Controlling Tail Risk Measures with Estimation Error

Tetsuya Kaji[1] and Hyungjune Kang[2]

[1]University of Chicago Booth School of Business*
[2]Two Sigma Investments†

March 2, 2020

Abstract

Risk control plays an important role in economic decision making and banking regulations. In practice, controlling risks such as Value-at-Risk and expected shortfall entails estimation risk that is often ignored. This paper provides a novel way to hedge against estimation risk in general risk management setups. Our starting point is the assumption that a valid (one-sided) confidence interval is available. We then define a class of tail risk measures that control the probabilities of undesirable events, and bound the true but unobservable risk probability by observable quantities using the Bonferroni inequality. An empirical application to Value-at-Risk and expected shortfall illustrates how the proposed method can be applied to practical risk control problems. In situations where estimation precision cannot be quantified, a multiplier of 2 is recommended to account for estimation risk.

Keywords: risk measures, estimation risk, value-at-risk, expected shortfall, banking regulations.

JEL Codes: D81, G28, G32, C58.

*tkaji@chicagobooth.edu
†hjkang.1006@gmail.com

1 Introduction

Risk control plays an important role in economic decision making (Markowitz, 1952; Roy, 1952). In banking regulation, Basel II imposes capital requirements based on Value-at-Risk (VaR); Basel 2.5 adds requirements based on expected shortfall (ES) (Chen, 2014).[1] In search of the critical aspects of risk, many risk measures and properties thereof are studied (Artzner et al., 1999; Acerbi, 2002; Dowd and Blake, 2006; Ahmadi-Javid, 2012).

In practice, true risk measures are not observed and hence need to be estimated. McNeil et al. (2005, pp. 40–41) warn the danger of a naive use of estimated risk, pointing out that they are subject to estimation error, model risk, and market liquidity. For the latter two, there are proposed ways to account in risk management. Stahl (1997) argues that multiplying 3 to the estimated VaR hedges against the model risk as employed by Basel II (Chen, 2014);[2] Leippold and Vanini (2002) show that the multiplier of 1.5 is enough for the model risk on ES; see also Hendricks and Hirtle (1997), Lopez (1998), and Novak (2010). Also, many institutions use "liquidity-adjusted VaR" that takes into account realistic holding periods; Gârleanu and Pedersen (2007) analyze the feedback effect of such practice; Adrian and Brunnermeier (2016) propose a measure of systemic risk that subsumes liquidity risk.

In contrast, despite the vast literature of statistical inference on risk measures, a practical way to incorporate estimation error into generic risk management problems has not been established. In a related context, estimation error is notoriously problematic in Markowitz's (1952) mean-variance model (Michaud, 1989; Chopra and Ziemba, 1993), and many portfolio optimization methods that incorporate estimation error are proposed (Ceria and Stubbs, 2006; Michaud and Michaud, 2008). However, not all risk management problems are written in the form of portfolio optimization (notably banking regulation), and a general method to uphold desired risk guarantee is needed.

This paper proposes a general method to incorporate estimation error into the control of risk measures that are given by probability bounds. We start with the assumption that a valid (one-sided) confidence interval is available and provide a simple method to bound the true (but unobservable) probability associated with the risk by a quantity that is observable.

[1] Expected shortfall is originally named *conditional Value-at-Risk* (Rockafellar and Uryasev, 2000). We use *expected shortfall* since conditional Value-at-Risk sometimes refers to a different concept in the literature (Chernozhukov and Umantsev, 2001; Wang and Zhao, 2016).

[2] Stahl's (1997) justification was retrospective, after Basel II first employed the multiplier of 3.

We call the risk measure to which our method applies the *tail risk measure*, and show that VaR and ES—arguably the two most popular risk measures—are examples of it.

To motivate the paper, we note that there is a long history of concerns on estimation error in practically used risk estimates (Jorion, 1996; Hendricks, 1996; Pritsker, 1997a,b, 2006; Barone-Adesi and Giannopoulos, 2001; Berkowitz and O'Brien, 2002; Aussenegg and Miazhynskaia, 2006; Thiele, 2019). Chen (2008) notes that the effective sample size for ES at confidence level $1-p$ is the actual sample size times p^2, stating that the estimator's high volatility is a common challenge for statistical inference. Caccioli et al. (2017) note that because of high dimensionality of institutional portfolios and the lack of stationarity in the long run, portfolio optimization is plagued by (relatively) small sample sizes. In Section 2, we demonstrate that estimation error indeed acts adversally to risk control by simulation.

To recover the stated risk guarantee, we observe that many risk measures make statements about the tail event that happens with a small probability, and apply the Bonferroni inequality with a valid confidence interval. To illustrate the idea, consider VaR at confidence level $1-p$ defined by $\Pr(X \leq -\mathrm{VaR}_p) \leq p$, where X is considered a random variable that represents the return of a portfolio or the profits of a company. "Controlling VaR at $1-p$" thus means identifying the value that bounds the probability of a loss exceeding it by p and holding liquid assets worth thereof. True VaR is unobservable, however, as the distribution of X is unknown, and we need to use some estimate $\widehat{\mathrm{VaR}}_p$ in practice. The problem with such estimates is that there is no guarantee that they satisfy $\Pr(X \leq -\widehat{\mathrm{VaR}}_p) \leq p$; therefore, holding liquid assets worth $\widehat{\mathrm{VaR}}_p$ does not ensure that the probability of insolvency is capped by p. We simulate a generalized autoregressive conditional heteroskedasticity (GARCH) model in Section 2 and show that $\Pr(X \leq -\widehat{\mathrm{VaR}}_p)$ in fact exceeds p, sometimes by a huge margin, for three choices of $\widehat{\mathrm{VaR}}_p$.[3]

Now, suppose that a one-sided $(1-r)$-confidence interval for VaR_q is available, that is, $\Pr(\mathrm{VaR}_q \geq \widetilde{\mathrm{VaR}}_{q,r}) \leq r$ for some observable $\widetilde{\mathrm{VaR}}_{q,r}$. Then, by the Bonferroni inequality,

$$\Pr(X \leq -\widetilde{\mathrm{VaR}}_{q,r}) \leq \Pr(X \leq -\mathrm{VaR}_q) + \Pr(-\mathrm{VaR}_q \leq -\widetilde{\mathrm{VaR}}_{q,r}) = q + r.$$

Thus, $\widetilde{\mathrm{VaR}}_{q,r}$ controls VaR at confidence level $1-q-r$; we can choose $q+r \leq p$ to control VaR at confidence level p. This bound is conservative in finite samples and (q,r) are tuning

[3]Simulation of i.i.d. returns exhibits the same problem. See Kaji (2018) for details.

parameters that need to be specified in practice. We propose to set (q, r) so that the bound is asymptotically non-conservative. Note that under standard asymptotics, $\widetilde{\mathrm{VaR}}_{q,r}$ for any fixed r converges to true VaR_q. Therefore, it is optimal to gradually let $r \to 0$ and $q \to p$ as the sample size grows so that no portion of the probability is wasted in the limit. Meanwhile, shrinking r too fast would result in $\widetilde{\mathrm{VaR}}_{q,r}$ not converging to VaR_p. We characterize this tradeoff when $\widetilde{\mathrm{VaR}}_{q,r}$ is constructed from an asymptotically normal estimator and advocate the use of $r = 1/T$.

Building on the above intuition, we define a class of risk measures to which the same argument applies. We call such risk measures the *tail risk measures*. Roughly speaking, if a risk measure prescribes the probability of some bad event to be low, then it is compatible with the Bonferroni inequality. Precisely, let $\chi(F_X)$ be the measure of "badness" of the distribution F_X of X. A tail risk measure associated with χ gives the value c_q such that the probability of an event E in which $\chi(F_{X|E})$ exceeds c_q is capped by q. Thus, with a $(1-r)$ confidence interval for this tail risk measure, the upper bound thereof controls the tail risk measure with confidence level $(1-q-r)$.[4] We then show that two most popular risk measures, VaR and ES, have interpretations as tail risk measures.

