorem 5.2, we assumed that the sequence $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ is uniformly integrable. For such a sequence, "outliers" are confined to a certain magnitude, and this condition is important for the robustness result to hold. For robustness of ES on uniformly integrable sets, see e.g., Theorem 3.5 of Pesenti et al. (2016).

6 Conclusions

Capital allocation or sensitivity analysis based on both VaR and ES have been studied in the literature on many occasions. Because the problems are mathemtically equivalent, we decided to focus on the capital allocation formulation and its lingo associated with when we presented our results. Nonparametric estimation for an allocation based on VaR has a slower rate of convergence than that based on ES. This disadvantage is overcome in this paper by proposing to change the risk level via a connection with an ES-based allocation, so that the total capital still equals the reference quantile-based capital level. Therefore, the new allocation could be estimated nonparametrically at the standard rate of convergence. The asymptotic normality of the proposed nonparametric estimator for the new allocation is derived under the setup of a mixing sequence. In order to assess the performance of the estimation, a multivariate time series model and a bootstrap method based on residuals are proposed. A simulation study shows the effectiveness of the proposed inference. It is straightforward to generalize the idea in this paper to cover the case of netting agreement developed in Fermanian and Scaillet (2005).

7 Proofs

Proof of Proposition 2.1. First, note that $c_t(Y)$ is an increasing function in $t \in [0,1)$. By definition, $c_p(Y) \geqslant q_p(Y)$, thus $p^* \leqslant p$. Because $c_0(Y) = \mathbb{E}[Y] < q_p(Y)$, we know that $p^* \geqslant 0$. If Y is continuously distributed, then $c_t(Y) = \mathrm{ES}_t(Y)$ which is a continuous function of $t \in [0,1)$. Therefore, $c_{p^*}(Y) = q_p(Y)$. □

Proof of Theorem 3.1. Denote first $U_n(y) = \frac{1}{n}\sum_{t=1}^n I\big(G(Y_t) \leqslant y\big)$. Then, it follows from Proposition 4.4 of Berghaus et al. (2017) that for any $\delta \in (0,1/2)$, $\lambda \in (0,1)$ and $\delta_n \to 0$,

$$\left\{\begin{array}{l} \sup_{0<y<1} \frac{|\sqrt{n}(U_n(y)-y)|}{y^\delta(1-y)^\delta} = O_p(1), \quad \sup_{1/n^\lambda<y\leqslant 1-1/n^\lambda} \frac{|\sqrt{n}(U_n^-(y)-y)|}{y^\delta(1-y)^\delta} = O_p(1) \\ \\ \sup_{|u_1-u_2|+|v_1-v_2|\leqslant \delta_n} \frac{|\sqrt{n}\{C_n(u_1,v_1;j)-C(u_1,v_1;j)\}-\sqrt{\{C_n(u_2,v_2;j)-C(u_2,v_2;j)\}}|}{\max\{|u_1-u_2|^\delta+|v_1-v_2|^\delta, n^{-\delta}\}} = o_p(1). \end{array}\right. \tag{20}$$

By Proposition 2.1, we have

$$G^-(p) = -\frac{1}{1-p^*}\int_{G^-(p^*)}^{\infty} s\, d\{1-G(s)\} = G^-(p^*) + \frac{1}{1-p^*}\int_{G^-(p^*)}^{\infty}\{1-G(s)\}\, ds.$$

Hence we can write

$$\begin{aligned} 0 &= \tfrac{1}{1-\hat{p}^*}\int_{\hat{p}^*}^1 G_n^-(s)\,ds - G_n^-(p) \\ &= \tfrac{1}{1-\hat{p}^*}\int_{\hat{p}^*}^1 G_n^-(s)\,ds - \tfrac{1}{1-p^*}\int_{p^*}^1 G_n^-(s)\,ds + \tfrac{1}{1-p^*}\int_{p^*}^1 G_n^-(s)\,ds - G_n^-(p) \\ &= \tfrac{1}{1-\hat{p}^*}\int_{\hat{p}^*}^1 G_n^-(s)\,ds - \tfrac{1}{1-p^*}\int_{p^*}^1 G_n^-(s)\,ds - \tfrac{1}{1-p^*}\int_{G_n^-(p^*)}^{\infty} s\,d\big(1-G_n(s)\big) - G_n^-(p) \\ &= \tfrac{1}{1-\hat{p}^*}\int_{\hat{p}^*}^1 G_n^-(s)\,ds - \tfrac{1}{1-p^*}\int_{p^*}^1 G_n^-(s)\,ds + G_n^-(p^*) + \tfrac{1}{1-p^*}\int_{G_n^-(p^*)}^{G^-(p^*)}\big(1-G_n(s)\big)\,ds \\ &\quad + \tfrac{1}{1-p^*}\int_{G^-(p^*)}^{\infty}\big(1-G_n(s)\big)\,ds - G_n^-(p) \\ &= \tfrac{1}{1-\hat{p}^*}\int_{\hat{p}^*}^1 G_n^-(s)\,ds - \tfrac{1}{1-p^*}\int_{p^*}^1 G_n^-(s)\,ds + G_n^-(p^*) - G^-(p^*) \\ &\quad + \tfrac{1}{1-p^*}\int_{G_n^-(p^*)}^{G^-(p^*)}\big(1-G_n(s)\big)\,ds + \tfrac{1}{1-p^*}\int_{G^-(p^*)}^{\infty}\big(G(s)-G_n(s)\big)\,ds - \big(G_n^-(p)-G^-(p)\big) \\ &= (\hat{p}^*-p^*)\tfrac{-G_n^-(\tilde{p})(1-\tilde{p})+\int_{\tilde{p}}^1 G_n^-(s)\,ds}{(1-\tilde{p})^2} + \big(G_n^-(p^*)-G^-(p^*)\big) + \tfrac{1}{1-p^*}\int_{G_n^-(p^*)}^{G^-(p^*)}\big(1-G_n(s)\big)\,ds \\ &\quad + \tfrac{1}{1-p^*}\int_{p^*}^1\big(s-U_n(s)\big)\,dG^-(s) - \big(G^-(U_n^-(p)) - G^-(p)\big) \\ &= I_1 + \cdots + I_5, \end{aligned} \tag{21}$$

