

# CVaR sensitivity with respect to tail thickness

by Stoyan V. Stoyanov, Svetlozar T. Rachev,  
Frank J. Fabozzi

No. 29 | MAY 2011

WORKING PAPER SERIES IN ECONOMICS



## **Impressum**

Karlsruher Institut für Technologie (KIT)  
Fakultät für Wirtschaftswissenschaften  
Institut für Wirtschaftspolitik und Wirtschaftsforschung (IWW)  
Institut für Wirtschaftstheorie und Statistik (ETS)

Schlossbezirk 12  
76131 Karlsruhe

KIT – Universität des Landes Baden-Württemberg und  
nationales Forschungszentrum in der Helmholtz-Gemeinschaft

Working Paper Series in Economics  
**No. 29**, May 2011

ISSN 2190-9806

---

[econpapers.wiwi.kit.edu](http://econpapers.wiwi.kit.edu)

# CVaR sensitivity with respect to tail thickness

Stoyan V. Stoyanov

EDHEC Business School

EDHEC-Risk Institute–Asia

e-mail: stoyan.stoyanov@edhec-risk.com

Svetlozar T. Rachev\*

Department of Applied Mathematics & Statistics

Stony Brook University, USA

School of Economics and Business Engineering

Karlsruhe Institute of Technology, Germany, and

FinAnalytica USA

e-mail: rachev@ams.sunysb.edu

Frank J. Fabozzi

Yale School of Management

e-mail: frank.fabozzi@yale.edu

March 29, 2011

---

\*Dr Rachev is Frey Family Foundation Chair-Professor at Department of Applied Mathematics & Statistics, Stony Brook University, and Chief-Scientist, FinAnalytica. He gratefully acknowledges research support by grants from Division of Mathematical, Life and Physical Sciences, College of Letters and Science, University of California, Santa Barbara (where he is Professor Emeritus), the Deutschen Forschungsgemeinschaft and the Deutscher Akademischer Austausch Dienst.

# CVaR sensitivity with respect to tail thickness

## Abstract

We consider the sensitivity of conditional value-at-risk (CVaR) with respect to the tail index assuming regularly varying tails and exponential and faster-than-exponential tail decay for the return distribution. We compare it to the CVaR sensitivity with respect to the scale parameter for stable Paretian, the Student's  $t$ , and generalized Gaussian laws and discuss implications for the modeling of daily returns and marginal rebalancing decisions. Finally, we explore empirically the impact on the asymptotic variability of the CVaR estimator with daily returns which is a standard choice for the return frequency for risk estimation.

**keywords** fat-tailed distributions, regularly varying tails, conditional value-at-risk, marginal rebalancing, asymptotic variability

# 1 Introduction

There is substantial empirical evidence that financial returns exhibit fat-tails and excess kurtosis after accounting for the clustering of volatility and autocorrelation. Different models have been suggested to explain these empirical facts. Mandelbrot (1963) and Fama (1965) proposed the stable Paretian distribution which was later incorporated as a building block in GARCH-type processes, see for example Mittnik et al. (2002) and Mittnik and Paolella (2003).

Stable Paretian distributions are fat-tailed; they can explain the observed skewness and excess kurtosis of financial returns and represent a clear improvement over the Gaussian distribution, see Rachev and Mittnik (2000) for empirical studies and further references. Models based on Stable Paretian laws, however, do not explain the empirical fact that lower-frequency returns tend to have a higher tail index, see Samorodnitsky and Grabchak (2010). Stable distributions have an infinite variance and this creates technical difficulties in other areas such as option pricing and asset allocation modeling. Recently, tempered stable distributions have been suggested as an alternative model which has some of the attractive features of the class of stable laws but is more flexible, see Kim et al. (2008) and Kim et al. (2010) for additional details.

Apart from stable and tempered stable laws, other distributional models have been suggested in the literature and are used by practitioners. One alternative is the class of hyperbolic distributions which contains fat-tailed and skewed representatives and can be considered as a building block in financial models both in a discrete and continuous setting, see Bibby and Sorensen (2003). This class contains the Student's  $t$  distribution which is attractive to practitioners because of its simplicity.

Other examples include distributions used only for tail modeling without providing a basis for more general models. This category of models includes, for example, the approach based on extreme value theory. An application to value-at-risk modeling with a GARCH model for the clustering of volatility effect is available in Kuester et al. (2006). There are also more ad-hoc models such as the generalized normal distribution (GND). An application to high-frequency data is available in Chen et al. (2008). For additional examples with the Weibull, the Laplace-Gaussian mixture, and other distributions, see Haas et al. (2006) and Rachev and Mittnik (2000).

An important property distinguishing these classes of models is the rate of tail decay of the distribution. Stable Paretian distributions exhibit a power decay with a tail exponent  $\alpha$  in  $0 < \alpha < 2$ . Tempered stable laws and hyperbolic laws have a more involved exponential decay, while the GND can have an exponential and slower-than-exponential decay depending on the parameter values. Distinguishing statistically between different types of tail decay turns out to be a difficult problem. Heyde and Kou (2004)

report that 5,000 observations (20 years of daily data) are insufficient to distinguish between exponential-type and power-type tails. Heyde et al. (2006) conclude that downside risk measures based on tail expectations are not robust because of the sensitivity to the tail decay assumption. Outside of academic circles, the lack of robustness is interpreted as a sign of catastrophe – small changes in the tail parameter lead to huge changes in the tail risk measure.<sup>1</sup>

In this paper, our goal is to explore the implications of the lack of robustness for a popular measure of downside tail risk, the conditional value-at-risk (CVaR). The sensitivity of CVaR with respect to tail thickness is important for at least two practical problems – marginal rebalancing and stability with respect to estimation errors. We compare the relative importance of the tail thickness parameter to the scale parameter in the context of the power tail decay of the symmetric stable Paretian and the Student’s  $t$  distributions and the exponential and slower-than-exponential decay of the GND. The GND is chosen because of the varieties of tail decay (including a Gaussian tail) and because it leads to closed-form expressions for CVaR.

The paper is organized in the following way. We proceed with a description of the probability distributions and their CVaRs. In Section 3, we calculate an approximation of the the sensitivity of CVaR with respect to the tail index assuming regularly varying tails and GND-type tails. Finally, we discuss the impact of the tail decay assumption on the asymptotic variance of the CVaR estimator in Section 4 using daily returns. The choice of daily return frequency is motivated by the fact that this frequency is widely used for parameter estimation in risk management systems.

## 2 Fat-tailed models for asset returns

Mandelbrot (1963) strongly rejected normality as a distributional model for asset returns, conjecturing that financial returns behave like non-Gaussian stable returns. To distinguish between Gaussian and non-Gaussian stable distributions, the latter are commonly referred to as “stable Paretian” distributions or “Levy stable” distributions.

The class of stable distributions is defined by means of their characteristic functions. With very few exceptions, no closed-form expressions are known for their densities and cumulative distribution functions (c.d.f.). A random variable  $X$  is said to have a stable distribution,  $X \in S_\alpha(\beta, \sigma, \mu)$ , if its characteristic function  $\varphi_X(t) = Ee^{itX}$  has the following form

---

<sup>1</sup>The exact quote from Nassim Taleb’s blog at <http://sethkaufman.posterous.com/taleb-the-fourth-quadrant-a-map-of-the-limits> is: “Parametrizing a power law lends itself to monstrous estimation errors... Small changes in the “alpha” main parameter used by power laws leads to monstrously large effects in the tails.”

$$\varphi_X(t) = \begin{cases} \exp\{-\sigma^\alpha |t|^\alpha (1 - i\beta \frac{t}{|t|} \tan(\frac{\pi\alpha}{2})) + i\mu t\}, & \alpha \neq 1 \\ \exp\{-\sigma |t| (1 + i\beta \frac{2}{\pi} \frac{t}{|t|} \ln(|t|)) + i\mu t\}, & \alpha = 1 \end{cases} \quad (1)$$

where  $\frac{t}{|t|} = 0$  if  $t = 0$ .

The formula in (1) implies that stable laws are described by four parameters: (1)  $\alpha$ , called the index of stability, which determines the tail weight or density's kurtosis with  $0 < \alpha \leq 2$ , (2)  $\beta$ , called the skewness parameter, which determines the density's skewness with  $-1 \leq \beta \leq 1$ , (3)  $\sigma > 0$  which is a scale parameter, and (4)  $\mu \in \mathbb{R}$  which is a location parameter. Stable distributions allow for skewed distributions when  $\beta \neq 0$  and when  $\beta = 0$ , the distribution is symmetric around  $\mu$ . The Gaussian distribution is a stable distribution with  $\alpha = 2$ . Stable Paretian laws have fat tails, meaning that extreme events have high probability relative to a normal distribution when  $\alpha < 2$ . The power decay of the tail of non-Gaussian stable distributions is described by the following relation<sup>2</sup>

$$\lim_{\lambda \rightarrow \infty} \lambda^\alpha P(X \leq -\lambda) = \frac{\Gamma(\alpha) \sin(\pi\alpha/2)}{\pi} \left(\frac{1-\beta}{2}\right) \sigma^\alpha. \quad (2)$$

Of the four parameters,  $\alpha$  and  $\beta$  are most important as they identify two fundamental properties that are atypical of the normal distribution — heavy tails and asymmetry.

The Student's  $t$  distribution is very often used in practice for a very pragmatic reason – it is simple to work with. We consider the location-scale augmented version which is defined in the following way:  $X \in t(\nu, \sigma, \mu)$  if  $X = \sigma Y + \mu$ , where  $Y$  has the density

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad x \in \mathbb{R}. \quad (3)$$

where  $\nu > 0$  is the degrees of freedom parameter,  $\sigma > 0$ , and  $\mu \in \mathbb{R}$  are the scale and the location parameters, respectively. The density of the Student's  $t$  distribution decays like a power law. A more precise result concerning the decay of the tail of the distribution is given in the next theorem.