The only assumption we make is the availability of a valid confidence interval for the tail risk measure of interest. Estimation and inference of risk measures are a widely studied area, ranging from parametric to nonparametric methods. Embrechts et al. (1997) consider VaR estimation using extreme-value theory; Barone-Adesi et al. (1999) give filtered historical simulation to estimate VaR; Chen and Tang (2005) and Scaillet (2004) provide nonparametric estimators of VaR and ES; Chen (2008) discuss performances of two major nonparametric estimation methods; Linton and Xiao (2013) and Hill (2015) consider nonparametric ES estimation when the portfolio return may not have variance. Belomestny and Krätschmer (2012) derive asymptotic normality for plugin estimators of law-invariant coherent risk measures; Gao and Song (2008) establish asymptotic distribution of VaR and ES estimates. There is also large literature on estimation of VaR and ES *conditional on covariates* (Chernozhukov and Umantsev, 2001; Cai and Wang, 2008; Kato, 2012; Chernozhukov and Fernández-Val, 2011; Chun et al., 2012; Chernozhukov et al., 2017; Martins-Filho et al., 2018).

In the empirical application, we apply our method to control VaR and ES in a portfolio

[4] Note that there are two distinct concepts of "confidence levels:" one associated with the risk measure and the other with statistical inference thereof.

of three assets. We do not assume any portfolio optimization problem; instead, we directly derive a bound on VaR and ES of level 95%; hence, the insights are directly applicable to other problems such as banking regulation. In particular, we calculate the bounds on VaR and ES for the stocks of the Bank of America Corp. and Morgan Stanley and the index fund of the Dow Jones Industrial Averages. The risk of the entire portfolio is then bounded by the convex combination of the individual risks when the risk measure is subadditive. Our exercise reveals that even if we use a relatively small sample size of approximately one year ($T = 293$), our bounds on VaR are at most 60% bigger than the direct estimates of VaR of 95%, and our bounds on ES are at most 70% bigger than the direct estimates of ES of 95%. Thus, despite the use of the Bonferroni inequality, our method does not suggest extremely conservative risk management.

This paper is organized as follows. Section 2 demonstrates the problem of estimation error in risk control by simulation. Section 3 introduces the notion of a tail risk measure and examples thereof. Section 3.2 provides a method to control the true risk probability using the Bonferroni inequality. Section 4 illustrates the use of our method in the risk control of VaR and ES, using data from Yahoo! Finance.

2 Distortion of Risk Probability with Estimated Risk

The defining feature of VaR is that the probability of the loss $-X$ exceeding VaR is bounded by a desired level, $\Pr(-X > \mathrm{VaR}_p) \leq p$. However, if we substitute VaR_p with an estimator $\widehat{\mathrm{VaR}}_p$, the probability of the loss exceeding it is no longer bounded by p. We simulate a GARCH model to investigate the relationship between the actual risk probability $\Pr(-X > \widehat{\mathrm{VaR}}_p)$ and the intended risk probability p.

Our simulation consists of two iterations. In the first iteration, we draw $T = 200$ observations $\{X_1, \ldots, X_T\}$ from a GARCH(1,1) model:

$$\begin{cases} X_t = \sigma_t z_t, \\ \sigma_t^2 = \omega + \alpha z_{t-1}^2 + \beta \sigma_{t-1}^2, \end{cases}$$

where $\omega = 0.001$, $\alpha = 0.05$, $\beta = 0.9$, and $z_t \sim N(0, 1)$ i.i.d. We first estimate the GARCH parameters (ω, α, β) by maximum likelihood estimation (MLE) and fit innovations $\{\hat{z}_1, \ldots, \hat{z}_T\}$

and volatility $\hat{\sigma}_T$. We then estimate VaR_p for $p \in \{1\%, 5\%\}$ by $\widehat{\mathrm{VaR}}_p = -\hat{\sigma}_T \kappa_p$ where κ_p is given by 3 methods:

1 (Parametric). The p-quantile of a standard normal distribution.

2 (Semiparametric). Weissman's (1978) estimator of the p-quantile of $\{\hat{z}_1, \ldots, \hat{z}_T\}$:

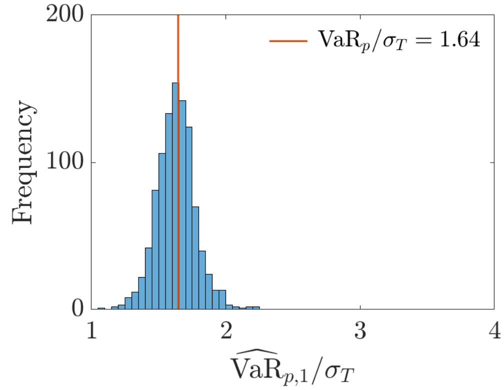
$$\kappa_p = \hat{z}_{(k)} - \left(\hat{z}_{(k)} - \frac{1}{k} \sum_{i=1}^{k} \hat{z}_{(i)} \right) \log \left(\frac{k}{pT} \right) \qquad \text{for } k = 10.$$

3 (Nonparametric). The empirical p-quantile of $\{\hat{z}_1, \ldots, \hat{z}_T\}$.

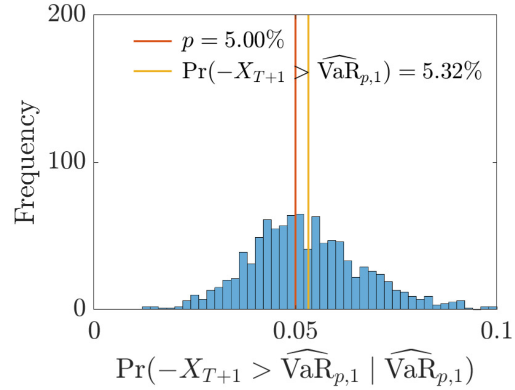
Since $z_T \sim N(0,1)$, we know the exact probability that $-X_T$ exceeds $\widehat{\mathrm{VaR}}_p$ conditional on the estimate, $\Pr(-X_{T+1} > \widehat{\mathrm{VaR}}_p \mid \widehat{\mathrm{VaR}}_p)$. In the second iteration, we repeat the first iteration for $S = 1{,}000$ times and compute the unconditional probability with which the new draw falls below the estimator, $\Pr(-X_{T+1} > \widehat{\mathrm{VaR}}_p)$. We see below that better estimation of VaR does not necessarily imply less distortion of risk probability.

Figures 1 and 2 are histograms of the estimated VaR and the risk probability for confidence levels 5% and 1%. Figure 1a shows the histogram of $\widehat{\mathrm{VaR}}_p$ scaled by the true volatility σ_T. Note that the variance of X_T differs in each iteration. If scaled by the true volatility, X_T/σ_T follows a standard normal. Therefore, $\widehat{\mathrm{VaR}}_p/\sigma_T$ is comparable across iterations. The red line indicates VaR_p/σ_T, which is simply the negative of a p-quantile of a standard normal distribution. Figure 1b is the histogram of the risk probability conditional on the estimate, $\Pr(-X_{T+1} > \widehat{\mathrm{VaR}}_p \mid \widehat{\mathrm{VaR}}_p)$. The red line indicates the intended risk probability p, and the orange line the actual risk probability $\Pr(-X_{T+1} > \widehat{\mathrm{VaR}}_p)$. Figure 1b suggests that if we maintain liquid assets worth $\widehat{\mathrm{VaR}}_p$, then we go insolvent 5.32% of the time, more often than the intended 5%. Figures 1c and 1d are the corresponding figures for the semiparametric estimator; Figures 1e and 1f for the nonparametric.