where $\tilde{p}$ is between p^* and $\hat{p}^*$.

Note first that equations (21), (8) and (20) imply that

$$\hat{p}^* \xrightarrow{p} p^* \quad \text{as} \quad n \to \infty. \tag{22}$$

By noting that

$$C_n(1, y; j) = 1 - U_n(1 - y) + O_p\left(\frac{1}{n}\right), \tag{23}$$

it follows from (20), (22), C3) and C4) that

$$\begin{aligned} \frac{-G_n^-(\tilde{p})(1-\tilde{p}) + \int_{\tilde{p}}^1 G_n^-(s)\,ds}{(1-\tilde{p})^2} &= \frac{-G^-(p^*)(1-p^*) + \int_{p^*}^1 G^-(s)\,ds}{(1-p^*)^2} + o_p(1) \\ &= \frac{G^-(p) - G^-(p^*)}{1-p^*} + o_p(1). \end{aligned} \tag{24}$$

Using (20), C3) and C4), we can show that

$$\sqrt{n}\{I_2 + I_3\} = \frac{\sqrt{n}}{1-p^*} \int_{G^-(U_n^-(p^*))}^{G^-(p^*)} \{1 - G(s) - (1-p^*)\}\,ds = o_p(1). \tag{25}$$

By (8), (20), (23) and the fact that $W(1, y; j) = W_d(y)$, we have

$$\sup_{0<s<1} \frac{\left|\sqrt{n}\big(U_n(s) - s\big) + W_d(1-s)\right|}{\{\min(s, 1-s)\}^\delta} = o_p(1) \tag{26}$$

for any $\delta \in (0, 1/2)$, which implies that

$$\sup_{0<s<1} \left|\sqrt{n}\big(U_n^-(s) - s\big) - W_d(1-s)\right| = o_p(1). \tag{27}$$

By (26) and C4), we have

$$\sqrt{n} I_4 = \frac{1}{1-p^*} \int_{p^*}^1 W_d(1-s)\,dG^-(s) + o_p(1). \tag{28}$$

Now, equation (27) suggests that

$$\sqrt{n} I_5 = -\frac{W_d(1-p)}{G'(G^-(p))} + o_p(1). \tag{29}$$

Hence, the asymptotic distribution of $\hat{p}^*$ follows from equations (21), (24), (25), (28)–(29), i.e., (9) holds.

Next write

$$
\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n} X_{j,t} I\big(Y_t > G_n^{-}(\hat{p}^*)\big) \\
&= \frac{1}{n}\sum_{t=1}^{n} F_j^{-}\big(F_j(X_{j,t})\big) I\big(1 - G(Y_t) \leqslant 1 - U_n^{-}(\hat{p}^*)\big) \\
&= \int_0^1 F_j^{-}(x)\, d\Big(C_n\big(x, 1 - U_n^{-}(\hat{p}^*); j\big) - C\big(x, 1 - U_n^{-}(\hat{p}^*); j\big)\Big) \\
&\quad + \int_0^1 F_j^{-}(x)\, d\Big(C\big(x, 1 - U_n^{-}(\hat{p}^*); j\big) - C\big(x, 1 - \hat{p}^*; j\big)\Big) \\
&\quad + \int_0^1 F_j^{-}(x)\, d\big(C(x, 1 - \hat{p}^*; j) - C(x, 1 - p^*; j)\big) + \int_0^1 F_j^{-}(x)\, dC(x, 1 - p^*; j) \\
&= -\int_0^{1/n} \Big(C_n\big(x, 1 - U_n^{-}(\hat{p}^*); j\big) - C\big(x, 1 - U_n^{-}(\hat{p}^*); j\big)\Big)\, dF_j^{-}(x) \qquad (30) \\
&\quad - \int_{1/n}^{1-1/n} \Big(C_n\big(x, 1 - U_n^{-}(\hat{p}^*); j\big) - C\big(x, 1 - U_n^{-}(\hat{p}^*); j\big)\Big)\, dF_j^{-}(x) \\
&\quad - \int_{1-1/n}^{1} \Big(C_n\big(x, 1 - U_n^{-}(\hat{p}^*); j\big) - C\big(x, 1 - U_n^{-}(\hat{p}^*); j\big)\Big)\, dF_j^{-}(x) \\
&\quad - \int_0^1 \Big(C\big(x, 1 - U_n^{-}(\hat{p}^*); j\big) - C\big(x, 1 - \hat{p}^*; j\big)\Big)\, dF_j^{-}(x) \\
&\quad - \int_0^1 \big(C(x, 1 - \hat{p}^*; j) - C(x, 1 - p^*; j)\big)\, dF_j^{-}(x) + \int_0^1 F_j^{-}(x)\, dC(x, 1 - p^*; j) \\
&= J_1 + \cdots + J_6.
\end{aligned}
$$

