

MULTIVARIATE HEAVY-TAILED MODELS FOR VALUE-AT-RISK ESTIMATION

CARLO MARINELLI

*Facoltà di Economia, Università di Bolzano
Piazza Università 1, I-39100 Bolzano, Italy*

STEFANO D'ADDONA*

*Department of International Studies, University of Rome 3
Via G. Chiabrera, 199, I-00145 Rome, Italy
daddona@uniroma3.it*

SVETLOZAR T. RACHEV

*Department of Applied Mathematics & Statistics
Stony Brook University, Stony Brook
NY 11794-3600, USA*

Received 25 April 2010
Accepted 15 December 2011

For purposes of Value-at-Risk estimation, we consider several multivariate families of heavy-tailed distributions, which can be seen as multidimensional versions of Paretian stable and Student's t distributions allowing different marginals to have different indices of tail thickness. After a discussion of relevant estimation and simulation issues, we conduct a backtesting study on a set of portfolios containing derivative instruments, using historical US stock price data.

Keywords: Value-at-Risk; multidimensional stable-like distribution; multidimensional t -like distribution; tail thickness; tail dependence; backtesting.

1. Introduction

The purpose of this paper is to assess the performance of some classes of multivariate laws with heavy tails in the estimation of Value-at-Risk for nonlinear portfolios. The inadequacy of Gaussian laws, in one or several dimensions, to model the distribution of risk factors, especially in view of applications to risk modeling, is well-documented in the empirical literature (see e.g. [4, 10] and references therein). Here we concentrate on models for risk factors that are multivariate extensions of the

*Corresponding author.

classical α -stable and Student's t distributions. In particular, we consider multivariate laws whose marginals may have different indices of tail thickness, and/or whose structures allow for tail dependence (i.e., roughly speaking, extreme movements of several risk factors may happen together).

Let us briefly recall how VaR is usually estimated for nonlinear portfolios (i.e. for portfolios containing derivative instruments), and what kind of improvements have been proposed. In the simplest setting, one uses a linear approximation of losses with normally distributed risk factors: denoting by L the loss over a certain time period, one sets $L \approx \langle \Delta, X \rangle$, where $X \sim N(m, Q)$ is a d -dimensional vector of Gaussian risk factors, Δ is an element of \mathbb{R}^d , and $\langle \cdot, \cdot \rangle$ stands for the usual scalar product of two vectors. Then $\langle \Delta, X \rangle$ follows a one-dimensional Gaussian distribution with mean $\langle \Delta, m \rangle$ and variance $\langle Q\Delta, \Delta \rangle$, so that (an approximation of) VaR can be obtained immediately. However, it is clear that such a scheme suffers from two major weaknesses: the linear approximation is inaccurate, as the payoff function of derivatives is usually highly non-linear, and the hypothesis that random factors are Gaussian is often inappropriate, as briefly mentioned above (the literature on this issue is very rich — see e.g. [5, 12, 13], to mention just a few classical references). Among the many improvements that have been suggested in literature, some focus on a better modeling of the nonlinear relation between L and X (e.g. by using quadratic approximations of the type $L \approx \langle \Delta, X \rangle + \langle \Gamma X, X \rangle$), but still assuming X Gaussian (see e.g. [10]), while others introduce alternative distributions of portfolio losses, often just in the univariate setting (see e.g. [15, 22]). To the best of our knowledge, however, there are only a small number of studies devoted to models that take into account both non-linearities and non-normality in a multivariate setting: Duffie and Pan [11] and Glasserman *et al.* [16] adopt the quadratic approximation and non-Gaussian risk factors. In particular, risk factors include a jump component in the first work, and are modeled by multivariate t distributions (or a modification thereof) in the latter. However, both works are devoted to different issues (analytic approximations and efficient simulation techniques, respectively), therefore they do not address the statistical issues related to the implementation of their models, and do not test their empirical performance on real data.

Our contributions are the following: we introduce a stable-like model for risk factors obtained by multivariate subordination of a Gaussian law on \mathbb{R}^d (see Sec. 2), such that each marginal (i.e. each risk factor) can have a different index of tail thickness. We construct estimators for the parameters of this distribution and we study their asymptotic behavior. An analogous program is carried out for a multivariate t -like law (see Sec. 3). In Secs. 4 and 5 we consider models of risk factors obtained by “deforming” the marginals of symmetric sub-Gaussian α -stable and multivariate t -distributed random vectors, respectively. Equivalently, using the language of copulas, we consider models of risk factors with symmetric α -stable resp. t -distributed marginals (possibly with different parameters) on which a sub-Gaussian α -stable resp. multivariate t copula is superimposed. In Sec. 7 we provide an extensive back-testing study of all parametric families of distributions using real data, on portfolios

containing both standard and exotic options, relying both on full revaluation of the portfolio value and on its quadratic approximation.

We conclude this introduction with a few words about notation: given a (possibly random) d -dimensional vector X , we shall denote its i th component, $1 \leq i \leq d$, by X_i . The inverse of an invertible function f will be denoted by f^{-1} . The Gaussian measure with mean m and covariance matrix Q will be denoted $N(m, Q)$. We shall write $X \sim \mathcal{L}$, with X a random variable and \mathcal{L} a probability measure, to mean that the law of X is \mathcal{L} . The α -stable measure on the real line with index α , skewness β , scale σ and location μ is denoted by $S_\alpha(\sigma, \beta, \mu)$, and we always use the parametrization adopted in [23].