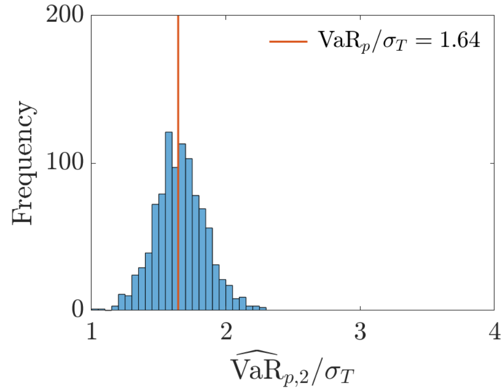
Being correctly specified, the parametric estimator achieves the smallest mean squared error (MSE) (Figures 1a, 1c, and 1e); however, its risk probability distortion $5.32\% - 5.00\% = 0.32\%$ is worse than that of the semiparametric estimator 0.15% (Figures 1b and 1d). Therefore, higher precision of VaR estimation does not necessarily imply less distortion of the risk probability. The distortion rates are worse for VaR for $p = 1\%$; for example, the semiparametric estimator distorts 0.53% when the intended probability is 1% (Figure 2d).



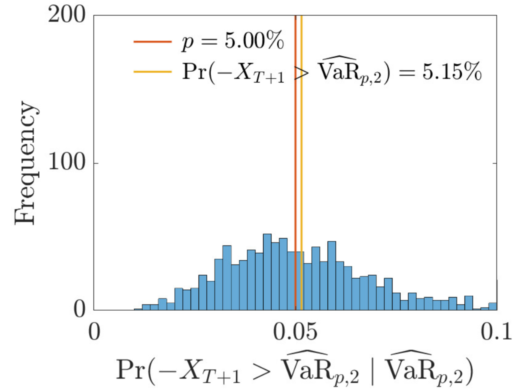
(a) The parametric estimator for VaR with $p = 5\%$. MSE $= 0.0207$.



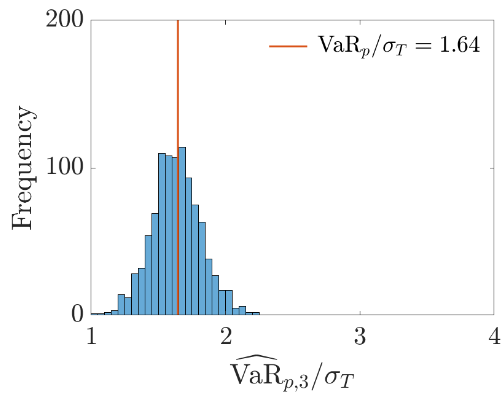
(b) The risk probability for the parametric estimator.



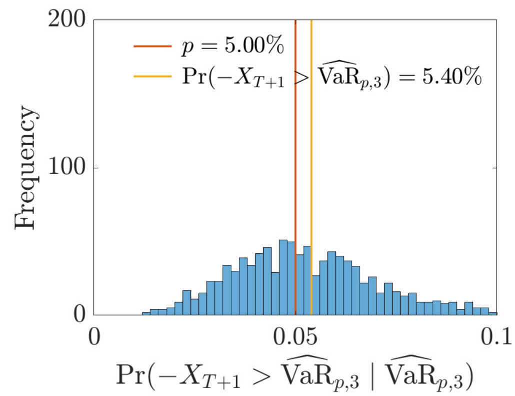
(c) The semiparametric estimator for VaR with $p = 5\%$. MSE $= 0.0355$.



(d) The risk probability for the semiparametric estimator.

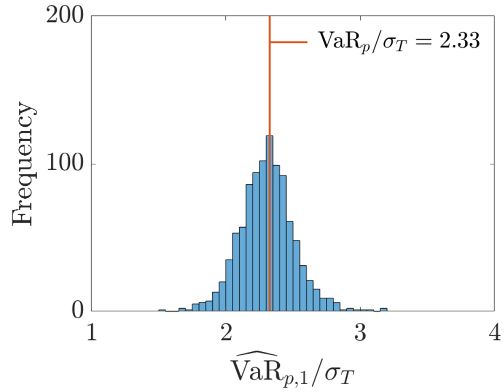


(e) The nonparametric estimator for VaR with $p = 5\%$. MSE $= 0.0339$.

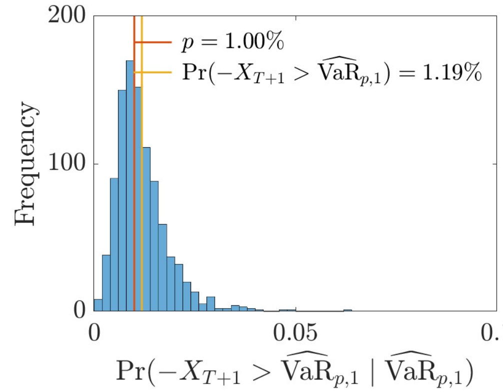


(f) The risk probability for the nonparametric estimator.

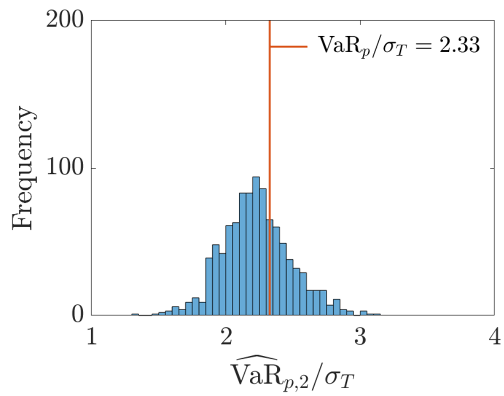
Figure 1: Simulation of GARCH(1,1) with $T = 200$ and $S = 1{,}000$ for $\mathrm{VaR}_{5\%}$.



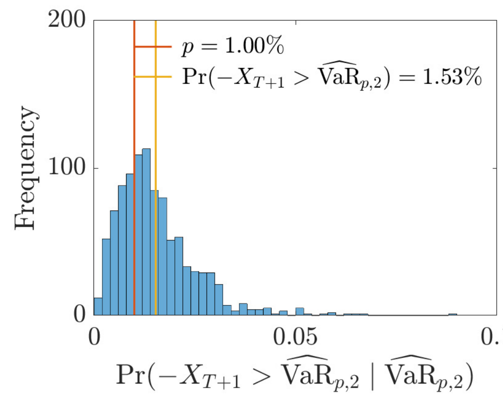
(a) The parametric estimator for VaR with $p = 1\%$. MSE $= 0.0414$.



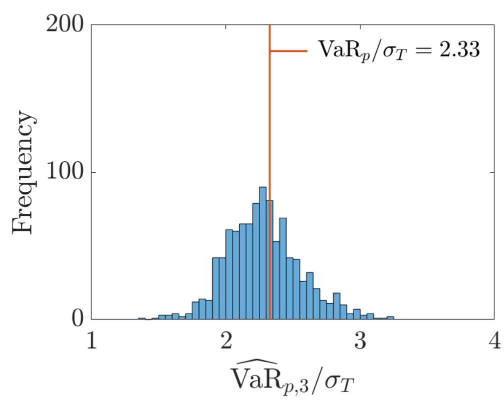
(b) The risk probability for the parametric estimator.



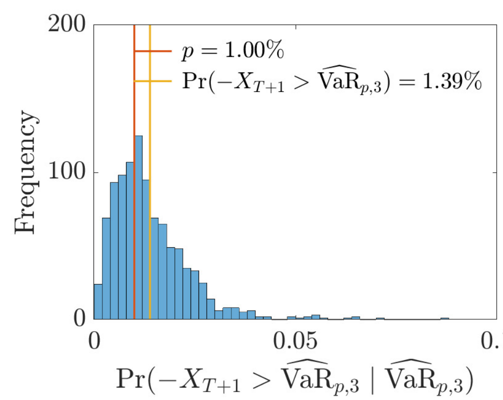
(c) The semiparametric estimator for VaR with $p = 1\%$. MSE $= 0.0749$.