Using (8), (20), (27), C3) and C4), we can show that

$$
\begin{aligned}
\sqrt{n} J_1 &= o_p(1), \quad \sqrt{n} J_3 = o_p(1), \\
\sqrt{n} J_2 &= -\int_{1/n}^{1-1/n} W\big(x, 1 - U_n^{-}(p^*); j\big)\, dF_j^{-}(x) + o_p(1), \\
\sqrt{n} J_4 &= W_d(1 - p^*) \int_0^1 C^{(2)}(x, 1 - p^*; j)\, dF_j^{-}(x) + o_p(1),
\end{aligned}
$$

and

$$
\sqrt{n} J_5 = \sqrt{n}(\hat{p}^* - p^*) \int_0^1 C^{(2)}(x, 1 - p^*; j)\, dF_j^{-}(x) + o_p(1),
$$

which imply that

$$
\begin{aligned}
&\frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Big(X_{j,t} I\big(Y_t > G_n^{-}(\hat{p}^*)\big) - \int_0^1 F_j^{-}(x)\, dC(x, 1 - p^*; j)\Big) \\
&= -\int_0^1 W(x, 1 - p^*; j)\, dF_j^{-}(x) + W_d(1 - p^*) \int_0^1 C^{(2)}(x, 1 - p^*; j)\, dF_j^{-}(x) \qquad (31) \\
&\quad + \sqrt{n}(\hat{p}^* - p^*) \int_0^1 C^{(2)}(x, 1 - p^*; j)\, dF_j^{-}(x) + o_p(1).
\end{aligned}
$$

Therefore, the asymptotic limit of $\hat{C}_j$ follows from (9), (31) and the fact that

$$\begin{aligned}\sqrt{n}(\hat{C}_j - C_j) &= \sqrt{n}\left(\frac{1}{1-\hat{p}^*}\frac{1}{n}\sum_{t=1}^{n} X_{j,t} I\big(Y_t > G_n^-(\hat{p}^*)\big) - C_j\right)\\ &= \frac{\sqrt{n}(\hat{p}^* - p^*)}{(1-\hat{p}^*)(1-p^*)}\frac{1}{n}\sum_{t=1}^{n} X_{j,t} I\big(Y_t > G_n^-(\hat{p}^*)\big)\\ &\quad + \frac{\sqrt{n}}{1-p^*}\left(\frac{1}{n}\sum_{t=1}^{n} X_{j,t} I\big(Y_t > G_n^-(\hat{p}^*)\big) - \mathbb{E}\Big(X_{j,1} I\big(G(Y_1) > p^*\big)\Big)\right),\end{aligned}$$

i.e., (10) holds.

Similarly to (30), we have

$$\begin{aligned} & \sqrt{n}\{\tfrac{1}{n}\textstyle\sum_{t=1}^{n} Y_t I(Y_t > G_n^-(\hat{p}^*)) - \int_{p^*}^1 G^-(y)\,dy\}\\ =\ & \textstyle\int_{U_n^-(\hat{p}^*)}^1 G^-(y)\,d\sqrt{n}\{U_n(y) - y\} + \sqrt{n}\int_{U_n^-(\hat{p}^*)}^{\hat{p}^*} G^-(y)\,dy + \sqrt{n}\int_{\hat{p}^*}^{p^*} G^-(y)\,dy\\ =\ & -\textstyle\int_{U_n^-(\hat{p}^*)}^1 \sqrt{n}\{U_n(y) - y\}\,dG^-(y) - G^-(p^*)\sqrt{n}(\hat{p}^* - p^*) + o_p(1)\\ =\ & \textstyle\int_{p^*}^1 W_d(1-y)\,dG^-(y) - G^-(p^*)\sqrt{n}(\hat{p}^* - p^*) + o_p(1).\end{aligned} \tag{32}$$