**Theorem 1.** *Suppose that  $X \in t(\nu, \sigma, 0)$ . Then, the following limit relation holds,*

$$\lim_{\lambda \rightarrow \infty} \lambda^\nu P(X \leq -\lambda) = \frac{\Gamma(\frac{\nu+1}{2}) \sigma^\nu}{\nu\Gamma(\nu/2)\sqrt{\pi}}. \quad (4)$$

---

<sup>2</sup>For more details on the properties of stable distributions, see Samorodnitsky and Taqqu (1994).

*Proof.* First, we express the probability  $P(X \leq -\lambda)$  in terms of the incomplete beta function taking advantage of the substitution  $t = \frac{\nu}{\nu+(x/\sigma)^2}$ ,

$$\begin{aligned} P(X \leq -\lambda) &= \frac{1}{\sigma} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \int_{\lambda}^{\infty} \left(1 + \frac{(x/\sigma)^2}{\nu}\right)^{-\frac{\nu+1}{2}} dx \\ &= \frac{1}{2} \frac{B_{\frac{\nu}{\nu+(\lambda/\sigma)^2}}(\nu/2, 1/2)}{B(\nu/2, 1/2)}, \end{aligned}$$

where  $B_x(a, b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$  is the incomplete beta function and  $B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$  is the beta function. The incomplete beta function is related to the hypergeometric function through the equality  $B_x(a, b) = \frac{x^a}{a}F(a, 1-b, a+1, x)$ , see Abramowitz and Stegun (1972). From the power series expansion of the hypergeometric function given in Abramowitz and Stegun (1972), we can construct the following first-order approximation of the incomplete beta function for small values of  $x$ ,

$$B_x(a, b) = \frac{x^a}{a} + o(x).$$

As a consequence,

$$P(X \leq -\lambda) = \frac{1}{\nu B(\nu/2, 1/2)} \left(1 + \frac{(\lambda/\sigma)^2}{\nu}\right)^{-\frac{\nu}{2}} + o((\lambda/\sigma)^{-\nu}), \quad (5)$$

which leads to the limit

$$\lim_{\lambda \rightarrow \infty} (\lambda/\sigma)^\nu P(X \leq -\lambda) = \frac{1}{\nu B(\nu/2, 1/2)}.$$

The result in (4) is obtained after expressing the beta function in terms of the gamma function and noticing that  $\Gamma(1/2) = \sqrt{\pi}$ . □

Concerning GND, we adopt the following definition:  $X \in GN(\kappa, \sigma, \mu)$  if it has the density

$$f(x) = \frac{\kappa}{2^{1+\frac{1}{\kappa}}\sigma\Gamma(1/\kappa)} e^{-\frac{1}{2}\left|\frac{x-\mu}{\sigma}\right|^\kappa}, \quad x \in R \quad (6)$$

where  $\kappa > 0$  determines the tail behavior,  $\sigma > 0$ , and  $\mu \in R$  are the scale and the location parameters, respectively. If  $\kappa = 1$ , then GND reduces to the Laplace distribution which has an exponential decay of the density; if  $\kappa = 2$ , then we obtain the Gaussian distribution. Therefore,  $\kappa \in (1, 2)$  determines a faster than exponential and slower than Gaussian decay of the density and  $\kappa \in (0, 1)$  determines a slower than exponential decay. If  $\kappa > 2$ , then



the tails of the distribution converge faster than the tails of the Gaussian distribution. At the limit, as  $\kappa \rightarrow \infty$ ,  $GN(\kappa, \sigma, \mu)$  converges to a uniform distribution defined on the interval  $(\mu - \sigma, \mu + \sigma)$ .

A more precise result on the tail decay of GND is provided in the next theorem.

**Theorem 2.** *Suppose that  $X \in GN(\kappa, \sigma, 0)$ . Then, the following limit relation holds,*

$$\lim_{\lambda \rightarrow \infty} \lambda^{\kappa-1} e^{\frac{1}{2} \left(\frac{\lambda}{\sigma}\right)^\kappa} P(X \leq -\lambda) = \frac{\sigma^{\kappa-1}}{2^{\frac{1}{\kappa}} \Gamma\left(\frac{1}{\kappa}\right)}. \quad (7)$$

*Proof.* First, we derive an expression for  $P(X \leq -\lambda)$  in which  $\lambda > 0$ ,

$$\begin{aligned} P(X \leq -\lambda) &= \int_{-\infty}^{-\lambda} \frac{\kappa}{2^{1+\frac{1}{\kappa}} \sigma \Gamma(1/\kappa)} e^{-\frac{1}{2} \left(\frac{-t}{\sigma}\right)^\kappa} dt \\ &= \frac{1}{2\Gamma\left(\frac{1}{\kappa}\right)} \int_{\lambda^\kappa/(2\sigma^\kappa)}^{\infty} y^{1/\kappa-1} e^{-y} dy \\ &= \frac{1}{2\Gamma\left(\frac{1}{\kappa}\right)} \Gamma\left(\frac{1}{\kappa}, \frac{1}{2} \left(\frac{\lambda}{\sigma}\right)^\kappa\right) \end{aligned}$$

where the second equality follows from the substitution  $y = \frac{1}{2} \left(-\frac{t}{\sigma}\right)^\kappa$  and  $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$  denotes the upper incomplete gamma function. The function  $\Gamma(s, x)$  has the following asymptotic behavior,

$$\lim_{x \rightarrow \infty} \frac{\Gamma(s, x)}{x^{s-1} e^{-x}} = 1,$$

see Abramowitz and Stegun (1972). Applying this property to the expression for  $P(X \leq -\lambda)$ , we obtain

$$\lim_{x \rightarrow \infty} \frac{\Gamma\left(\frac{1}{\kappa}, \frac{1}{2} \left(\frac{\lambda}{\sigma}\right)^\kappa\right)}{\left[\frac{1}{2} \left(\frac{\lambda}{\sigma}\right)^\kappa\right]^{1/\kappa-1} e^{-\frac{1}{2} \left(\frac{\lambda}{\sigma}\right)^\kappa}} = 1.$$

which leads to the result in equation (7). □

## 2.1 CVaR and fat-tailed distributions

In this section, we include the results for CVaR of fat-tailed distributions which we use in the paper. CVaR is defined as the average loss provided that the loss is larger than a quantile at a given probability level,

$$CVaR_\epsilon(X) = -\frac{1}{\epsilon} \int_0^\epsilon F_X^{-1}(p) dp \quad (8)$$

where  $F_X^{-1}(p) = \inf\{x : P(X \leq x) \geq p\}$  denotes the inverse distribution function of the random variable  $r_p$ . CVaR was suggested as a superior alternative to value-at-risk because it satisfies the coherence axioms in Artzner et al. (1998).

Working numerically with the definition in (8) is difficult because the quantile function is unbounded for probabilities close to zero. Therefore, we have to calculate the integral in the definition for every distributional assumption for  $X$ .

Stoyanov et al. (2004) calculated the CVaR for stable distributions. The result for the symmetric case is provided in the following theorem.

**Theorem 3.** *If  $X \in S_\alpha(\sigma, 0, \mu)$  with  $\alpha > 1$  and  $q_\epsilon \neq 0$  is the  $\epsilon$ -quantile of  $(X - \mu)/\sigma$ , then  $CVaR_\epsilon(X)$  admits the representation  $CVaR_\epsilon(X) = \sigma A_\epsilon + \mu$ , where*

$$A_\epsilon = \frac{\alpha}{1 - \alpha} \frac{|q_\epsilon|}{\pi \epsilon} \int_0^{\pi/2} g(\theta) \exp\left(-|q_\epsilon|^{\alpha/(\alpha-1)} v(\theta)\right) d\theta \quad (9)$$

where

$$g(\theta) = \frac{\sin(\alpha - 2)\theta}{\sin \alpha \theta} - \frac{\alpha \cos^2 \theta}{\sin^2 \alpha \theta}$$

$$v(\theta) = \left(\frac{\cos \theta}{\sin \alpha \theta}\right)^{\frac{\alpha}{\alpha-1}} \frac{\cos(\alpha - 1)\theta}{\cos \theta}$$

If  $q_\epsilon = 0$ , then  $A_\epsilon = \frac{2\Gamma(\frac{\alpha-1}{\alpha})}{\pi}$ .

A much simpler and easy to establish expression exists for the CVaR of the Student's  $t$  distribution. The result is provided in the next theorem and a proof can be found in Stoyanov and Rachev (2008a).

**Theorem 4.** *If  $X \in t(\nu, \sigma, \mu)$ , with  $\nu > 1$ , then  $CVaR_\epsilon(X)$  admits the representation  $CVaR_\epsilon(X) = \sigma B_\epsilon + \mu$ , where*

$$B_\epsilon = \frac{1}{\epsilon} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{\sqrt{\nu}}{(\nu-1)\sqrt{\pi}} \left(1 + \frac{q_\epsilon^2}{\nu}\right)^{\frac{1-\nu}{2}}, \quad (10)$$

in which  $q_\epsilon$  is the  $\epsilon$ -quantile of  $(X - \mu)/\sigma$ .

Finally, we calculate the CVaR for the GND which, as far as we know, is not available elsewhere.

**Theorem 5.** *If  $X \in GND(\kappa, \sigma, \mu)$ , then for any  $\epsilon < 0.5$ ,  $CVaR_\epsilon(X)$  admits the representation  $CVaR_\epsilon(X) = \sigma C_\epsilon + \mu$ , where*

$$C_\epsilon = \frac{1}{\epsilon} \frac{2^{\frac{1}{\kappa}-1}}{\Gamma(\frac{1}{\kappa})} \Gamma\left(\frac{2}{\kappa}, \frac{(-q_\epsilon)^\kappa}{2}\right) \quad (11)$$

in which  $q_\epsilon$  is the  $\epsilon$ -quantile of  $(X - \mu)/\sigma$  and  $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-t} dt$  is the upper incomplete gamma function.