(d) The risk probability for the semiparametric estimator.



(e) The nonparametric estimator for VaR with $p = 1\%$. MSE $= 0.0786$.



(f) The risk probability for the nonparametric estimator.

Figure 2: Simulation of GARCH(1,1) with $T = 200$ and $S = 1{,}000$ for $\mathrm{VaR}_{1\%}$.

3 Tail Risk Measures

3.1 Definition

In this section, we define the tail risk measure and give examples. Notably, we show that VaR and ES are tail risk measures. Let $(\Omega, \mathcal{F}, P)$ be a probability space that is rich enough to contain all random variables and events introduced below. Let $X : \Omega \to \mathbb{R}$ be a random variable representing the return of a portfolio or the profits of a company. Denote by F_X and Q_X the distribution and quantile functions of X. We set U to be the uniformly distributed random variable on $(0, 1)$ that satisfies $Q_X(U) = X$ almost surely. For a random variable $Y : \Omega \to \mathbb{R}$, denote by $\sigma(Y) \subset \mathcal{F}$ the σ-algebra generated by Y. For an event $E \in \mathcal{F}$, denote by $F_{X|E}$ the conditional distribution function of X conditional on the occurrence of E.

The tail risk measure is a quantity such that the probability of some characteristic of X exceeding it is bounded. The characteristic is given by a function χ that maps a distribution to a real number. The bound on the probability is denoted by p and we call $1 - p$ the *confidence level*.

Definition. For $p \in (0, 1)$ and a map $F_{X|E} \mapsto \chi(F_{X|E}) \in \mathbb{R}$, the *tail risk measure of confidence level* $1 - p$ is the infimum of c_p that satisfies

$$\sup_{E \in \sigma_0} \big\{ \Pr(\omega \in E) : \chi(F_{X|E}) \geq c_p \big\} \leq p$$

for some fixed $\sigma_0 \subset \sigma(U)$. If the supremum is attained at $E = \varnothing$, the infimum of c_p is defined to be ∞.

Intuitively, the tail risk measure controls the probability of a bad event χ. The supremum with respect to E represents a search for the worst event over a set of events σ_0; χ represents the measure of badness of the distribution $F_{X|E}$ given E; c_p gives the threshold above which χ will not exceed with probability $1 - p$. A regulator of the banking system may want to bound the probability of financial crisis at some level. A portfolio manager may want to minimize the probability of the portfolio incurring a significant loss. This notion of bounding the probability is the key idea that the tail risk measure captures.

Note that the function χ does not depend on E or p, and the supremum is taken over a subset of the σ-algebra generated by U, not $\sigma(X)$ or $\mathcal{F}$. In general, risk measures are defined as functions that map random variables to real numbers; dependence on E gives too much freedom to qualify as a risk measure. Also, we want p to represent some probability of a fixed concept, so not allowing χ to depend on p is natural as well. The point on the supremum is explained at the end of this section.

First, we show that popular risk measures are tail risk measures. We illustrate how VaR, ES, and some distortion risk measures are written as tail risk measures. Then, we explain why variance cannot (and should not) be interpreted as a tail risk measure.

Example 1 (Value-at-Risk). For $X : \Omega \to \mathbb{R}$, *VaR of confidence level* $1-p$ is defined as

$$\mathrm{VaR}_p(X) := \inf\{x \in \mathbb{R} : \Pr(X < -x) \leq p\}.$$

In words, VaR gives the upper bound on the loss that holds with probability at least $1-p$, or equivalently, the worst loss one can expect with probability $1-p$ or higher. This quantity can be cast as the infimum of c_p that satisfies

$$\sup_{E \in \sigma(U)} \left\{\Pr(\omega \in E) : -\sup \mathrm{supp}(F_{X|E}) \geq c_p\right\} \leq p,$$

where $\mathrm{supp}(F_{X|E})$ denotes the support of $F_{X|E}$. Therefore, VaR of level $1-p$ is a tail risk measure of level $1-p$ with $\sigma_0 = \sigma(U)$ and $\chi(F_{X|E}) = -\sup \mathrm{supp}(F_{X|E})$. $\square$

Example 2 (Expected shortfall). For $X : \Omega \to \mathbb{R}$, *ES of confidence level* $1-p$ is defined as (Acerbi et al., 2001)

$$\mathrm{ES}_p(X) := \frac{1}{p} \int_0^p \mathrm{VaR}_\gamma(X) d\gamma.$$

Note that VaR is (the negative of) the quantile of X. Thus, for continuous X, ES is equal to $\mathbb{E}[-X \mid X \leq -\mathrm{VaR}_p]$, the *expected* loss in the worst event with probability p. ES is known to be coherent (Acerbi and Tasche, 2002). It can be cast as the infimum of c_p that satisfies

$$\sup_{E \in \sigma(U)} \left\{\Pr(\omega \in E) : \mathbb{E}[-X \mid \omega \in E] \geq c_p\right\} \leq p.$$

Thus, ES of level $1-p$ is a tail risk measure of level $1-p$ with $\sigma_0 = \sigma(U)$ and $\chi(F_{X|E}) = \mathbb{E}[-X \mid E] = \int -x dF_{X|E}$. $\square$

Remark. If the distribution of X is discontinuous at p, that is, there does not exist c such that $\Pr(X \leq c) = p$, then the event $U \in (0, \alpha)$ is not in $\sigma(X)$. For this reason, we use $\sigma(U)$ instead of $\sigma(X)$ in the definition.

Example 3 (Tail Value-at-Risk)**.** A quantity that is closely related to expected shortfall is tail Value-at-Risk (TVaR). For $X : \Omega \to \mathbb{R}$, *TVaR of confidence level* $1 - p$ is given by

$$\mathrm{TVaR}_p(X) := \mathbb{E}[-X \mid -X \geq \mathrm{VaR}_p(X)].$$

As mentioned above, this coincides with expected shortfall if X is continuous at $\mathrm{VaR}_p(X)$. TVaR is not coherent (Acerbi, 2004). This is cast as the infimum of c_p that satisfies

$$\sup_{E \in \sigma(X)} \bigl\{\Pr(\omega \in E) : \mathbb{E}[-X \mid \omega \in E] \geq c_p\bigr\} \leq p.$$

Note that the only difference from ES is how the supremum is taken; TVaR takes supremum over $\sigma(X)$ while ES over $\sigma(U)$. Thus, TVaR of level $1-p$ is a tail risk measure of level $1-p$ with $\sigma_0 = \sigma(X)$ and $\chi(F_{X|E}) = \mathbb{E}[-X \mid E]$. $\square$

Example 4 (Distortion risk measure)**.** Let $K : [0, 1] \to [0, 1]$ be nondecreasing with $K(0) = 0$ and $K(1) = 1$. The *distortion risk measure* for the distortion function K is defined as

$$\rho_K(X) := \int_0^1 \mathrm{VaR}_\gamma(X) dK(\gamma) = \int_0^1 -Q_X dK.$$

Let us focus on distortion functions that satisfy an additional assumption: there exists $p \in (0, 1)$ such that $K(u) < 1$ for $u < p$ and $K(u) = 1$ for $u \geq p$. VaR is a distortion risk measure with $K(u) = \mathbb{1}\{u \geq p\}$; ES is a distortion risk measure with $K(u) = \max\{u/p, 1\}$; thus, both are distortion risk measures that satisfy this assumption. With this, we may write $\rho_K(X)$ as the infimum of c_p such that

$$\sup_{E \in \sigma(U)} \left\{\Pr(\omega \in E) : \int_0^1 -Q_{X|E}(u) dK\left(\frac{u}{p}\right) \geq c_p\right\} \leq p,$$

where $Q_{X|E}$ denotes the generalized inverse function of $F_{X|E}$. $\square$

To motivate why the supremum is taken over (a subset of) $\sigma(U)$, let us look at risk measures that are not tail risk measures.