Hence, the asymptotic distribution of $\sum_{j=1}^d \hat{C}_j$ follows from (9), (26), (32) and the facts that $W(1, y; j) = W_d(y)$ and

$$\begin{aligned} & \sqrt{n}\textstyle\sum_{j=1}^{n}(\hat{C}_j - C_j)\\ =\ & \sqrt{n}\left(\tfrac{1}{1-\hat{p}^*}\tfrac{1}{n}\textstyle\sum_{t=1}^{n} Y_t I\big(Y_t > G_n^-(\hat{p}^*)\big) - \sum_{j=1}^{d} C_j\right)\\ =\ & \tfrac{\sqrt{n}(\hat{p}^*-p^*)}{(1-\hat{p}^*)(1-p^*)}\tfrac{1}{n}\textstyle\sum_{t=1}^{n} Y_t I\big(Y_t > G_n^-(\hat{p}^*)\big)\\ & + \tfrac{\sqrt{n}}{1-p^*}\left(\tfrac{1}{n}\textstyle\sum_{t=1}^{n} Y_t I\big(Y_t > G_n^-(\hat{p}^*)\big) - \mathbb{E}\Big(Y_1 I\big(G(Y_1) > p^*\big)\Big)\right),\end{aligned}$$

i.e., (11) holds. □

Proof of Theorem 5.1. (i) Assume that $\mathbb{E}[X|Y = t]/t \to r$ as $t \to \infty$. Without loss of generality $r > 0$ is further assumed. Then, for any $\delta > 0$ there exists a sufficiently large t_0 such that

$$1 - \delta < \frac{\mathbb{E}[X|Y = t]}{rt} < 1 + \delta \quad \text{for all } t > t_0.$$

Integrating the latter with respect to the distribution of Y on (t_1, ∞), where $t_1 > t_0$, one arrives at

$$\frac{\mathbb{E}[X|Y > t_1]}{\mathbb{E}[Y|Y > t_1]} = \frac{\int_{t_1}^{\infty} \mathbb{E}[X|Y = t]\mathbb{P}(Y \in \mathrm{d}t)}{\int_{t_1}^{\infty} t\mathbb{P}(Y \in \mathrm{d}t)} \in \big((1-\delta)r, (1+\delta)r\big).$$

Letting $t_1 \to \infty$ and noting that δ is arbitrary, we have that $\frac{\mathbb{E}[X|Y>t]}{\mathbb{E}[Y|Y>t]} \to r$ as $t \to \infty$. Thus,

$$\frac{\mathbb{E}[X|Y = y(t)]}{\mathbb{E}[X|Y > t]} = \frac{\mathbb{E}[X|Y = y(t)]}{y(t)} \times \frac{\mathbb{E}[Y|Y > t]}{\mathbb{E}[X|Y > t]} \to r \times r^{-1} = 1 \quad \text{as } t \to \infty,$$

because $y(t) \to \infty$ as $t \to \infty$.

(ii) Note that $\mathbb{E}[X|Y=t] < \infty$ for all $t > t_0$ and thus, a simple change of variable implies that

$$\begin{aligned} \frac{\mathbb{E}(X|Y=t)}{t} &= \frac{1}{t}\int_0^\infty \mathbb{P}(X > z|Y=t)\,\mathrm{d}z \\ &= \int_0^\infty \mathbb{P}(X > tx|Y=t)\,\mathrm{d}x \\ &\to \int_0^\infty g_1(x)\,\mathrm{d}x \quad \textit{as } t \to \infty, \end{aligned}$$

where the last implication is due to the fact that (16) holds uniformly on $\mathbb{R}_+$.

(iii) Similar to the proof of Part (ii), as (17) holds uniformly on $[x_0, \infty)$, one may integrate (17) to get

$$\mathbb{E}(X|Y=t) = x_0 + \int_{x_0}^\infty \mathbb{P}(X > x|Y=t)\,\mathrm{d}x \to x_0 + \int_{x_0}^\infty g_2(x)\,\mathrm{d}x \quad \textit{as } t \to \infty,$$

where the identity holds due to the fact that $\mathbb{E}[X|Y=t] < \infty$ for all $t > t_0$.

(iv) The marginal density f_2 of Y satisfies

$$\frac{f_2(ty)}{tv(t)} = \int_{-\infty}^\infty \frac{f(x,ty)}{tv(t)}\mathrm{d}x = \int_{-\infty}^\infty \frac{f(tz,ty)}{v(t)}\mathrm{d}z \to \int q(z,y)\mathrm{d}z \qquad \text{as } t \to \infty,$$

due to some simple change of variables and Potter's bound of a regular variation in (18) in the last step. It follows that

$$\begin{aligned} \mathbb{E}[X|Y=t]/t &= \int_{-\infty}^\infty xf(x,t)\mathrm{d}x/\big(tf_2(t)\big) \\ &= \int_{-\infty}^\infty \frac{zf(tz,t)}{v(t)}\mathrm{d}z \times \frac{tv(t)}{f_2(t)} \\ &\to \frac{\int_{-\infty}^\infty zq(z,1)\mathrm{d}z}{\int_{-\infty}^\infty q(z,1)\mathrm{d}z} \qquad \text{as } t \to \infty, \end{aligned}$$

where the derivations are again based on change of variables and Potter's bound of a regular variation in (18). Consequently, our claim holds as $\int_{-\infty}^\infty zq(z,1)\mathrm{d}z$ and $\int_{-\infty}^\infty q(z,1)\mathrm{d}z$ are finite and nonzero.