*Proof.* Since  $X$  has a density,  $CVaR_\epsilon(X) = \sigma CVaR_\epsilon(Y) + \mu = -\sigma E(Y|Y \leq q_\epsilon) + \mu$ , where  $Y = (X - \mu)/\sigma$  denotes the standardized distribution. Assuming  $\epsilon < 0.5$ , the conditional expectation is calculated directly,

$$\begin{aligned} -E(Y|Y \leq q_\epsilon) &= -\frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} \frac{x\kappa}{2^{1+1/\kappa}\Gamma(1/\kappa)} e^{-\frac{1}{2}(-x)^\kappa} dx \\ &= \frac{1}{\epsilon} \frac{2^{\frac{1}{\kappa}-1}}{\Gamma(\frac{1}{\kappa})} \int_{(-q_\epsilon)^\kappa/2}^\infty y^{2/\kappa-1} e^{-y} dy. \end{aligned} \quad (12)$$

The second expression is obtained after the substitution  $y = (-x)^\kappa/2$  and the integral is recognized as the upper incomplete gamma function.  $\square$

### 3 CVaR and tail thickness sensitivity

Taking advantage of the positive homogeneity and the translation invariance property, the CVaR derivatives with respect to the distribution parameters of the models considered in Section 2 equal

$$\begin{aligned} \frac{\partial CVaR_\epsilon(X)}{\partial \theta_1} &= \sigma \frac{\partial CVaR_\epsilon(Y)}{\partial \theta_1} \\ \frac{\partial CVaR_\epsilon(X)}{\partial \sigma} &= CVaR_\epsilon(Y) \\ \frac{\partial CVaR_\epsilon(X)}{\partial \mu} &= -1 \end{aligned} \quad (13)$$

where  $\theta_1$  stands for  $\alpha$ ,  $\nu$  or  $\kappa$  and  $Y = (X - \mu)/\sigma$ .

In this section, we explore the derivative of CVaR with respect to the tail thickness parameter, which corresponds to the term  $\partial CVaR_\epsilon(Y)/\partial \theta_1$  in equation (13), and we compare it to the derivative with respect to the scale parameter. In the context of marginal rebalancing, this comparison can indicate which parameter is marginally more significant for a marginal risk reduction under the technical assumption that the return distributions of all portfolios belong to one and the same parametric class.

We begin with a general result proving that under the assumption of regularly varying tails and one technical condition, the sensitivity of CVaR with respect to tail thickness is bounded. Therefore, under this general hypothesis which includes the pure power decay, small changes in the tail thickness parameter cannot lead to huge changes in the risk measure. The main result is provided in the next theorem.

**Theorem 6.** Suppose that  $X$  has a regularly varying left tail,  $P(X \leq -u) \sim u^{-\alpha} L(u, \alpha)$  with  $\alpha > 1$ , where  $L(u, \alpha)$  is a slowly varying function for any  $\alpha$ . Then, for a sufficiently small  $\epsilon > 0$ , the derivative  $\partial CVaR_\epsilon(X)/\partial\alpha$  is approximated by

$$h_{\epsilon, \alpha}(X) = -\frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) \log(-x) dx + \frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) \frac{\partial \log(L(-x, \alpha))}{\partial \alpha} dx \quad (14)$$

where  $q_\epsilon$  denotes the  $\epsilon$ -quantile of  $X$ . Furthermore, if  $\partial L(-x, \alpha)/\partial\alpha$  is a slowly varying function, then  $\partial CVaR_\epsilon(X)/\partial\alpha$  is bounded for any  $\epsilon > 0$ .

*Proof.* We start from the definition of CVaR, take advantage of the substitution  $y = F(x)$  and after integration by parts, we obtain

$$CVaR_\epsilon(X) = -q_\epsilon(\alpha) + \frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon(\alpha)} F(x) dx. \quad (15)$$

In this transformation, we assume that  $F(x)$  is continuous at  $x = q_\epsilon(\alpha)$ . Using the assumed asymptotic behavior for the left tail, we obtain

$$CVaR_\epsilon(X) \approx -q_\epsilon(\alpha) + \frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon(\alpha)} (-x)^{-\alpha} L(-x, \alpha) dx. \quad (16)$$

where  $\epsilon$  is sufficiently small. The expression in (14) is derived by differentiating the right hand-side of (16) using the chain rule

$$\frac{d}{d\alpha} \left( \int_{-\infty}^{f(\alpha)} G(x, \alpha) dx \right) = \int_{-\infty}^{f(\alpha)} \frac{\partial G(x, \alpha)}{\partial \alpha} dx + G(f(\alpha), \alpha) f'(\alpha). \quad (17)$$

The boundedness of  $\partial CVaR_\epsilon(X)/\partial\alpha$  is a consequence of the assumed regular variation. The strategy is the following one. We consider the following bound,

$$|h_{\epsilon, \alpha}(X)| \leq \left| \frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) \log(-x) dx \right| + \left| \frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) \frac{\partial \log(L(-x, \alpha))}{\partial \alpha} dx \right|, \quad (18)$$

and focus on the first term demonstrating that it is bounded. The same reasoning holds for the second term. Finally,  $\partial CVaR_\epsilon(X)/\partial\alpha$  is bounded because it is in a neighbourhood of  $h_{\epsilon, \alpha}(X)$ .

The integrand of the first term in (14) can be represented as

$$F(x) \log(-x) \sim (-x)^{-\alpha} L(-x) \log(-x)$$

for a sufficiently small  $x$ . The term  $L(-x) \log(-x)$  is a product of two slowly varying functions and, as a consequence, is a slowly varying function.

Therefore,  $F(x) \log(-x)$  can be viewed as a regularly varying tail with the same index  $\alpha$ . Consider the following upper bound

$$\left| \int_{-\infty}^b F(x) \log(-x) dx \right| \leq |F(b) \log(-b)| + \left| \int_{-\infty}^b x d[F(x) \log(-x)] \right|.$$

The first term is a constant. By assumption,  $\alpha > 1$  and as a consequence the second term

$$\left| \int_{-\infty}^b x d[F(x) \log(-x)] \right| < \infty$$

for any  $b$  using a theorem in Feller (1971).

The same conclusion holds for the second term in (18). First, notice that

$$\left| \int_{-\infty}^{q_\epsilon} F(x) \frac{\partial \log(L(-x, \alpha))}{\partial \alpha} dx \right| \leq \int_{-\infty}^{q_\epsilon} F(x) \left| \frac{\partial \log(L(-x, \alpha))}{\partial \alpha} \right| dx.$$

The same reasoning can be used to demonstrate that the upper bound is finite using the assumption that  $\partial \log(L(-x, \alpha))/\partial \alpha$  is a slowly varying function. □

In the theorem, we assumed that the slowly varying component is also a function of the tail thickness parameter. Under the assumption that  $L(u, \alpha) = C_\alpha L(u)$ , we obtain the following simpler result.

**Corollary 1.** *If  $P(X \leq -u) \sim C_\alpha u^{-\alpha} L(u)$  with  $\alpha > 1$ , where  $C_\alpha$  is a constant, then*

$$h_{\epsilon, \alpha}(X) = -\frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) \log(-x) dx + \frac{B_\alpha}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) dx, \quad (19)$$

in which  $B_\alpha = d \log C_\alpha / d\alpha$ . An equivalent representation in terms of conditional moments is also available,

$$\begin{aligned} h_{\epsilon, \alpha}(X) = & E(X \log(-X) | X \leq q_\epsilon) - q_\epsilon \log(-q_\epsilon) \\ & + B_\alpha (q_\epsilon - E(X | X \leq q_\epsilon)). \end{aligned}$$

*Proof.* Integrating by parts and taking advantage of  $(x \log(-x) + 1)' = \log(-x)$ , we obtain the conditional moments representation. □

In the particular case of stable Paretian distributions, we obtain the following corollary.

**Corollary 2.** *If  $X \in S_\alpha(0, 1, 0)$  with  $\alpha > 1$ , then*

$$h_{\epsilon, \alpha}(X) = -\frac{C_\alpha}{\epsilon} \left[ \frac{(-q_\epsilon)^{1-\alpha} \log(-q_\epsilon)}{\alpha - 1} + \frac{(-q_\epsilon)^{1-\alpha}}{(\alpha - 1)^2} \right] + \frac{C'_\alpha}{\epsilon} \frac{(-q_\epsilon)^{1-\alpha}}{\alpha - 1}$$

where  $C_\alpha = \Gamma(\alpha) \sin(\pi\alpha/2)/(2\pi)$ ,  $C'_\alpha = dC_\alpha/d\alpha$  and  $\Gamma(x)$  stands for the gamma function.

*Proof.* If  $X \sim S_\alpha(0, 1, 0)$ , then  $P(X \leq -u) \sim C_\alpha u^{-\alpha}$ , where

$$C_\alpha = \Gamma(\alpha) \sin(\pi\alpha/2)/(2\pi).$$

The corollary is obtained after a direct calculation of the two integrals in (19).  $\square$

It is possible to prove a similar set of results for GND. The result below is the analogue of Theorem 6 assuming GND-type tails.

**Theorem 7.** *Suppose that  $X$  has a tail with an asymptotic behavior,  $P(X \leq -u) \sim C_\kappa u^{1-\kappa} e^{-\frac{u^\kappa}{2}}$  with  $\kappa > 0$ , where  $C_\kappa$  is a constant. Then, for a sufficiently small  $\epsilon > 0$ , the derivative  $\partial CVaR_\epsilon(X)/\partial\kappa$  is approximated by*

$$h_{\epsilon, \kappa}(X) = -\frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) \log(-x) \left( 1 + \frac{1}{2}(-x)^\kappa \right) dx + \frac{B_\kappa}{\epsilon} \int_{-\infty}^{q_\epsilon} F(x) dx \quad (20)$$

where  $q_\epsilon$  denotes the  $\epsilon$ -quantile of  $X$  and  $B_\kappa = d \log C_\kappa / d\kappa$ . Furthermore, the derivative  $\partial CVaR_\epsilon(X)/\partial\kappa$  is bounded for any  $\epsilon > 0$  on condition that  $B_\kappa$  is bounded.