Example 5 (Variance)**.** For $X : \Omega \to \mathbb{R}$ with a second moment, *variance* is defined by

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}[X])^2].$$

Variance is not a feature of X conditional on an event, but a feature that captures how spread the entire distribution is. As such, it cannot be represented as a characteristic of X conditional on an event in $\sigma(U)$, which restricts how X distributes; hence, it is not a tail risk measure. On the other hand, if we allow the supremum to be taken over a more general sub-σ-algebra, we can represent variance in a similar form. Assume that $\mathcal{F}$ is rich enough that it contains a random variable V independent of U. Then, variance is the infimum of c_p such that

$$\sup_{E \in \sigma(V)} \left\{ \Pr(\omega \in E) : \mathbb{E}[(X - \mathbb{E}[X \mid E])^2 \mid E] \geq c_p \right\} \leq p$$

for arbitrary $p \in [0, 1]$, since $\mathrm{Var}(X \mid E) = \mathrm{Var}(X)$. However, it is not sensible to regard a decision maker who aims to control variance as wanting to control the conditional variance on any event at confidence level 0. Put differently, the event that attains the supremum is the one in which $\Pr(V \in V(E)) = p$ and hence the conditional distribution of X is no different from its marginal. Therefore, considering such an event as the "worst" case is pointless. □

Example 6 (Entropic Value-at-Risk)**.** For $X : \Omega \to \mathbb{R}$ with a finite moment-generating function $M_X(z) := \mathbb{E}[e^{zX}]$ for $z \leq 0$, Ahmadi-Javid (2012) considers *entropic Value-at-Risk (EVaR)* of confidence level $1 - p$ by[5]

$$\mathrm{EVaR}_p(X) := \inf_{z<0} \frac{\log(M_X(z)/p)}{-z}.$$

Since M_X depends on the entire distribution of X, it is not a tail risk measure. Similarly as variance, if we allow a richer σ-algebra, EVaR can be written as the infimum of c_α such that

$$\sup_{E \in \sigma(V)} \left\{ \Pr(\omega \in E) : \inf_{z<0} \frac{\log(M_{X|E}(z)/p)}{-z} \geq c_p \right\}$$

for arbitrary $p \in [0, 1]$. However, this does not capture any tail event regarding X. □

[5]Note that we measure the risk of the negative tail of X while Ahmadi-Javid (2012) measures that of the positive, hence the sign difference.

3.2 Controlling the Tail Risk Measure

This section provides a method to control the tail risk measures with an observable quantity. In particular, we assume that a one-sided confidence interval for a tail risk measure c_q is available,

$$\Pr(c_q < \tilde{c}_{q,r}) \geq 1 - r$$

for observable $\bar{c}_{q,r}$. Inuitively, this gives rise to the Bonferroni bound

$$\Pr(\chi(F_{X|E}) \geq \tilde{c}_{q,r}) \leq \Pr(\chi(F_{X|E}) \geq c_q) + \Pr(c_q \geq \tilde{c}_{q,r}) \leq q + r.$$

Thus, we may control the tail risk measure of confidence level $1 - p$ by setting $q + r \leq p$.

Theorem 1. *Let c_q be the true tail risk measure of confidence level $1 - q$, and $\tilde{c}_{q,r}$ the $(1 - r)$-confidence bound of c_q, that is, $\Pr(\tilde{c}_{q,r} \leq c_q) \leq r$. Then, we have*

$$\sup_{E \in \sigma_0} \Pr\big(\omega \in E \wedge \chi(F_{X|E}) \geq \tilde{c}_{q,r}\big) \leq q + r.$$

In other words, $\tilde{c}_{q,r}$ controls c_{q+r}.

Proof. By the Bonferroni inequality and the properties of c_q and $\tilde{c}_{q,r}$,

$$\sup_{E \in \sigma_0} \Pr\big(\omega \in E \wedge \chi(F_{X|E}) \geq \tilde{c}_{q,r}\big) \leq \sup_{E \in \sigma_0} \big\{\Pr(\omega \in E) : \chi(F_{X|E}) \geq c_q\big\} + \Pr(\tilde{c}_{q,r} \leq c_q) \leq q + r.$$

■

Remark. Although the theorem is valid for any fixed choice of (q, r), we cannot "hunt" for the smallest $\tilde{c}_{q,r}$. This changes the distribution of $\tilde{c}_{q,r}$ (due to randomness introduced by hunting), and the probability bound is no longer valid.

The confidence interval is assumed to be one-sided but need not be so. The upper bound of a two-side confidence interval by construction works as a bound for the one-sided confidence interval. It becomes, however, more conservative than necessary. Also, our results only assumes availability of a confidence interval but not of an estimator. We may use a large body of the literature that concern inference of risk measures (Gao and Song, 2008; Chen, 2008; Linton and Xiao, 2013; Hill, 2015; Belomestny and Krätschmer, 2012).

3.3 Practical Choices of (q, r)

Given a valid confidence interval, our method gives a bound on the risk probability with a finite sample guarantee. As the sample size increases, however, we should spend less and less on the construction of the confidence interval. Suppose that c_q admits an asymptotically normal estimator such that

$$\sqrt{T}(\hat{c}_q - c_q) \rightsquigarrow N(0, \sigma_q^2)$$

for some σ_q^2. The one-sided $(1-r)$-confidence bound for c_q is of the form

$$\hat{c}_q + \frac{\sigma_q}{\sqrt{T}}\kappa_r$$

for the corresponding $(1-r)$-confidence bound κ_r of the standard normal distribution. Thus, the confidence bound shrinks at rate $T^{-1/2}$. On the other hand, κ_r blows up as $r \to 0$.

In most practical situations, we can assume that $\hat{c}_q$ and σ_q are continuous at $q = p$, so the rate at which r shrinks must make sure that $\kappa_r/\sqrt{T}$ goes to zero. Let ϕ and Φ be the pdf and cdf of a standard normal. Applying L'ôpital's rule to Φ/ϕ reveals that $\Phi(-\kappa_r) \approx \phi(\kappa_r)/\kappa_r$ for large κ_r. Using $\phi(\kappa_r)/e^{\kappa_r} \lesssim \phi(\kappa_r)/\kappa_r \lesssim \phi(\kappa_r)$, we obtain

$$\sqrt{1 - 2\log(\sqrt{2\pi}r)} - 1 \lesssim \kappa_r \lesssim \sqrt{-2\log(\sqrt{2\pi}r)}.$$

In other words, we can let $r \to 0$ at the speed such that $\sqrt{-2\log r}/\sqrt{T} \to 0$, or $r \gg e^{-T/2}$. Therefore, r can indeed go very fast to zero; e.g., any rational function in T converges slower than $e^{-T/2}$. For the sample size of order hundrets to thousands, $r = 1/T$ would be a good choice. A general recommendation is to plot $\kappa_r/\sqrt{T}$ for given T across r and pick a small r for which $\kappa_r/\sqrt{T}$ is small enough. This does not constitute p-hacking as no data is involved in the decision.