(v) Assume that Z and W are both $\mathrm{N}(0,1)$ random variables with correlation coefficient $\rho \in [-1,1]$. Note that for $w \in \mathbb{R}$, $Z|W=w$ follows a normal distribution with mean ρw and variance $1-\rho^2$. That is, $Z|W=w$ is identically distributed as $T+\rho w$, where T is $\mathrm{N}(0,1-\rho^2)$ distributed. If $\rho = 0$, then Z and W are independent and we have $Q_p(f(Z)|g(W)) = \mathbb{E}[f(Z)] =$

$C_{p^*}(f(Z)|g(W))$. In the following, we assume that $\rho \neq 0$. Define the functions f and g on $\mathbb{R}$ by $f(z) = C_1 e^{az}$ and $g(w) = C_2 e^{bw}$ for some $C_1, C_2, a, b > 0$. Let ϕ and Φ be the pdf and cdf of the standard normal distribution. Recall now the elementary relationship

$$\lim_{w\to\infty} \frac{\phi(w)}{w(1-\Phi(w))} = 1. \tag{33}$$

For $w \in \mathbb{R}$, let $z_w \in \mathbb{R}$ be such that $g(z_w) = \mathbb{E}[g(W)|W > w]$. Thus,

$$\lim_{w\to\infty} \frac{g(z_w)}{g(w)} = \lim_{w\to\infty} \frac{\mathbb{E}[g(W)|W > w]}{g(w)} = \lim_{w\to\infty} \frac{-g(w)\phi(w)}{-g(w)\phi(w) + (1-\Phi(w))g'(w)},$$

where the L'Hôpital's rule is applied (if the limit exists). By the definition of g and (33),

$$\lim_{w\to\infty} \frac{(1-\Phi(w))g'(w)}{g(w)\phi(w)} = \lim_{w\to\infty} \frac{b}{w} = 0,$$

which in turn gives that

$$\lim_{w\to\infty} e^{b(z_w - w)} = \lim_{w\to\infty} \frac{g(z_w)}{g(w)} = 1,$$

implying that $z_w - w \to 0$ as $w \to \infty$.

On the other hand,

$$\begin{aligned}
\lim_{w\to\infty} \frac{\mathbb{E}[f(Z)|W > w]}{\mathbb{E}[f(Z)|W = w]} &= \lim_{w\to\infty} \frac{1}{1-\Phi(w)} \frac{\int_w^\infty \mathbb{E}[f(Z)|W = x]\mathrm{d}\Phi(x)}{\mathbb{E}[f(Z)|W = w]} \\
&= \lim_{w\to\infty} \frac{\int_w^\infty \mathbb{E}[f(T+\rho x)]\phi(x)\mathrm{d}x}{(1-\Phi(w))\mathbb{E}[f(T+\rho w)]} \\
&= \lim_{w\to\infty} \frac{-\mathbb{E}[f(T+\rho w)]\phi(w)}{-\mathbb{E}[f(T+\rho w)]\phi(w) + (1-\Phi(w))\mathbb{E}[f'(T+\rho w)\rho]},
\end{aligned}$$

where the L'Hôpital's rule is applied once again. By the definition of f and (33),

$$\lim_{w\to\infty} \frac{(1-\Phi(w))\mathbb{E}[f'(T+\rho w)]}{\mathbb{E}[f(T+\rho w)]\phi(w)} = \lim_{w\to\infty} \frac{\mathbb{E}[f'(T+\rho w)]}{w\mathbb{E}[f(T+\rho w)]} = \lim_{w\to\infty} \frac{a}{w} = 0.$$

Therefore,

$$\lim_{w\to\infty} \frac{\mathbb{E}[f(Z)|W > w]}{\mathbb{E}[f(Z)|W = w]} = 1. \tag{34}$$

Moreover, using the fact $z_w - w \to 0$ as $w \to \infty$, we obtain that

$$\lim_{w\to\infty} \frac{\mathbb{E}[f(Z)|W = z_w]}{\mathbb{E}[f(Z)|W = w]} = \lim_{w\to\infty} \frac{\mathbb{E}[f(T+\rho z_w)]}{\mathbb{E}[f(T+\rho w)]} = \lim_{w\to\infty} \frac{\mathbb{E}[f(T+\rho w)]e^{a\rho(z_w - w)}}{\mathbb{E}[f(T+\rho w)]} = 1.$$