*Proof.* Using the assumed asymptotic behavior of  $P(X \leq x)$  and the expression in (15), we obtain

$$CVaR_\epsilon(X) \approx -q_\epsilon(\kappa) + \frac{1}{\epsilon} \int_{-\infty}^{q_\epsilon(\kappa)} C_\kappa (-x)^{1-\kappa} e^{-\frac{(-x)^\kappa}{2}} dx. \quad (21)$$

We calculate the derivative  $\frac{\partial G(x, \kappa)}{\partial\kappa}$  in (17),

$$\begin{aligned} \frac{\partial G(x, \kappa)}{\partial\kappa} &= C'_\kappa (-x)^{1-\kappa} e^{-\frac{(-x)^\kappa}{2}} - C_\kappa (-x)^{1-\kappa} e^{-\frac{(-x)^\kappa}{2}} \log(-x) \left( 1 + \frac{1}{2}(-x)^\kappa \right) \\ &= \frac{C'_\kappa}{C_\kappa} F(x) - C_\kappa F(x) \log(-x) \left( 1 + \frac{1}{2}(-x)^\kappa \right), \end{aligned}$$

where  $C'_\kappa = dC_\kappa/d\kappa$ . As a next step, we apply the chain rule in (17) to the expression in (21) and we obtain the result in equation (20).

Concerning the boundedness of the derivative, we consider the two terms in (20) separately. Assume that  $q_\epsilon < 0$  which is not restrictive since the

theorem considers the case of small  $\epsilon$ . If  $B_\kappa$  is bounded, then the second term in (20) is bounded because the approximation through the asymptotic behavior of the tail

$$\begin{aligned} \int_{-\infty}^{q_\epsilon} F(x) dx &\approx C_\kappa \int_{-q_\epsilon}^{\infty} u^{1-\kappa} e^{-\frac{u^\kappa}{2}} du \\ &= C_\kappa \frac{2^{\frac{2}{\kappa}-1}}{\kappa} \int_{(-q_\epsilon)^\kappa/2}^{\infty} z^{2/\kappa-2} e^{-z} dz < \infty \end{aligned}$$

which follows from the substitution  $z = u^\kappa/2$ .

The first term can be handled in a similar manner taking advantage of the inequality  $\log(x) \leq 1 + x$ .

$$\begin{aligned} \int_{-\infty}^{q_\epsilon} F(x) \log(-x) \left(1 + \frac{1}{2}(-x)^\kappa\right) dx &\leq \int_{-q_\epsilon}^{\infty} F(-u)(1+u) \left(1 + \frac{1}{2}u^\kappa\right) du \\ &\approx C_\kappa \int_{-q_\epsilon}^{\infty} u^{1-\kappa} e^{-\frac{u^\kappa}{2}} (1+u) \left(1 + \frac{1}{2}u^\kappa\right) du \\ &= C_\kappa \frac{2^{\frac{2-\kappa}{\kappa}}}{\kappa} \int_{(-q_\epsilon)^\kappa/2}^{\infty} t^{\frac{1-\kappa}{\kappa}} e^{-t} (1 + 2^{1/\kappa} t^{1/\kappa}) (1+t) dt < \infty \end{aligned}$$

The last inequality follows because the last integral reduces to a sum of integrals of the type  $\int_x^\infty t^\alpha e^{-t} dt$  which is finite for  $x > 0$ . □

**Corollary 3.** *Suppose that  $X \in GN(\kappa, 1, 0)$ . Then, the constant  $C_\kappa = (2^{1/\kappa} \Gamma(1/\kappa))^{-1}$  and the result in equation (20) holds with*

$$B_\kappa = \frac{1}{\kappa^2} \left( \log 2 + \frac{\Gamma'(1/\kappa)}{\Gamma(1/\kappa)} \right) < \infty.$$

Furthermore,  $|B_\kappa| < \infty$  for any  $\kappa > 0$ .

*Proof.* The constant  $C_\kappa$  is calculated in equation (7) and the calculation of  $B_\kappa$  is straightforward.  $B_\kappa$  is bounded for any positive  $\kappa$  because the derivative of the gamma function is bounded for any  $\kappa > 0$ . □

Since the derivative of CVaR with respect to tail thickness for regularly varying tails and for the GND tails is finite, it is interesting to check how it compares to the derivative with respect to the scale parameter. We consider the following ratio

$$\tilde{\sigma}(\epsilon) = \frac{CVaR_\epsilon(Y)}{|\partial CVaR_\epsilon(Y)/\partial \theta_1|} \quad (22)$$

where  $Y = (X - \mu)/\sigma$  and  $\theta_1$  is the tail thickness parameter, which does not depend on the scale parameter. It can be interpreted as the value of the scale parameter which balances the sensitivities to tail thickness and scale. We demonstrate that for the three distribution classes,  $\tilde{\sigma}(\epsilon)$  becomes arbitrarily small when  $\epsilon$  decreases. This implies that even though  $\partial CVaR_\epsilon(Y)/\partial\theta_1$  is finite for  $\epsilon > 0$ , it can become arbitrarily larger than  $\partial CVaR_\epsilon(Y)/\partial\sigma$  for sufficiently small  $\epsilon$ . This is not a surprising result since the deeper we go into the tail, the more sensitive CVaR becomes with respect to tail thickness. As a by-product, we obtain that  $\partial CVaR_\epsilon(Y)/\partial\theta_1$  increases indefinitely when  $\epsilon \rightarrow 0$  because  $\lim_{\epsilon \rightarrow 0} CVaR_\epsilon(Y) = \infty$ .

**Theorem 8.** *Let  $X \in S_\alpha(\sigma, 0, \mu)$  with  $\alpha > 1$ . Then,  $\lim_{\epsilon \rightarrow 0} \tilde{\sigma}(\epsilon) = 0$ . Furthermore, the rate of convergence is given by*

$$\lim_{\epsilon \rightarrow 0} \log(-q_\epsilon) \tilde{\sigma}(\epsilon) = \alpha \quad (23)$$

where  $q_\epsilon$  is the  $\epsilon$ -quantile of  $Y = (X - \mu)/\sigma$ .

*Proof.* We consider the limit as  $\epsilon$  approaches zero, therefore we can calculate the numerator taking advantage of the asymptotic tail behavior. Integration by parts and the assumption that  $\alpha > 1$  lead to

$$CVaR_\epsilon(Y) = -q_\epsilon + \frac{1}{\epsilon} \frac{C_\alpha (-q_\epsilon)^{1-\alpha}}{\alpha - 1}$$

where  $q_\epsilon$  is the  $\epsilon$ -quantile of  $Y$ . Taking advantage of the result in Corollary 2, we obtain

$$\begin{aligned} \tilde{\sigma}(\epsilon) &= - \frac{-q_\epsilon + \frac{1}{\epsilon} \frac{C_\alpha (-q_\epsilon)^{1-\alpha}}{\alpha - 1}}{-\frac{C_\alpha}{\epsilon} \left[ \frac{(-q_\epsilon)^{1-\alpha} \log(-q_\epsilon)}{\alpha - 1} + \frac{(-q_\epsilon)^{1-\alpha}}{(\alpha - 1)^2} \right] + \frac{C'_\alpha}{\epsilon} \frac{(-q_\epsilon)^{1-\alpha}}{\alpha - 1}} \\ &= - \frac{(\alpha - 1)\epsilon(-q_\epsilon)^\alpha + C_\alpha}{-C_\alpha \left[ \log(-q_\epsilon) + \frac{1}{(\alpha - 1)} \right] + C'_\alpha}. \end{aligned}$$

The asymptotic result in equation (2) indicates that  $\epsilon(-q_\epsilon)^\alpha = P(Y \leq q_\epsilon)(-q_\epsilon)^\alpha \rightarrow C_\alpha$ . Therefore,  $\tilde{\sigma}(\epsilon) \rightarrow 0$  because the logarithm in the denominator is unbounded as  $\epsilon \rightarrow 0$ .

The rate of convergence is established by the same argument taking advantage of the expression,

$$\log(-q_\epsilon) \tilde{\sigma}(\epsilon) = - \frac{(\alpha - 1)\epsilon(-q_\epsilon)^\alpha + C_\alpha}{-C_\alpha \left[ 1 + \frac{1}{\log(-q_\epsilon)(\alpha - 1)} \right] + \frac{C'_\alpha}{\log(-q_\epsilon)}}. \quad (24)$$

As a consequence,  $\log(-q_\epsilon) \tilde{\sigma}(\epsilon) \rightarrow \frac{(\alpha - 1)C_\alpha + C_\alpha}{C_\alpha} = \alpha$ .  $\square$



A similar result can be established for the Student's  $t$  distribution. It is contained in the next theorem.

**Theorem 9.** *Let  $X \in t(\nu, \sigma, \mu)$  with  $\nu > 1$ . Then,  $\lim_{\epsilon \rightarrow 0} \tilde{\sigma}(\epsilon) = 0$ . Furthermore, the rate of convergence is given by*

$$\lim_{\epsilon \rightarrow 0} \log(-q_\epsilon) \tilde{\sigma}(\epsilon) = \nu, \quad (25)$$

where  $q_\epsilon$  is the  $\epsilon$ -quantile of  $Y = (X - \mu)/\sigma$ .