4 An Empirical Application to VaR and ES

Consider the problem of controlling VaR and ES of a portfolio with three assets: the stock of the Bank of America Corp. (BAC), the stock of Morgan Stanley (MS), and the index fund of the Dow Jones Industrial Averages (DJI). The choice of assets is due to the ease of

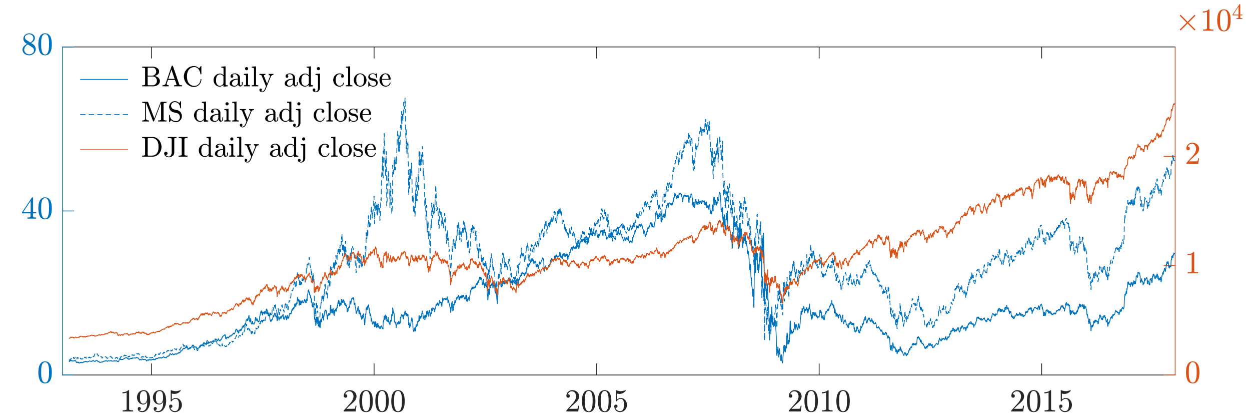


Figure 3: The daily adjusted close values of the Bank of America Corp., Morgan Stanley, and the Dow Jones Industrial Averages from Feb. 23, 1993 to Dec. 31, 2017.

access to the data, but it can be any assets, e.g., on a bank's balance sheet. We use the daily adjusted close values from February 23, 1993 to December 31, 2017. The price data are retrieved from Yahoo! Finance. Figure 3 shows the adjusted close values of these assets.

Denoting by Y_t the daily close value, the daily return is calculated as $X_t = (Y_t - Y_{t-1})/Y_{t-1}$. Figure 4 shows the daily returns of the three assets. The daily portfolio return is given by

$$X = w_1 X_{\mathrm{BAC}} + w_2 X_{\mathrm{MS}} + w_3 X_{\mathrm{DJI}},$$

where X_{BAC} is the return of the Bank of America, X_{MS} of Morgan Stanley, X_{DJI} of the Dow Jones, and (w_1, w_2, w_3) are weights of each asset. We want to obtain bounds on VaR and ES of level $1-p$ of the portfolio return X.

We model each daily return as a GARCH(1,1) process:

$$\begin{cases} X_t = \mu + \sigma_t z_t, \\ \sigma_t^2 = \omega + \alpha z_{t-1}^2 + \beta \sigma_{t-1}^2, \end{cases}$$

where $\{z_t\}$ are i.i.d. random variables and $\omega, \alpha, \beta \geq 0$ are GARCH parameters.[6] We use Gao and Song (2008) to estimate VaR and ES of individual assets: For $r = 1/T$ and $q = p - 3r$,

1. Estimate the GARCH parameters $(\mu, \omega, \alpha, \beta)$ by quasi-maximum likelihood estimation

[6] X_t is stationarity if $\alpha + \beta < 1$ and i.i.d. if $\alpha = \beta = 0$.

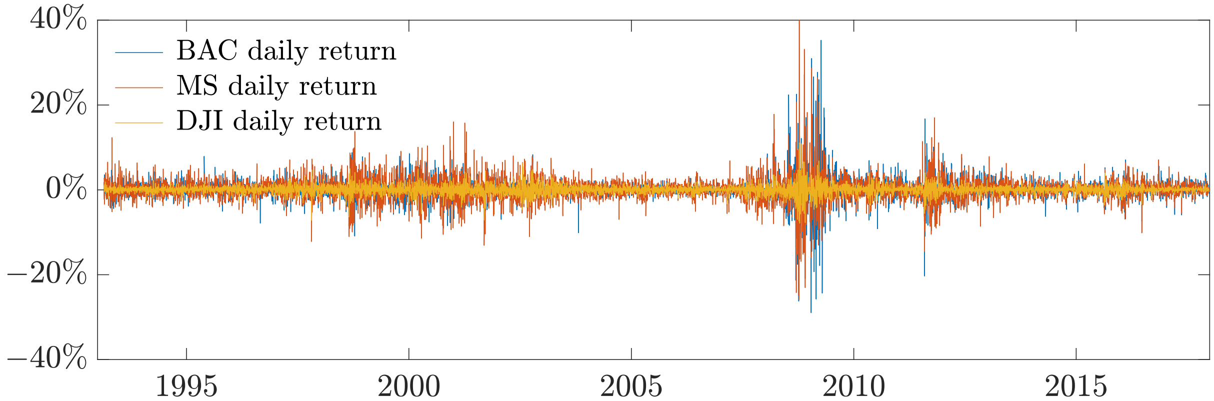


Figure 4: The daily returns of the Bank of America Corp., Morgan Stanley, and the Dow Jones Industrial Averages from Feb. 23, 1993 to Dec. 31, 2017.

(QMLE) assuming normality of z_t, and fit $\hat{\sigma}_T$ and $\{\hat{z}_1, \ldots, \hat{z}_{t-1}\}$.

2. Let κ_q be the empirical q-quantile of $\hat{z}_t$ and compute the q-trimmed average of $\hat{z}_t$ by $\eta_q = \sum_{t=1}^{T} \hat{z}_t \mathbb{1}\{\hat{z}_t < \kappa_q\} / \sum_{t=1}^{T} \mathbb{1}\{\hat{z}_t < \kappa_q\}$.

3. Estimate VaR_q and ES_q by $\widehat{\mathrm{VaR}}_q = \hat{\sigma}_T \kappa_q$ and $\widehat{\mathrm{ES}}_q = \hat{\sigma}_T \eta_q$.

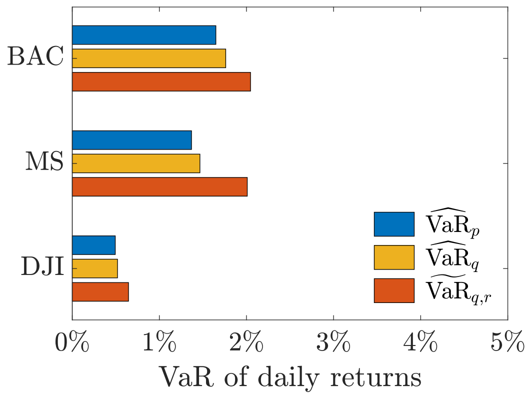
4. Estimate $\mathrm{Var}(\widehat{\mathrm{VaR}}_q)$ and $\mathrm{Var}(\widehat{\mathrm{ES}}_q)$ using analytical formulae provided in Gao and Song (2008, Theorems 3.2 and 3.3).

5. Construct the one-sided $(1-r)$-confidence bounds $\widetilde{\mathrm{VaR}}_{q,r}$ and $\widetilde{\mathrm{ES}}_{q,r}$.