By (34), we have

$$\lim_{p\uparrow 1} \frac{C_{p^*}(f(Z)|g(W))}{Q_p(f(Z)|g(W))} = \lim_{w\to\infty} \frac{\mathbb{E}[f(Z)|g(W) > g(w)]}{\mathbb{E}[f(Z)|g(W) = g(z_w)]} = \lim_{w\to\infty} \frac{\mathbb{E}[f(Z)|W > w]}{\mathbb{E}[f(Z)|W = z_w]} = 1.$$

Therefore, by taking $X = f(Z)$ and $Y = g(W)$, (14) holds for any jointly log-normally distributed X and Y. $\square$

Proof of Theorem 5.2. (i) By the convergence in distribution of $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ to (X, Y), we have $q_t(X_n) \to q_t(X)$ for all $t \in (0, 1)$ such that the quantile function of X is continuous and $q_t(Y_n) \to q_t(Y)$ for all $t \in (0, 1)$. Because $q_t(Y_n)$ is an increasing function in t, and $q_t(Y)$ is continuous in t, we have that $q_t(Y_n) \to q_t(Y)$ as $n \to \infty$ holds uniformly on any interval $[a, b] \subset (0, 1)$.

Because $\{(X_n, Y_n)\}_{n\in\mathbb{N}}$ is uniformly integrable, $c_t(Y_n) \to c_t(Y)$ as $n \to \infty$ for all $t \in [0, 1)$. As a consequence, the functions $f_n : [0, 1) \to \mathbb{R}$, $t \mapsto c_t(Y_n) - q_p(Y_n)$ converge to $c_t(Y) - q_p(Y)$ as $n \to \infty$ for each $t \in (0, 1)$. Note that both $q_t(Y)$ and $c_t(Y)$ are continuous and strictly increasing functions of $t \in [0, 1)$. Because the function $f : [0, 1) \to \mathbb{R}$, $t \mapsto c_t(Y) - q_p(Y)$ is continuous and strictly increasing in t, we have $p_n^* \to p^*$ as $n \to \infty$.

(ii) We utilize the result from Part(i). The uniform convergence of $q_t(Y_n) \to q_t(Y)$ as $n \to \infty$ in a neighbourhood of p^* and the convergence $p_n^* \to p^*$ as $n \to \infty$ yield that

$$|q_{p_n^*}(Y_n) - q_{p^*}(Y)| \leqslant |q_{p_n^*}(Y_n) - q_{p_n^*}(Y)| + |q_{p_n^*}(Y) - q_{p^*}(Y)| \to 0.$$

Writing $A = \{Y > q_{p^*}(Y)\}$, $A_n = \{Y_n > q_{p_n^*}(Y_n)\}$ and $B_n = \{Y_n > q_{p^*}(Y)\}$ for all $n \in \mathbb{N}$, we have

$$C_{p_n^*}(X_n|Y_n) = \frac{\mathbb{E}\left[X_n I(A_n)\right]}{\mathbb{P}(A_n)} = \frac{\mathbb{E}\left[X_n I(A_n)\right]}{1 - p_n^*}.$$

Note that $(X_n, Y_n) \to (X, Y)$ in distribution as $n \to \infty$. In addition, the mapping $(x, y) \mapsto xI(y > q_{p^*}(Y))$ has discontinuity points $D = \{(x, q_{p^*}(Y)) \in \mathbb{R}^2 : x \neq 0\}$ with $\mathbb{P}((X, Y) \in D) = 0$. By the Continuous Mapping Theorem (for example, see Theorem 2.7 of Bilingsley (1999)), $\{X_n I(B_n)\}$ converges to $XI(A)$ in distribution as $n \to \infty$.

Note that for any $x \in \mathbb{R}$, $|\mathbb{P}(X_n I(A_n) \leqslant x) - \mathbb{P}(X_n I(B_n) \leqslant x)| \leqslant \mathbb{E}[|I(A_n) - I(B_n)|]$. Whenever $q_{p_n^*}(Y_n) \leqslant q_{p^*}(Y)$, then $B_n \subset A_n$ and

$$\mathbb{E}[|I(A_n) - I(B_n)|] = \mathbb{P}(A_n \setminus B_n) = \mathbb{P}(q_{p_n^*}(Y_n) < Y_n \leqslant q_{p^*}(Y)).$$

Further, for any $\varepsilon > 0$ and a sufficiently large n, the following is true

$$\mathbb{P}(q_{p_n^*}(Y_n) < Y_n \leqslant q_{p^*}(Y)) \leqslant \mathbb{P}(q_{p^*}(Y) - \varepsilon < Y_n \leqslant q_{p^*}(Y)),$$

which converges to $\mathbb{P}(q_{p^*}(Y)) - \varepsilon < Y \leqslant q_{p^*}(Y))$ when $n \to \infty$, as Y is continuously distributed. As ε is arbitrarily chosen, we have $\mathbb{E}[|I(A_n) - I(B_n)|] \to 0$ as $n \to \infty$. Similarly, if $q_{p_n^*}(Y_n) > q_{p^*}(Y)$, we also have $\mathbb{E}[|I(A_n) - I(B_n)|] \to 0$ when $n \to \infty$. In summary, $|\mathbb{P}(X_n I(A_n) \leqslant$