*Proof.* The numerator of  $\tilde{\sigma}(\epsilon)$  is computed explicitly in (10). We take advantage of an intermediate result in equation (5) and (19) in order to calculate the denominator. We can use the result in equation (19) because the asymptotic behavior in (4) implies a tail decay of the type  $P(X \leq \lambda) \sim C_\nu x^{-\nu}$ , where  $C_\nu = (\nu B(\nu/2, 1/2))^{-1}$ . After canceling the constant  $(B(\nu/2, 1/2)/\epsilon)^{-1}$  appearing in both the numerator and the denominator, we obtain

$$\begin{aligned} \tilde{\sigma}(\epsilon) &= \frac{\frac{\sqrt{\nu}}{(\nu-1)} \left(1 + \frac{q_\epsilon^2}{\nu}\right)^{\frac{1-\nu}{2}}}{\frac{1}{\nu} \int_{-\infty}^{q_\epsilon} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu}{2}} \log(-t) dt - B_\nu \int_{-\infty}^{q_\epsilon} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu}{2}} dt} \\ &= \frac{\frac{\sqrt{\nu}}{2} \left(1 + \frac{q_\epsilon^2}{\nu}\right)^{-\frac{1}{2}-\frac{\nu}{2}} \frac{2(-q_\epsilon)}{\nu}}{\frac{1}{\nu} \left(1 + \frac{q_\epsilon^2}{\nu}\right)^{-\frac{\nu}{2}} \log(-q_\epsilon) - B_\nu \left(1 + \frac{q_\epsilon^2}{\nu}\right)^{-\frac{\nu}{2}}} \\ &= \frac{\frac{\sqrt{\nu}}{\nu}}{\frac{1}{\nu} \left(\frac{1}{q_\epsilon^2} + \frac{1}{\nu}\right)^{\frac{1}{2}} \log(-q_\epsilon) - B_\nu \left(\frac{1}{q_\epsilon^2} + \frac{1}{\nu}\right)^{\frac{1}{2}}} \end{aligned}$$

in which the second equality follows after applying l'Hôpital's rule. The constant  $B_\nu = d \log C_\nu / d\nu$  can be calculated explicitly,

$$B_\nu = -\frac{1}{\nu} - \frac{\Gamma'(\nu/2)}{2\Gamma(\nu/2)} + \frac{\Gamma'((\nu+1)/2)}{2\Gamma((\nu+1)/2)}.$$

It follows that  $|B_\nu| < \infty$  if  $\nu > 0$  since the derivative of the gamma function is bounded if  $\nu > 0$ . As a consequence,  $\tilde{\sigma}(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .

The asymptotic behavior is derived from

$$\log(-q_\epsilon) \tilde{\sigma}(\epsilon) = \frac{\frac{\sqrt{\nu}}{\nu}}{\frac{1}{\nu} \left(\frac{1}{q_\epsilon^2} + \frac{1}{\nu}\right)^{\frac{1}{2}} - \frac{B_\nu}{\log(-q_\epsilon)} \left(\frac{1}{q_\epsilon^2} + \frac{1}{\nu}\right)^{\frac{1}{2}}}.$$

As a consequence,  $\log(-q_\epsilon) \tilde{\sigma}(\epsilon) \rightarrow \nu$  as  $\epsilon \rightarrow 0$ . □

A similar result can be established for GND. It is contained in the next theorem.

**Theorem 10.** *Let  $X \in GN(\kappa, \sigma, \mu)$ . Then,  $\lim_{\epsilon \rightarrow 0} \tilde{\sigma}(\epsilon) = 0$ . Furthermore, the rate of convergence is given by*

$$\lim_{\epsilon \rightarrow 0} \log(-q_\epsilon) \tilde{\sigma}(\epsilon) = \kappa, \quad (26)$$

where  $q_\epsilon$  is the  $\epsilon$ -quantile of  $Y = (X - \mu)/\sigma$ .

*Proof.* We use the first equality in equation (12) to calculate the numerator of  $\tilde{\sigma}(\epsilon)$ . The denominator is calculated by taking advantage of the result in equation (20) and the asymptotic behavior of the tail given in (7) which implies that  $C_\kappa = [2^{1/\kappa} \Gamma(1/\kappa)]^{-1}$  and  $B_\kappa = d \log C_\kappa / d\kappa$ . After canceling out the terms involving  $C_\kappa$ , we obtain

$$\begin{aligned} \tilde{\sigma}(\epsilon) &= \frac{\frac{\kappa}{2} \int_{-\infty}^{q_\epsilon} (-x) e^{-\frac{(-x)^\kappa}{2}} dx}{-\int_{-\infty}^{q_\epsilon} (-x)^{1-\kappa} e^{-\frac{(-x)^\kappa}{2}} \left( B_\kappa - \log(-x) \left( 1 + \frac{(-x)^\kappa}{2} \right) \right) dx} \\ &= \frac{\frac{\kappa}{2} (-q_\epsilon) e^{-\frac{(-q_\epsilon)^\kappa}{2}}}{-(-q_\epsilon)^{1-\kappa} e^{-\frac{(-q_\epsilon)^\kappa}{2}} \left( B_\kappa - \log(-q_\epsilon) \left( 1 + \frac{(-q_\epsilon)^\kappa}{2} \right) \right)} \\ &= \frac{\frac{\kappa}{2}}{-B_\kappa (-q_\epsilon)^{-\kappa} + \log(-q_\epsilon) \left( (-q_\epsilon)^{-\kappa} + \frac{1}{2} \right)} \end{aligned}$$

in which the second equality follows after applying l'Hôpital's rule. The constant  $B_\kappa$  is bounded and can be calculated, see Corollary 3. As a consequence  $\tilde{\sigma}(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .

The rate of convergence is readily established from the equality

$$\log(-q_\epsilon) \tilde{\sigma}(\epsilon) = \frac{\frac{\kappa}{2}}{-B_\kappa (\log(-q_\epsilon))^{-1} (-q_\epsilon)^{-\kappa} + \left( (-q_\epsilon)^{-\kappa} + \frac{1}{2} \right)}.$$

In effect,  $\log(-q_\epsilon) \tilde{\sigma}(\epsilon) \rightarrow \kappa$ . □

The convergence rates in (23), (25), and (26) imply that  $\tilde{\sigma}(\epsilon)$  decreases relatively slowly when  $X$  follows one of the three distribution models considered in the paper. The limit in the three cases equals the corresponding tail exponent meaning that within a given distribution class, convergence is asymptotically slower for larger values of the tail exponent.

Figure 1 provides a numerical illustration of the convergence rate. The top plot shows  $\tilde{\sigma}(\epsilon)$  as a function of the tail index  $\alpha$  for different choices of tail probability in the stable Paretian case. The derivative  $\partial CVaR_\epsilon(Y) / \partial \alpha$

is calculated numerically from the expression in Theorem 3. The plot illustrates that the convergence rate slows down at different values of  $\epsilon$  depending on the tail index. Thus,  $\epsilon = 0.01$ , which would be a typical choice for the tail probability in practice, leads to  $\tilde{\sigma}(\epsilon) > 0.2$  when  $\alpha > 1.4$ . This has interesting practical implications which we discuss in the next section. The reason for this behavior is that in (24)  $C_\alpha \rightarrow 0$  as  $\alpha \rightarrow 2$ , implying that the logarithmic decay appears asymptotically for very small  $\epsilon$  when  $\alpha$  is close to 2.

The middle and the bottom plots in Figure 1 show  $\tilde{\sigma}(\epsilon)$  calculated numerically using the expressions in Theorems 4 and 5 assuming the Student's  $t$  distribution and GND, respectively. We notice a similar behavior. Even  $\epsilon = 0.001$ , which is very deep in the tail for all practical purposes, leads to  $\tilde{\sigma}(\epsilon) > 1$  and  $\tilde{\sigma}(\epsilon) > 0.25$  for the Student's  $t$  distribution and GND, respectively. Comparing these numbers directly has to be done carefully because they correspond to different parametric classes in which the tail thickness parameter has a different function. For example, for one and the same tail probability, we notice that  $\tilde{\sigma}(\epsilon)$  for the Student's  $t$  model is larger than  $\tilde{\sigma}(\epsilon)$  for the stable Paretian model. This is in agreement with the fact that the Student's  $t$  tails for  $\nu > 3$  decay more quickly than stable Paretian tails even though both tails decay asymptotically according to a power law. We cannot, however, conclude that any other tail decay slower than the power decay leads to a smaller  $\tilde{\sigma}(\epsilon)$  for a fixed  $\epsilon$ . Even though there is no contradiction comparing the top plot to the middle plot, this conclusion would be wrong even for the tails decaying according to a power law for two reasons. First, the power decay is asymptotic,  $P(X \leq -u) \sim u^{-\alpha}L(u, \alpha)$ . Second, the slowly varying component  $L(u, \alpha)$  has an impact and can be different for different parametric families.

The bottom plot illustrates that a faster-than-exponential tail decay can lead to higher tail sensitivities relative to the scale sensitivity if compared to the slower power tail decay of the Student's  $t$  distribution even for tail probabilities as small as  $\epsilon = 0.001$ .

Therefore, the tail thickness sensitivity of CVaR is a feature of the parametric model and a special focus on a particular type of tail decay is not relevant in this context. In order to be able to compare across the three parametric models, it is necessary first to fit the distribution parameters on real data.