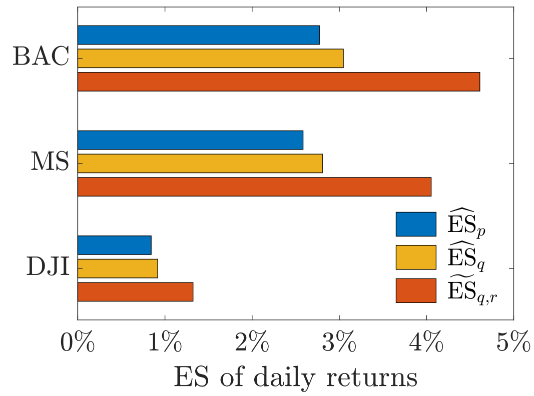
Since the formulae in Gao and Song (2008, Theorems 3.2 and 3.3) are of marginal distributions, we allot probability $1/T$ to each confidence bound, hence $q = p - 3r$; if there are many assets, this may lead to a very conservative bound. Instead, we may use the bootstrap procedure in Gao and Song (2008) and estimate the joint distribution of VaR and ES estimates to avoid overconservatism.[7]

Figure 5 shows the estimates of VaR and ES, which are also summarized in Table 1. Figures 5a and 5b use the data from November 1, 2016 to December 31, 2017, consisting of $T = 293$ daily returns; Figures 5c and 5d from February 23, 1993 to December 31, 2017, totalling $T = 6{,}260$ observations. The blue bars in Figure 5a show the direct estimates of

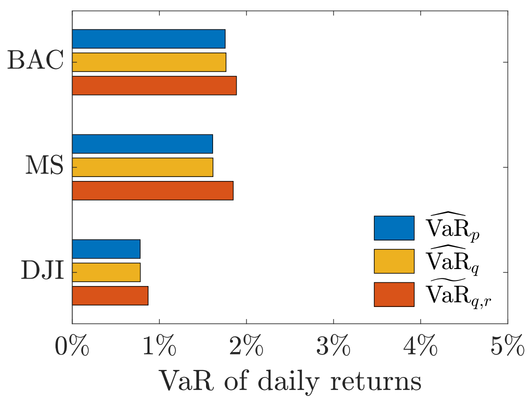
[7]To the best of our knowledge, however, the validity of bootstrap in this context is not fully understood in the literature.



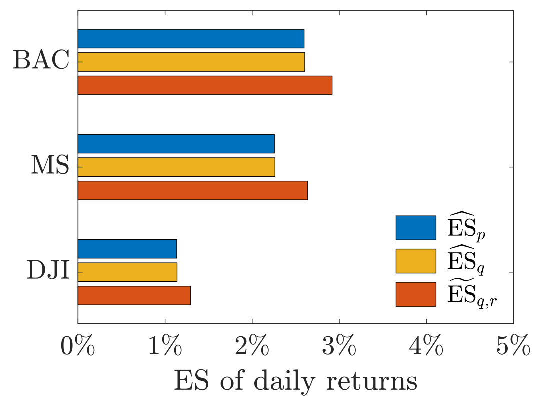
(a) VaR with $(p, q, r) = (5\%, 3.98\%, 0.34\%)$ and $T = 293$.



(b) ES with $(p, q, r) = (5\%, 3.98\%, 0.34\%)$ and $T = 293$.



(c) VaR with $(p, q, r) = (5\%, 4.95\%, 0.016\%)$ and $T = 6{,}260$.



(d) ES with $(p, q, r) = (5\%, 4.95\%, 0.016\%)$ and $T = 6{,}260$.

Figure 5: VaR and ES of level 95% for the daily returns with $(p, q, r) = (5\%, p - 3r, 1/T)$. The top two use the data from Nov. 1, 2016 to Dec. 31, 2017 ($T = 293$); the bottom two from Feb. 23, 1993 to Dec. 31, 2017 ($T = 6{,}260$).

VaR of level 95%, which have no guarantee on the risk probability. The orange bars are the interim estimates of VaR of level $q = p - 3r = 3.98\%$. The red bars are the upper bounds of the one-sided $(1 - r)$-confidence intervals of VaR_q for $r = 1/T = 0.34\%$; they satisfy the promise of the risk probability provided that the confidence intervals are valid.

The numbers in parentheses in Table 1 are "multipliers" relative to the standard estimates; in columns (1–3) they are ratios relative to column (1), and in columns (4–6)

Table 1: Estimates of VaR and ES of level 95%. The numbers in parentheses are the multipliers relative to $\widehat{\text{VaR}}_p$ and $\widehat{\text{ES}}_p$.

	$\widehat{\text{VaR}}_p$ (1)	$\widehat{\text{VaR}}_q$ (2)	$\widetilde{\text{VaR}}_{q,r}$ (3)	$\widehat{\text{ES}}_p$ (4)	$\widehat{\text{ES}}_q$ (5)	$\widetilde{\text{ES}}_{q,r}$ (6)
Nov. 1, 2016 to Dec. 31, 2017 ($T = 293$)						
BAC	1.65%	1.76%	2.05%	2.77%	3.05%	4.61%
	(1.000)	(1.069)	(1.241)	(1.000)	(1.099)	(1.664)
MS	1.37%	1.47%	2.01%	2.58%	2.81%	4.05%
	(1.000)	(1.071)	(1.468)	(1.000)	(1.086)	(1.569)
DJI	0.49%	0.52%	0.65%	0.84%	0.92%	1.32%
	(1.000)	(1.052)	(1.308)	(1.000)	(1.089)	(1.570)
Feb. 24, 1993 to Dec. 31, 2017 ($T = 6{,}260$)						
BAC	1.76%	1.77%	1.89%	2.60%	2.60%	2.92%
	(1.000)	(1.005)	(1.074)	(1.000)	(1.003)	(1.124)
MS	1.61%	1.62%	1.85%	2.25%	2.26%	2.63%
	(1.000)	(1.002)	(1.146)	(1.000)	(1.003)	(1.168)
DJI	0.78%	0.78%	0.87%	1.13%	1.14%	1.29%
	(1.000)	(1.002)	(1.117)	(1.000)	(1.003)	(1.138)

ratios relative to column (4); for example, in column (3), $\widetilde{\text{VaR}}_{q,r}/\widehat{\text{VaR}}_p$, and in column (6), $\widetilde{\text{ES}}_{q,r}/\widehat{\text{ES}}_p$. VaR has overall very low multipliers, meaning that estimation error can be controlled without having to worry about being too conservative. Although ES has bigger multipliers, the magnitudes are still comparable to the multipliers needed for model risk (3 for VaR and 1.5 for ES). Whenever we can construct a confidence interval, we advocate the use of our method since the multipliers heavily depend on the sample size; however, when statistical inference is not available or infeasible, using a fixed multiplier of 2 may be a practically reasonable choice for estimation error of VaR and ES.

In order to control VaR of level 95%, we need to identify the value which the loss exceeds with probability 5% or less. Assuming that VaR is subadditive (Ibragimov, 2005, p. 25; McNeil et al., 2005, Theorem 6.8), we can give one such value as

$$w_1 2.05\% + w_2 2.01\% + w_3 0.65\%,$$

where $(2.05\%, 2.01\%, 0.65\%)$ are the red bars from Figure 5a and (w_1, w_2, w_3) are weights of the portfolio on BAC, MS, and DJI. In other words, by holding a liquid asset worth this value, the probability of insolvency is capped by 5%. Note that even without the assumption of subadditivity, we can directly compute a bound on the risk of the entire portfolio by applying the same exercise to the transformed data $\{w_1 X_{\text{BAC}} + w_2 X_{\text{MS}} + w_3 X_{\text{DJI}}\}$; however, if we use subadditivity, it eliminates the need to reestimate the risks when we change the weights.

Since ES is subadditive by construction (McNeil et al., 2005, Proposition 6.9), we can control ES of level 95% by

$$w_1 4.61\% + w_2 4.05\% + w_3 1.32\%,$$

where the numbers come from the red bars in Figure 5b.

To allow for short positions (negative weights), observe that the risk associated with shorting X_t is equal to the risk of longing $-X_t$. Then, we can apply the same method with the risk estimates on the other tail.

Acknowledgements

We thank Anna Mikusheva, Isaiah Andrews, Victor Chernozhukov, and Deborah Lucas for helpful comments, and August Shen for research assistance.

Funding

This work is supported by the Richard N. Rosett Faculty Fellowship and the Liew Family Faculty Fellowship at the University of Chicago Booth School of Business.

Conflict of Interest

The authors have no conflict of interest to declare.

References

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

——— (2004): "Coherent Representations of Subjective Risk Aversion," in *Risk Measures for the 21st Century*, ed. by G. Szegö, New York: Wiley, 147–207.