$x) - \mathbb{P}(X_n I(B_n) \leqslant x)| \to 0$ as $n \to \infty$. Because $\{X_n I(B_n)\}$ converges to $XI(A)$ in distribution as $n \to \infty$, we conclude that $\{X_n I(A_n)\}$ also converges to $XI(A)$ in distribution as $n \to \infty$. Noting that $\{X_n I(A_n)\}_{n\in\mathbb{N}}$ is uniformly integrable, we have $\mathbb{E}[X_n I(A_n)] \to \mathbb{E}[XI(A)]$ when $n \to \infty$. Consequently,

$$\lim_{n\to\infty} C_{p_n^*}(X|Y) = \lim_{n\to\infty} \frac{\mathbb{E}[X_n I(A_n)]}{1-p_n^*} = \frac{\mathbb{E}[XI(A)]}{1-p^*} = C_{p^*}(X|Y).$$

□

Acknowledgments

We thank two reviewers and the co-editor, Professor Vadim Linetsky, for helpful comments. Peng's research was partly supported by the Simons Foundation. Wang acknowledges financial support form the Natural Sciences and Engineering Research Council of Canada (RGPIN-2018-03823/RGPAS-2018-522590).

References

Acharya, V.V. (2009). A Theory of Systemic Risk and Design of Prudential Bank Regulation. *Journal of Financial Stability*, **5**(3), 224–255.

Acharya, V., Engle, R. and Richardson, M. (2012). Capital Shortfall: A New Approach to Ranking and Regulating Systemic Risks. *American Economic Review*, **102**(3), 59-64.

Adrian, T. and Brunnermeier, M. K. (2016). CoVaR. *American Economic Review*, **106**, 1705–1741.

BCBS (2016). *Standards. Minimum Capital Requirements for Market Risk. January 2016.* Basel Committee on Banking Supervision. Basel: Bank for International Settlements.

Berghaus, B., Bücher, A. and Volgushev, S. (2017). Weak Convergenc of the Empirical Copula Process with Respect to Weighted Metrics. *Bernoulli*, **23**(1), 743–772.

Bernard, C. and Vanduffel, S. (2015). A New Approach to Assessing Model Risk in High Dimensions. *Journal of Banking and Finance*, **58**, 166–178.

Billingsley, P. (1999). *Convergence of Probability Measures*, Second Edition, Wiley.

Cai, J., Liu, H. and Wang, R. (2018). Asymptotic Equivalence of Risk Measures under Dependence Uncertainty. *Mathematical Finance*, **28**, 29–49.

Cambanis, S., Huang, S. and Simons, G. (1981). On the Theory of Elliptically Contoured Distributions. *Journal of Multivariate Analysis*, **11**(3), 368–385.

Chen, X. and Fan, Y. (2006). Estimation of Copula-based Semiparametric Time Series Models. *Journal of Econometrics*, **130**(2), 307–335.

Chen, C., Iyengar, G. and Moallemi, C. C. (2013). An Axiomatic Approach to Systemic Risk. *Management Science*, **59**(6), 1373–1388.

Cheng, M. Y. and Peng, L. (2002). Regression Modeling for Nonparametric Estimation of Distribution and Quantile Functions. *Statistica Sinica*, **12**, 1043–1060.

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

Denault, M. (2001). Coherent Allocation of Risk Capital. *Journal of Risk*, **4**, 1–34.

EIOPA (2011). Equivalence Assessment of the Swiss Supervisory System in Relation to Articles 172, 227 and 260 of the Solvency II Directive, *EIOPA-BoS-11-028*, available at https://eiopa.europa.eu/consultations/consultation-papers/index.html.

Embrechts, P., Puccetti, G. and Rüschendorf, L. (2013). Model Uncertainty and VaR Aggregation. *Journal of Banking and Finance*, **37**(8), 2750–2764.

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

Embrechts, P., Wang, B. and Wang, R. (2015). Aggregation-robustness and Model Uncertainty of Regulatory Risk Measures. *Finance and Stochastics*, **19**(4), 763–790.

Emmer, S., Kratz, M. and Tasche, D. (2015). What is the Best Risk Measure in Practice? A Comparison of Standard Measures. *Journal of Risk*, **18**(2), 31–60.

Fermanian, J. D. and Scaillet, O. (2003). Nonparametric Estimation of Copulas for Time Series. *Journal of Risk*, **5**, 25–54.

Fermanian, J. D. and Scaillet, O. (2005). Sensitivity Analysis of VaR and Expected Shortfall for Portfolios under Netting Agreements. *Journal of Banking and Finance*, **29**, 927–958.

Fissler, T. and Ziegel, J. F. (2016). Higher Order Elicitability and Osband's Principle. *Annals of Statistics*, **44**(4), 1680–1707.