## 4 Tail thickness sensitivity and CVaR estimator

In this section, we consider the link between CVaR sensitivity with respect to distribution parameters and the problem of CVaR estimation from a sample of observations. Stoyanov and Rachev (2008b) prove that the asymptotic distribution of the empirical CVaR estimator is a totally skewed to the right

stable law if the tails are regularly varying with an index  $\alpha \in (1, 2)$ . However, if we assume a parametric model which results in a CVaR expression differentiable with respect to the distribution parameters, the limit behavior changes if the estimator of the distribution parameters is asymptotically normal. Under this assumption, the classical central limit theorem argument leads to

$$n^{1/2}(CVaR_\epsilon(X, \hat{\theta}) - CVaR_\epsilon(X, \theta)) \xrightarrow{w} N(0, \nabla' \Sigma \nabla) \quad (27)$$

where  $\theta$  is a  $k$ -dimensional vector of parameters,

$$\nabla = \left( \frac{\partial CVaR_\epsilon(X, \theta)}{\partial \theta_1}, \dots, \frac{\partial CVaR_\epsilon(X, \theta)}{\partial \theta_k} \right),$$

and  $\Sigma = \{\sigma_{ij}\}$  denotes the asymptotic covariance matrix of the parameter estimator  $\hat{\theta}$ . An example of an asymptotically normal estimator for the three distribution classes which we consider in the paper is the maximum likelihood method if  $\theta$  is in the interior of the parameter space.<sup>3</sup>

The asymptotic variance in (27),

$$\nabla' \Sigma \nabla = \sum_{i=1}^k \left( \frac{\partial CVaR_\epsilon(X, \theta)}{\partial \theta_i} \right)^2 \sigma_i^2 + \sum_{i \neq j} \frac{\partial CVaR_\epsilon(X, \theta)}{\partial \theta_i} \frac{\partial CVaR_\epsilon(X, \theta)}{\partial \theta_j} \sigma_{ij},$$

where  $\sigma_i^2 = \sigma_{ii}$ , contains both the derivative of the risk measure and the asymptotic covariance matrix of the estimator. As a consequence, it makes sense to compare the terms  $(\partial CVaR/\partial \theta_{i_1})^2 \sigma_{i_1}^2$  and  $(\partial CVaR/\partial \theta_{i_2})^2 \sigma_{i_2}^2$  corresponding to the tail thickness and the scale parameters respectively for different values of the tail probability  $\epsilon$ . In fact, we consider the following ratio,

$$r = \frac{|\partial CVaR/\partial \theta_{i_1}| \sigma_{i_1}}{|\partial CVaR/\partial \theta_{i_2}| \sigma_{i_2}} \quad (28)$$

where  $i_1$  corresponds to the tail thickness parameter and  $i_2$  corresponds to the scale parameter. The ratio in (28) compares the relative size of the corresponding two terms in the asymptotic variance of CVaR. Generally, a smaller derivative can be compensated for by a larger asymptotic variance  $\sigma_{ii}$  and vice versa. Moreover, it is interesting to see how the ratio changes with the tail probability. On the basis of the results in Section 3, the numerator is expected to grow faster than the denominator for smaller  $\epsilon$ .

---

<sup>3</sup>In the case of stable distributions, there are no closed-form expressions for the density function. We employ the FFT-based approach pioneered in Mittnik et al. (1999) to construct an approximation of the density in order to calculate numerically the likelihood function. Alternative approaches exist as well, see Doganoglu and Mittnik (1998) for a polynomial-based approximation.

We carry out an empirical study in which we compare the two components in the asymptotic variance of the CVaR estimator corresponding to the tail thickness and the scale parameters for stable Paretian, Student's  $t$ , and GND laws. We also calculate the ratio in (22) for the fitted distributions and compare it to the fitted scale parameter in order to verify how CVaR tail thickness sensitivity compares with CVaR scale sensitivity in a real-life situation.

In the empirical study, we consider the daily returns of the constituents of the German DAX 30 index in the period from January 1, 2008 to January 1, 2010 which includes the latest financial crisis. In order to clean the clustering of volatility effect, we fit first a GARCH(1,1) model,

$$\begin{aligned} y_t &= \mu + e_t \\ e_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= K + a\sigma_{t-1}^2 + be_{t-1} \end{aligned} \tag{29}$$

assuming a Student's  $t$  distributed residual, which is consistent with the assumption of a fat-tailed residual process. Then, we fit a stable law and a GND to the observed residuals. Finally, we calculate a forecast for the volatility

$$\hat{\sigma}_{T+1}^2 = \hat{K} + \hat{a}\hat{\sigma}_{T-1}^2 + \hat{b}e_{T-1}$$

from the fitted values and multiply the scale of the fitted distribution in order to obtain the distribution parameters of the conditional return distribution for the time instant  $T + 1$  as of time  $T$ .

Table 1 provides a summary about the fitted parameters for the 30 constituents. We report the interval containing 95% of the fitted values which is obtained by truncating the smallest 2.5% and the largest 2.5% of them.

In order to compare the relative impact of tail thickness, we consider three cases – with a small, medium, and large tail thickness parameter. We choose three stocks with a stable Paretian tail index equal to  $\alpha = 1.69$ ,  $\alpha = 1.8$ , and  $\alpha = 1.99$ , respectively. Note that a smaller tail index corresponds to a fatter tail.

Next, we identify three stocks from the DAX 30 universe with stable tail indices close to these values and we fit a stable Paretian, Student's  $t$ , and GND law using the maximum likelihood estimator.<sup>4</sup> The corresponding fitted tail thickness and scale parameters are given in Table 2. Certainly, the fitted tail indices imply different tail decay. For example, in the small tail index case,  $\hat{\kappa} = 1.07$  implies an exponential decay while the fitted Student's

---

<sup>4</sup>The three stocks are the following ones: Merck Group is the stock with a tail index equal to  $\alpha = 1.69$ , Fresenius SE is the the stock with a tail index equal to  $\alpha = 1.8$ , and Fresenius Medical Care is the the stock with a tail index equal to  $\alpha = 1.99$ .

$t$  and stable Paretian laws imply a power decay with different indices. Comparing the fitted scale parameters in all three cases to  $\tilde{\sigma}(\epsilon)$  in Figure 1, we find that even for  $\epsilon = 0.01$  the fitted scale parameters are about one order of magnitude smaller than  $\tilde{\sigma}(\epsilon)$ . Since  $\tilde{\sigma}(\epsilon)$  is the value of the scale parameter leading to equal CVaR tail thickness and scale sensitivities, this implies that the CVaR derivative with respect to the tail thickness parameter is an order of magnitude smaller than the CVaR derivative with respect to the scale parameter irrespective of the model for tail probabilities as small as  $\epsilon = 0.01$ . Therefore, at least as far as stable Paretian, Student's  $t$ , or GND laws are concerned, small errors in the scale parameter lead to larger deviations in CVaR than small errors in the tail thickness parameter.

This conclusion holds when fitting the distributions to daily data. If the frequency is lower, then, generally, the fitted scale parameter increases while the fitted tails are less thick and vice versa for higher frequencies. As a consequence, it is not clear if the same conclusion will hold for higher or lower frequency returns.

In the context of marginal rebalancing, this observation has the following implication assuming that the portfolio daily return distribution has a similar tail behavior and scale. Reducing marginally the scale parameter of the portfolio return distribution through a marginal change in the portfolio weights is a more efficient strategy to control portfolio CVaR than reducing marginally tail thickness.

Figure 2 compares the three models in terms of the ratio given in (28). The left column contains plots of the ratio calculated for tail probabilities  $\epsilon \in [0.01, 0.4]$  for the three stocks with a large, medium, and small tail index and the right column contains the corresponding plots of the asymptotic standard deviation of the CVaR estimator. A thicker tail (a smaller tail index) implies a relatively higher asymptotic variability due to the tail index parameter relative to the scale parameter. The ratio is smallest for GND if  $\epsilon < 0.3$  in the medium and small tail index cases which indicates that the asymptotic variance of the tail index estimator for GND is much smaller. An assumption of a fatter tail, however, does not necessarily lead to a CVaR estimator with a larger asymptotic variance for all  $\epsilon$ . The asymptotic variance of the GND based CVaR is larger than that of the Student's  $t$  based CVaR for almost all  $\epsilon$  in the small tail index case.

For the three models, decreasing the tail probability results in more variable CVaR estimators and also in a higher variability due to the tail index parameter, which is a natural result. We can also conclude that for very small tail probabilities, e.g.  $\epsilon < 0.01$ , the models assuming fatter tails lead to a larger proportion in the asymptotic standard deviation due to the tail thickness parameter and also to relatively larger total asymptotic standard deviation.

Another impact of the different tail assumptions, as suggested by the criticism in Heyde et al. (2006), concerns the estimated CVaR. Figure 3

shows the estimated CVaR as a function of the tail probability  $\epsilon \in [0.01, 0.1]$  for the three stocks with a large, medium, and small tail index compared to the historical CVaR estimated directly from the sample with no parametric assumptions. Figure 4 shows the fitted tails compared to a non-parametric kernel estimate of the tail. Generally, the medium and the small cases indicate that a more fat-tailed model implies a higher CVaR if  $\epsilon$  is small enough, in this example if  $\epsilon < 0.05$ . It is, however, apparent that questioning the functional form of the tail decay is not appropriate. On the middle and bottom plots of Figure 3, Student's  $t$  based CVaR generally underestimates the historical CVaR and stable Paretian CVaR generally overestimates the historical CVaR. Therefore, it is possible to find conservative and optimistic models even assuming only a power tail decay. As a result, the focus should be on choosing an appropriate parametric model based on arguments other than the particular type of tail decay. For example, a methodology for CVaR back-testing represents one possible approach. It may result in models with different functional forms of tail decay being statistically indistinguishable, which is the case with the top plot on Figure 3.

Nevertheless, it is clear that the smaller  $\epsilon$  is, the more significant the bias in the CVaR estimate can be and, therefore, leading to a higher model risk arising from a potential misspecification of the distributional model. As a consequence, constructing a realistic heavy-tailed CVaR model is a non-trivial exercise especially when the return distribution deviates significantly from the Gaussian law and when the tail probability is a small number, e.g.  $\epsilon < 0.03$ .