Acerbi, C., C. Nordio, and C. Sirtori (2001): "Expected Shortfall as a Tool for Financial Risk Management," https://arxiv.org/abs/cond-mat/0102304.

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

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

Ahmadi-Javid, A. (2012): "Entropic Value-at-Risk: A New Coherent Risk Measure," *Journal of Optimization Theory and Applications*, 155, 1105–1123.

Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999): "Coherent Measures of Risk," *Mathematical Finance*, 9, 203–228.

Aussenegg, W. and T. Miazhynskaia (2006): "Uncertainty in Value-at-Risk Estimates under Parametric and Non-parametric Modeling," *Financial Markets and Portfolio Management*, 20, 243–264.

Barone-Adesi, G. and K. Giannopoulos (2001): "Non-parametric VaR Techniques. Myths and Realities," *Economic Notes*, 30, 167–181.

Barone-Adesi, G., K. Giannopoulos, and L. Vosper (1999): "VaR without Correlations for Portfolios of Derivative Securities," *Journal of Futures Markets*, 19, 583–602.

Belomestny, D. and V. Krätschmer (2012): "Central Limit Theorems for Law-Invariant Coherent Risk Measures," *Journal of Applied Probability*, 49, 1–21.

Berkowitz, J. and J. O'Brien (2002): "How Accurate Are Value-at-Risk Models at Commercial Banks?" *Journal of Finance*, 57, 1093–1111.

Caccioli, F., I. Kondor, and G. Papp (2017): "Portfolio Optimization under Expected Shortfall: Contour Maps of Estimation Error," *Quantitative Finance*, 0, 1–19.

Cai, Z. and X. Wang (2008): "Nonparametric Estimation of Conditional VaR and Expected Shortfall," *Journal of Econometrics*, 147, 120–130.

Ceria, S. and R. A. Stubbs (2006): "Incorporating Estimation Errors into Portfolio Selection: Robust Portfolio Construction," *Journal of Asset Management*, 7, 109–127.

Chen, J. M. (2014): "Measuring Market Risk under the Basel Accords: VaR, Stressed VaR, and Expected Shortfall," *IEB International Journal of Finance*, 184–201.

Chen, S. X. (2008): "Nonparametric Estimation of Expected Shortfall," *Journal of Financial Econometrics*, 6, 87–107.

Chen, S. X. and C. Y. Tang (2005): "Nonparametric Inference of Value-at-Risk for Dependent Financial Returns," *Journal of Financial Econometrics*, 3, 227–255.

Chernozhukov, V. and I. Fernández-Val (2011): "Inference for Extremal Conditional Quantile Models, with an Application to Market and Birthweight Risks," *Review of Economic Studies*, 78, 559–589.

Chernozhukov, V., I. Fernández-Val, and T. Kaji (2017): "Extremal Quantile Regression," in *Handbook of Quantile Regression*, ed. by R. Koenker, V. Chernozhukov, X. He, and L. Peng, Chapman and Hall/CRC, chap. 18, 333–362.

Chernozhukov, V. and L. Umantsev (2001): "Conditional Value-at-Risk: Aspects of Modeling and Estimation," *Empirical Economics*, 26, 271–292.

Chopra, V. K. and W. T. Ziemba (1993): "The Effect of Errors in Means, Variances, and Covariances on Optimal Portfolio Choice," *Journal of Portfolio Optimization*, 19, 6–11.

Chun, S. Y., A. Shapiro, and S. Uryasev (2012): "Conditional Value-at-Risk and Average Value-at-Risk: Estimation and Asymptotics," *Operations Research*, 60, 739–756.

Dowd, K. and D. Blake (2006): "After VaR: The Theory, Estimation, and Insurance Applications of Quantile-Based Risk," *Journal of Risk and Insurance*, 73, 193–229.

Embrechts, P., C. Klüppelberg, and T. Mikosch (1997): *Modelling Extremal Events for Insurance and Finance*, Berlin: Springer.

Gao, F. and F. Song (2008): "Estimation Risk in GARCH VaR and ES Estimates," *Econometric Theory*, 24, 1404–1424.

Gârleanu, N. and L. H. Pedersen (2007): "Liquidity and Risk Management," *American Economic Review*, 97, 193–197.

Hendricks, D. (1996): "Evaluation of Value-at-Risk Models Using Historical Data," *FRBNY Economic Policy Review*, 2, 39–70.

Hendricks, D. and B. Hirtle (1997): "Bank Capital Requirements for Market Risk: The Internal Models Approach," *FRBNY Economic Policy Review*, 1–12.

Hill, J. B. (2015): "Expected Shortfall Estimation and Gaussian Inference for Infinite Variance Time Series," *Journal of Financial Econometrics*, 13, 1–44.

Ibragimov, R. (2005): "New Majorization Theory In Economics And Martingale Convergence Results In Econometrics," Ph.D. thesis, Yale University.

Jorion, P. (1996): "Risk^2: Measuring the Risk in Value at Risk," *Financial Analysts Journal*, 52, 47–56.

Kaji, T. (2018): "Essays on Asymptotic Methods in Econometrics," Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Kato, K. (2012): "Weighted Nadaraya-Watson Estimation of Conditional Expected Shortfall," *Journal of Financial Econometrics*, 10, 265–291.

Leippold, M. and P. Vanini (2002): "Half as Many Cheers—The Multiplier Reviewed," *Wilmott Magazine*, 2, 104–107.

Linton, O. and Z. Xiao (2013): "Estimation of and Inference about the Expected Shortfall for Time Series with Infinite Variance," *Econometric Theory*, 29, 771–807.

Lopez, J. A. (1998): "Regulatory Evaluation of Value-at-Risk Models," *Journal of Risk*, 1, 37–63.

Markowitz, H. (1952): "Portfolio Selection," *Journal of Finance*, 7, 77–91.

Martins-Filho, C., F. Yao, and M. Torero (2018): "Nonparametric Estimation of Conditional Value-at-Risk and Expected Shortfall based on Extreme Value Theory," *Econometric Theory*, 34, 23–67.

McNeil, A. J., R. Frey, and P. Embrechts (2005): *Quantitative Risk Management: Concepts, Techniques and Tools*, Princeton: Princeton University Press.

Michaud, R. (1989): "The Markowitz Optimization Enigma: Is 'Optimized' Optimal?" *Financial Analysts Journal*, 45, 31–42.

Michaud, R. and R. Michaud (2008): "Estimation Error and Portfolio Optimization: A Resampling Solution," *Journal of Investment Management*, 6, 8–28.

Novak, S. Y. (2010): "A Remark Concerning Value-at-Risk," *International Journal of Theoretical and Applied Finance*, 13, 507–515.

Pritsker, M. (1997a): "Evaluating Value at Risk Methodologies: Accuracy versus Computational Time," *Journal of Financial Services Research*, 12, 201–242.

——— (1997b): "Towards Assessing the Magnitude of Value-at-Risk Errors due to Errors in the Correlation Matrix," *Financial Engineering News*, 14–16.

——— (2006): "The Hidden Dangers of Historical Simulation," *Journal of Banking & Finance*, 30, 561–582.

Rockafellar, R. T. and S. Uryasev (2000): "Optimization of Conditional Value-at-Risk," *Journal of Risk*, 2, 21–41.

Roy, A. D. (1952): "Safety First and the Holding of Assets," *Econometrica*, 20, 431–449.

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

Stahl, G. (1997): "Three Cheers," *Risk*, 10, 67–69.

Thiele, S. (2019): "Detecting Underestimates of Risk in VaR Models," *Journal of Banking and Finance*, 101, 12–20.

Wang, C.-S. and Z. Zhao (2016): "Conditional Value-at-Risk: Semiparametric Estimation and Inference," *Journal of Econometrics*, 195, 86–103.

Weissman, I. (1978): "Estimation of Parameters and Larger Quantiles Based on the k Largest Observations," *Journal of the American Statistical Association*, 73, 812–815.