Francq, C. and Zakoïan, J. (2004). Maximum Likelihood Estimation of Pure GARCH and ARMA-GARCH Processes. *Bernoulli*, **10**(4), 605–637.

Hall, P. (1990). Using the bootstrap to estimate mean squared error and select smoothing parameter in nonparametric problems. *Journal of Multivariate Analysis*, **32**, 177–203.

Fu, M.C., Hong, L.J. and Hu, J.Q. (2009). Conditional Monte Carlo Estimation of Quantile Sensitivities. *Management Science*, **55**(12), 2019–2027.

Glasserman, P. (2005). Measuring Marginal Risk Contributions in Credit Portfolios. *Journal of Computational Finance*, **9**(1), 1–41.

Glasserman, P. and Li, J. (2005). Importance Sampling for Portfolio Credit Risk. *Management Science*, **51**(11), 1643–1656.

Gourieroux, C., Laurent, J.P. and Scaillet, O. (2000). Sensitivity Analysis of Values at Risk. *Journal of Empirical Finance*, **7**, 225–245.

Hong, L.J. (2009). Estimating Quantile Sensitivities. *Operations Research*, **57**(1), 118–130.

Hong, L.J. and Liu, G. (2009). Simulating Sensitivities of Conditional Value at Risk. *Management Science*, **55**(2), 281–293.

Huber, P. J. and Ronchetti E. M. (2009). *Robust Statistics.* Second ed., Wiley Series in Probability and Statistics. Wiley, New Jersey. First ed.: Huber, P. (1980).

IAIS (2014). *Consultation Document. Risk-based Global Insurance Capital Standard. December 2014.* International Association of Insurance Supervisors.

Jiang, G. and Fu, M.C. (2015). Technical Note – On Estimating Quantile Sensitivities via Infinitesimal Perturbation Analysis. *Operations Research*, **63**(2), 435–441.

Kalkbrener, M. (2005). An Axiomatic Approach to Capital Allocation. *Mathematical Finance*, **15**(3), 425–437.

Kou, S. and Peng, X. (2016). On the Measurement of Economic Tail Risk. *Operations Research*, **64**(5), 1056–1072.

Kou, S., Peng, X. and Heyde, C. C. (2013). External Risk Measures and Basel Accords. *Mathematics of Operations Research*, **38**(3), 393–417.

Krätschmer, V., Schied, A. and Zähle, H. (2014). Comparative and Quantitiative Robustness for Law-invariant Risk Measures. *Finance and Stochastics*, **18**(2), 271–295.

Landsman, Z. M. and Valdez, E. A. (2003). Tail Conditional Expectations for Elliptical Distributions. *North American Actuarial Journal*, **7**(4), 55–71.

Ling, S. (2007). Self-weighted and Local Quasi-maximum Likelihood Estimators for ARMA-GARCH/IGARCH models. *Journal of Econometrics*, **140**, 849–873.

Liu, G.W. and Hong, L.J. (2009). Kernel Estimation of Quantile Sensitivities. *Naval Research Logistics*, **56**, 511–525.

McNeil, A. J., Frey, R. and Embrechts, P. (2015). *Quantitative Risk Management: Concepts, Techniques and Tools.* Revised Edition. Princeton, NJ: Princeton University Press.

Nelsen, R. B. (2006). *An Introduction to Copulas, Second edition*, Springer, New York.

Pesenti, S. M., Millossovich, P. and Tsanakas, A. (2016). Robustness Regions for Masures of Risk Aggregation. Dependence Modeling, 4(1), pp. 348–367.

Rogers, L. C. G. and Veraart, L. A. M. (2013). Failure and Rescue in an Interbank Network. *Management Science*, **59**(4), 882–898.

Scaillet, O. (2004). Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall. *Mathematical Finance*, **14**(1), 115–129.

Scaillet, O. (2005). Nonparametric Estimation of Conditional Expected Shortfall. *Insurance and Risk Management Journal*, **74**(1), 639–660.

Siller, T. (2013). Measuring Marginal Risk Contributions in Credit Portfolios. *Quantitative Finance*, **13**(12), 1915–1923.

Sandström, A. (2010). *Handbook of Solvency for Actuaries and Risk Managers: Theory and Practice.* CRC Press/Taylor & Francis, Boca Raton, Florida.

Sklar, A. (1959). Fonctions de Répartion à n Dimensions et Leurs Marges. *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

Tasche, D. (2009). Capital Allocation for Credit Portfolios with Kernel Estimators. *Quantitative Finance*, **9**(5), 581–595.

Tsanakas, A. and Millossovich, P. (2016). Sensitivity Analysis using Risk Measures. *Risk Analysis*, **36**(1), 30–48.

Zhu, K. and Ling, S. (2011). Global Self-weighted and Local Quasi-maximum Exponential Likelihood Estimators for ARMA-GARCH/IGARCH models. *Annals of Statistics*, **39**, 2131–2163.

Ziegel, J. (2016). Coherence and Elicitability. *Mathematical Finance*, **26**, 901–918.