## 5 Conclusion

In this paper, we considered the problem of lack of robustness of CVaR with respect to the tail decay hypothesis in the context of stable Paretian, the Student's  $t$ , and GND laws allowing for a power, exponential, and faster-than-exponential tail decay. We proved that under the assumption of regular variation and GND-type tail decay, CVaR has a bounded derivative with respect to the tail index for  $\epsilon > 0$ . Therefore, as far as the sensitivity to the tail index is concerned, small variations in the tail index do not lead to abnormally large variations in CVaR. We demonstrated that even though bounded, the tail index derivative increases indefinitely in absolute value when  $\epsilon \rightarrow 0$  at a higher rate than the derivative with respect to the scale parameter for the three distribution models. This implies that the value of the scale parameter,  $\tilde{\sigma}(\epsilon)$ , resulting in equal derivatives converges to zero. The convergence rate, however, is determined by the logarithm of the corresponding quantile function leading to a slow rate. The practical implication of this result for the three distribution models based on daily returns is that the estimated scale parameter appears to be an order of magnitude smaller

than  $\tilde{\sigma}(\epsilon)$  for  $\epsilon \geq 0.01$  resulting in a tail index derivative an order of magnitude smaller than the scale parameter derivative. This result indicates that estimating properly the scale of the distribution is more important than having a very precise estimate for the tail index. For example, not taking into account properly the clustering of volatility effect can have an adverse impact on CVaR. Furthermore, an effective way of reducing marginally portfolio CVaR is the traditional approach of selling proportionately parts of the holdings in the risky stocks and buying a risk-free asset.

Finally, we compared the impact of the tail index sensitivity for smaller  $\epsilon$  on the asymptotic standard deviation of the CVaR estimator based on the maximum likelihood method for the distribution parameters. For the models assuming a fatter tail, decreasing the tail probability leads to a higher asymptotic variance and also to a higher proportion of the asymptotic variance due to the tail index relative to the scale parameter. This indicates that the main practical difficulty with fat-tailed models is not so much related to the variability of the risk measure with respect to the tail index but with the variance of the tail index estimator. Further on, whether CVaR is adequate or not depends, generally, on the parametric hypothesis and not necessarily on the particular functional form of the tail decay. As a consequence, special care should be taken to back-test fat-tailed CVaR models with small tail probabilities, especially when the data are very heavy-tailed.



## References

- Abramowitz, M. and I. A. Stegun (1972), *Handbook of Mathematical Functions*, Dover, New York.
- Artzner, P., F. Delbaen, J.-M. Eber and D. Heath (1998), ‘Coherent measures of risk’, *Math. Fin.* **6**, 203–228.
- Bibby, B. M. and M. Sorensen (2003), ‘Hyperbolic processes in finance’, *Handbook of heavy-tailed distributions in finance*, S. T. Rachev (ed) pp. 212–248.
- Chen, C. R., Y. Su and Y. Huang (2008), ‘Hourly index return autocorrelation and conditional volatility in an ear-gjr-garch model with generalized error distribution’, *Journal of Empirical Finance* **15**, 789–798.
- Doganoglu, T. and S. Mittnik (1998), ‘An approximation procedure for asymmetric stable paretian densities’, *Computational Statistics* **13**, 463–475.
- Fama, E. (1965), ‘The behavior of stock market prices’, *Journal of Business* **38**, 34–105.
- Feller, W. (1971), *An Introduction to Probability Theory and its Application. 2nd Edition*, Wiley, New York.
- Haas, M., S. Mittnik and M. Paoletta (2006), ‘Modeling and predicting market risk with laplace-gaussian mixture distributions’, *Journal of Financial Econometrics* **16**, 1145–1162.
- Heyde, C. and S. Kou (2004), ‘On the controversy over tailweight of distributions’, *Operations Research Letters* **32**, 399–408.
- Heyde, C., S. Kou and X. Peng (2006), ‘What is a good risk measure: Bridging the gaps between data, coherent risk measures, and insurance risk measures’, *Preprint, Columbia University*.
- Kim, Y. S., S. T. Rachev, M. L. Bianchi and F. J. Fabozzi (2008), ‘Financial market models with Lévy processes and time-varying volatility’, *Journal of Banking and Finance* **32**, 7, 1363–1378.
- Kim, Y. S., S. T. Rachev, M. L. Bianchi and F. J. Fabozzi (2010), ‘Tempered stable and tempered infinitely divisible GARCH models’, *Journal of Banking and Finance*, forthcoming.
- Kuester, K., S. Mittnik and M. Paoletta (2006), ‘Value-at-risk prediction: A comparison of alternative strategies’, *Journal Financial Econometrics* **4**(1), 53–89.

- Mandelbrot, B. (1963), ‘The variation of certain speculative prices’, *Journal of Business* **26**, 394–419.
- Mittnik, S. and M. Paoletta (2003), ‘Prediction of financial downside-risk with heavy-tailed conditional distributions’, *Handbook of Heavy Tailed Distributions in Finance*, S. T. Rachev (ed) pp. 385–404.
- Mittnik, S., M. Paoletta and S. Rachev (2002), ‘Stationarity of stable power-GARCH processes’, *Journal of Econometrics* **106**, 97–107.
- Mittnik, S., S. T. Rachev, T. Doganoglu and D. Chenyao (1999), ‘Maximum likelihood estimation of stable paretian models’, *Mathematical and Computer Modelling* **29**, 257–293.
- Rachev, S. T. and S. Mittnik (2000), *Stable Paretian Models in Finance*, John Wiley & Sons, Series in Financial Economics.
- Samorodnitsky, G. and M. Grabchak (2010), ‘Do financial returns have finite or infinite variance? A paradox and an explanation’, *Quantitative Finance*, forthcoming .
- Samorodnitsky, G. and M.S. Taqqu (1994), *Stable Non-Gaussian Random Processes*, Chapman & Hall, New York, London.
- Stoyanov, S., G. Samorodnitsky, S. Rachev and S. Ortobelli (2004), ‘Computing the portfolio conditional value-at-risk in the  $\alpha$ -stable case’, *Department of Econometrics and Statistics, University of Karlsruhe, Germany, Technical report* .
- Stoyanov, S. V. and S. T. Rachev (2008a), ‘Asymptotic distribution of the sample average value-at-risk’, *Journal of Computational Analysis and Applications* **10**, 465–483.
- Stoyanov, S. V. and S. T. Rachev (2008b), ‘Asymptotic distribution of the sample average value-at-risk in the case of heavy-tailed returns’, *Journal of Applied Functional Analysis* **3**, 443–461.

	Stable Paretian	Student's $t$	GND
Tail thickness	[1.53, 2]	[2.66, 19.1]	[0.87, 1.79]
Scale	[0.007, 0.017]	[0.011, 0.048]	[0.0055, 0.019]

Table 1: 95% confidence intervals for the tail thickness and the scale parameters of the three distribution models.

	Large	Medium	Small
Stable Paretian, $(\hat{\alpha}, \hat{\sigma})$	(1.99, 0.0107)	(1.80, 0.0092)	(1.69, 0.0084)
Student's $t$ , $(\hat{\nu}, \hat{\sigma})$	(11.90, 0.0152)	(4.70, 0.0152)	(3.52, 0.0159)
GND, $(\hat{\kappa}, \hat{\sigma})$	(1.57, 0.0117)	(1.18, 0.0075)	(1.02, 0.0058)

Table 2: The fitted tail thickness and scale parameters of the stable Paretian, Student's  $t$ , and GND laws of the large, medium, and small tail index cases.

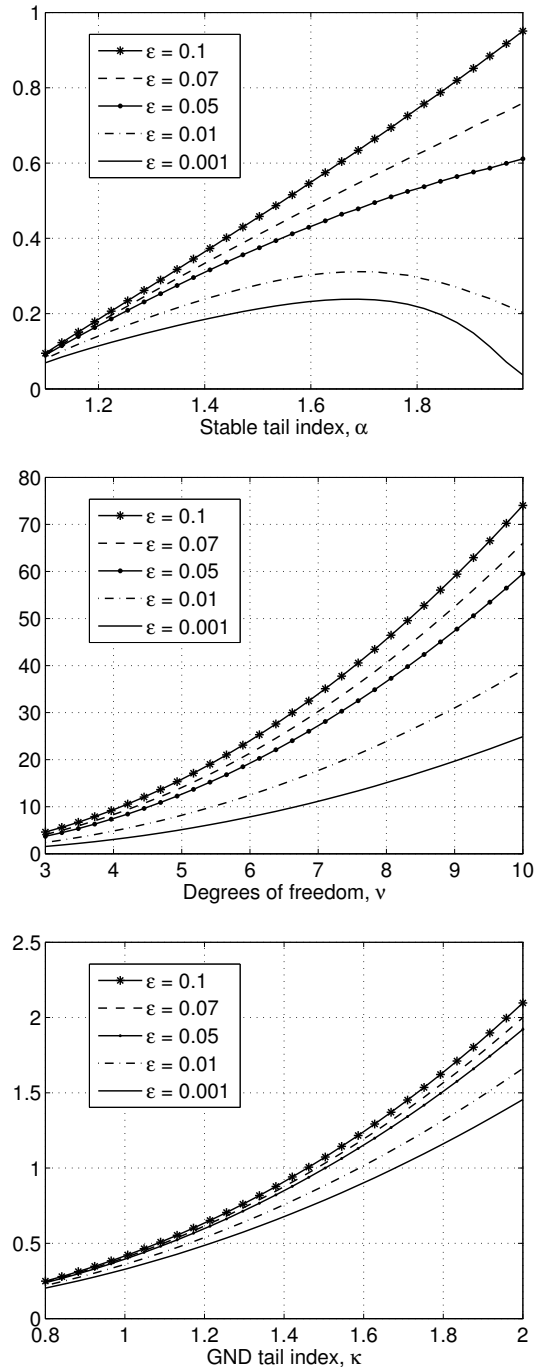


Figure 1: The ratio  $\tilde{\sigma}(\epsilon)$  as a function of tail thickness for different values of the tail probability for stable Paretian laws (top), the Student's  $t$  distribution (middle), and GND (bottom).

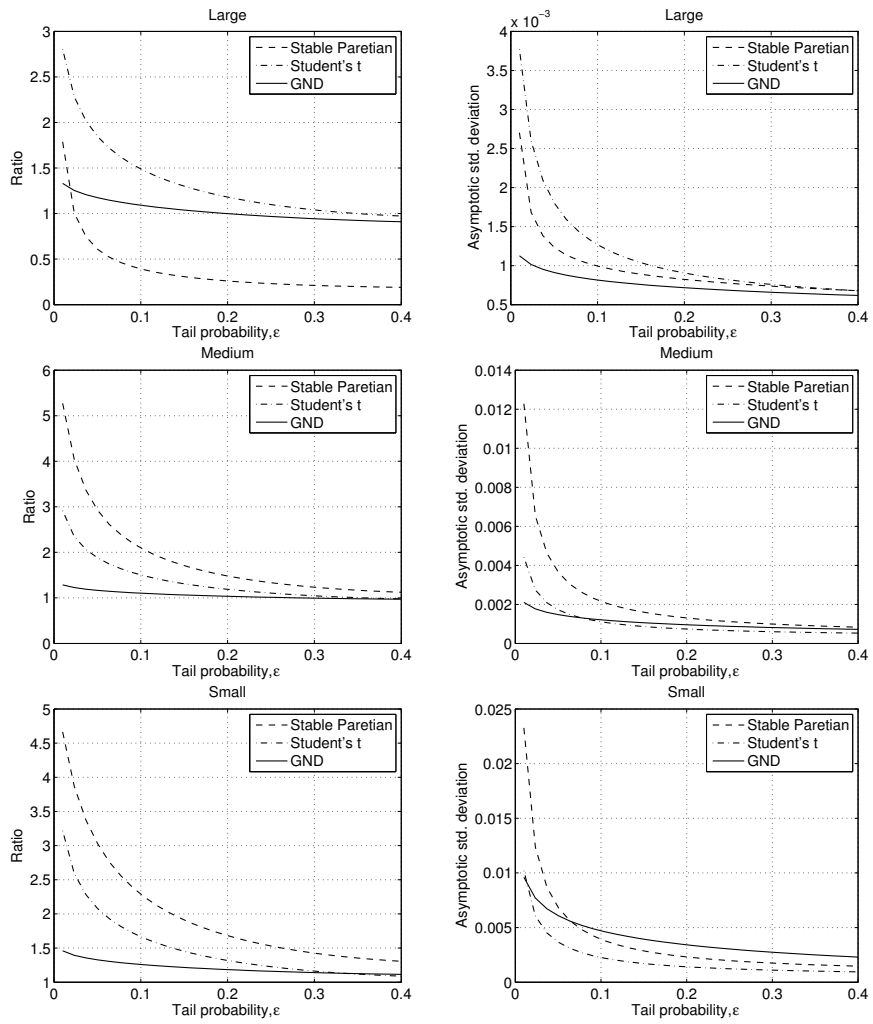


Figure 2: The ratio given in (28) (left column) and the asymptotic standard deviation of the CVaR estimator (right column) for tail probability  $\epsilon \in [0.01, 0.4]$  assuming stable Paretian, Student's  $t$ , and GND laws for stocks with a large, medium, and small value of the fitted tail index.

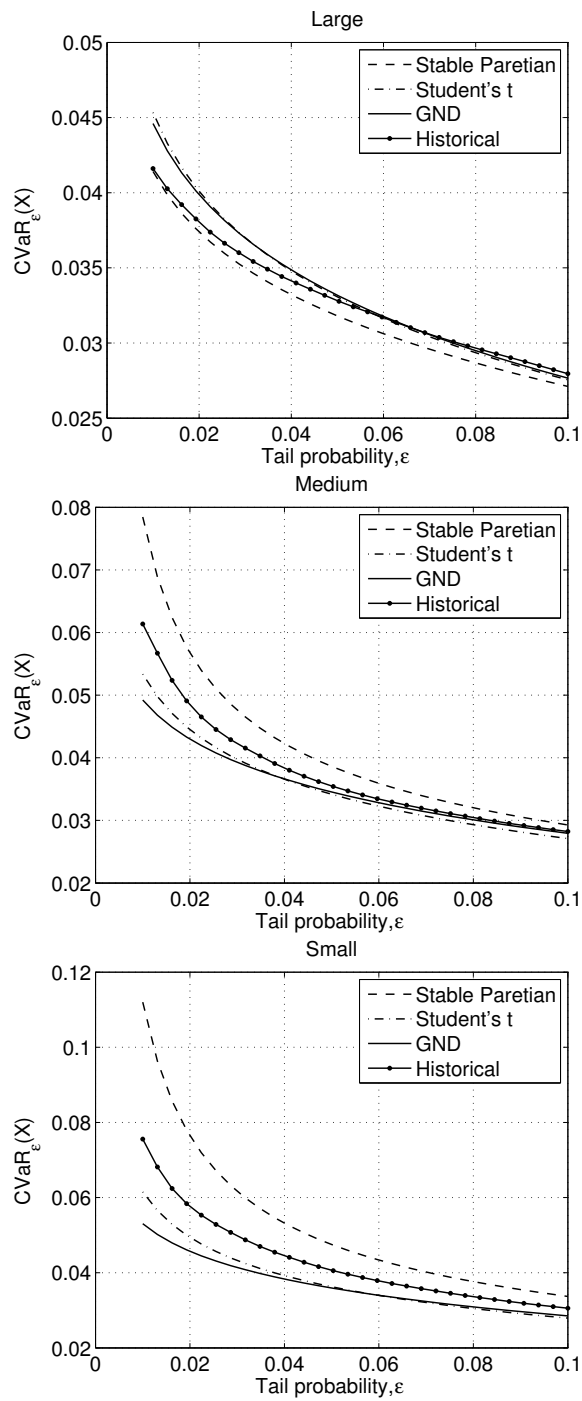


Figure 3: Estimated CVaR as a function of the tail probability  $\epsilon \in [0.01, 0.1]$  for stocks with a large, medium, and small value of the fitted tail index assuming stable Paretian, Student's  $t$ , GND, and the non-parametric historical method.

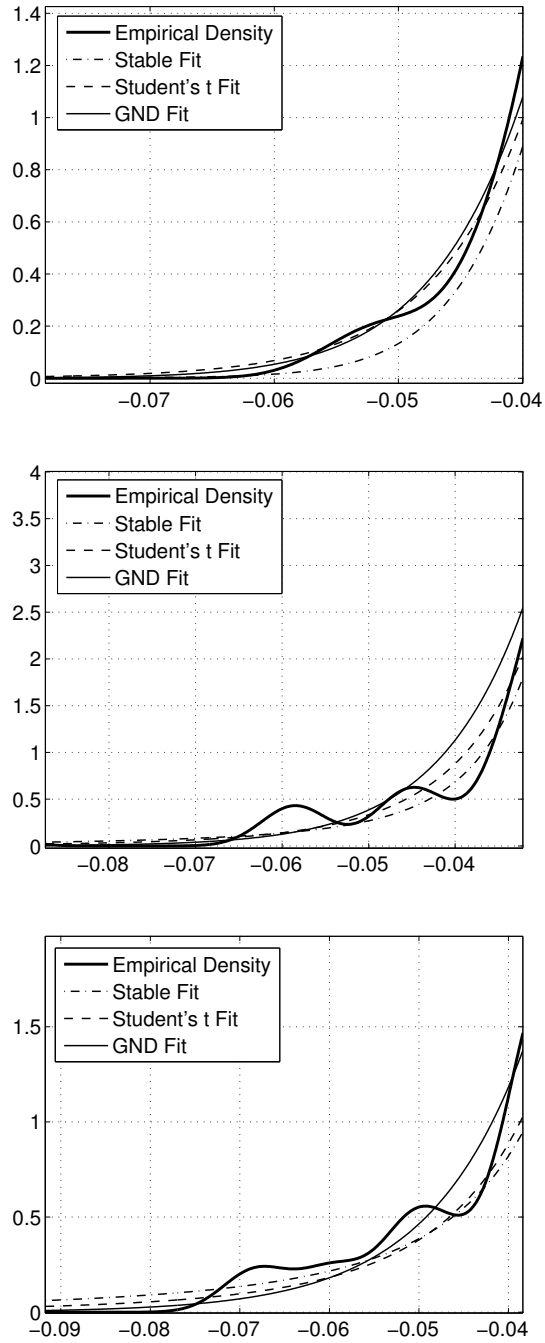


Figure 4: The tails of the fitted stable Paretian, Student's  $t$ , and GND laws compared to a kernel estimate of the empirical tail for stocks with a large (top), medium (middle), and small (bottom) value of the fitted tail index.



# Working Paper Series in Economics

---

recent issues

- No. 29** *Stoyan V. Stoyanov, Svetlozar T. Rachev, Frank J. Fabozzi: CVaR sensitivity with respect to tail thickness, May 2011*
  
- No. 28** *Young Shin Kim, Svetlozar T. Rachev, Michele Leonardo Bianchi, Frank J. Fabozzi: Tempered stable and tempered infinitely divisible GARCH models, May 2011*
  
- No. 27** *Takashi Kanamura, Svetlozar T. Rachev, Frank J. Fabozzi: A profit model for spread trading with an application to energy futures, May 2011*
  
- No. 26** *Michele Leonardo Bianchi, Svetlozar T. Rachev, Young Shin Kim, Frank J. Fabozzi: Tempered infinitely divisible distributions and processes, May 2011*
  
- No. 25** *Sebastian Kube, Michel André Maréchal and Clemens Puppe: The currency of reciprocity - gift-exchange in the workplace, April 2011*
  
- No. 24** *Clemens Puppe and Attila Tasnádi: Axiomatic districting, April 2011*
  
- No. 23** *Dinko Dimitrov and Clemens Puppe: Non-bossy social classification, April 2011*
  
- No. 22** *Kim Kaivanto and Eike B. Kroll: Negative recency, randomization device choice, and reduction of compound lotteries, April 2011*
  
- No. 21** *Antje Schimke and Thomas Brenner: Long-run factors of firm growth - a study of German firms, April 2011*
  
- No. 20** *Aaron B. Scholz: Spatial network configurations of cargo airlines, April 2011*
  
- No. 19** *Arne Beck: Public bus transport in Germany - a proposal to improve the current awarding system, April 2011*