2. Multivariate Stable-Like Risk Factors

2.1. Description of the model

Given a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $X : \Omega \rightarrow \mathbb{R}^d$ be a random vector of risk factors such that

$$X = A^{1/2}G, \tag{2.1}$$

where $A = \text{diag}(A_1, \dots, A_d)$ is a diagonal random matrix with independent entries,

$$A_i \sim S_{\alpha_i/2} \left(\left(\cos \frac{\pi\alpha_i}{4} \right)^{2/\alpha_i}, 1, 0 \right) \quad \forall i = 1, \dots, d, \tag{2.2}$$

and G is a \mathbb{R}^d -valued Gaussian random vector, independent of A , with mean zero and covariance matrix Q . In (2.2) we assume $\alpha_i \in]1, 2[$ for all $i = 1, \dots, d$.

Note that (2.1) and (2.2) imply that, for each $i = 1, \dots, d$, the i th marginal of X has distribution $S_{\alpha_i}(\sigma_i, 0, 0)$, where $\mathbb{E}[G_i^2] = 2\sigma_i^2$. In particular, risk factors are allowed to have different indices of tail thickness α_i , and they are dependent through the Gaussian component G .

2.2. Estimation

Let $X(t)$, $t = 1, \dots, n$ be independent samples from the distribution of X . For $p < \min_{1 \leq i \leq d} (\alpha_i)/2$, define the (improper) sample p th moment as

$$M_p(n) = n^{-1} \sum_{t=1}^n X_i(t)^{\langle p \rangle} X_j(t)^{\langle p \rangle},$$

where $X^{\langle p \rangle} := |X|^p \text{sgn}(X)$. Note that, by Cauchy-Schwarz' inequality, we have

$$\mathbb{E}|X_i X_j|^p \leq (\mathbb{E}|X_i|^{2p})^{1/2} (\mathbb{E}|X_j|^{2p})^{1/2} < \infty,$$

thus also, by Kolmogorov's strong law of large numbers,

$$\lim_{n \rightarrow \infty} M_p(n) = \mathbb{E}(X_i X_j)^{\langle p \rangle} \quad \text{a.s.}$$

Since the random matrix A and the random vector G are independent, one has

$$\mathbb{E}(X_i X_j)^{\langle p \rangle} = (\mathbb{E}A_i^{p/2})(\mathbb{E}A_j^{p/2})\mathbb{E}(G_i G_j)^{\langle p \rangle},$$

where (see e.g. [23, p. 18])

$$\begin{aligned} \mathbb{E}A_i^{p/2} &= \frac{2^{p/2}\Gamma(1-p/\alpha_i)}{p \int_0^\infty u^{-p/2-1} \sin^2 u du} \left(1 + \tan^2 \frac{\alpha_i \pi}{4}\right)^{\frac{p}{2\alpha_i}} \left(\cos \frac{\pi \alpha_i}{4}\right)^{\frac{p}{\alpha_i}} \cos \frac{p\pi}{4} \\ &=: C_{\alpha_i, p}. \end{aligned}$$

The constant $C_{\alpha, p}$ can be computed explicitly, recalling that

$$\int_0^\infty u^{-p/2-1} \sin^2 u du = -2^{p/2-1} \cos \frac{\pi p}{4} \Gamma(-p/2).$$

Let us now define the function

$$\begin{aligned} f_p :]-1, 1[&\rightarrow \mathbb{R} \\ q &\mapsto \mathbb{E}[(Z_1 Z_2)^{\langle p \rangle}], \end{aligned}$$

where Z_1, Z_2 are jointly normal random variables with covariance matrix

$$\begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix}.$$

For any given $p < \min_i(\alpha_i)/2$, matching the theoretical signed p th moments of $X_i X_j$ with their sample counterparts, we obtain the following estimator for the matrix Q :

$$\hat{Q}_{ij} = 2\sigma_i \sigma_j f_p^- \left(\frac{n^{-1} \sum_{t=1}^n X_i(t)^{\langle p \rangle} X_j(t)^{\langle p \rangle}}{2^p \sigma_i^p \sigma_j^p C_{\alpha_i, p} C_{\alpha_j, p}} \right), \quad i, j = 1, \dots, d.$$

If $\{\sigma_i\}_i$ and $\{\alpha_i\}_i$ are not known a priori, but we rather have only consistent estimators $\{\hat{\sigma}_{in}\}_i$ and $\{\hat{\alpha}_{in}\}_i$, respectively, one can easily deduce (by several applications of the continuous mapping theorem), that the estimator of Q obtained replacing α_i with $\hat{\alpha}_{in}$ and σ_i with $\hat{\sigma}_{in}$ in the above expression is still consistent.

Remark 2.1. (i) As far as the estimation of the covariance matrix Q is concerned, the heavy tailed assumption does not imply any extra computational burden.

(ii) For our purposes, it is enough to choose $p = 1/2$, as we always assume $\alpha_i > 1$ for all $i = 1, \dots, d$ (as is well-known, this is equivalent to assuming that all returns have finite mean).

(iii) Unfortunately we are not aware of an explicit expression for the function $q \mapsto f_p(q)$. However, it can be expressed as an integral with respect to a Gaussian

measure in \mathbb{R}^2 :

$$\begin{aligned} f_p(\rho) &= \frac{1}{2\pi\sqrt{\det Q}} \int_{\mathbb{R}^2} (x_1 x_2)^{\langle p \rangle} e^{-\frac{1}{2}\langle Q^{-1}x, x \rangle} dx \\ &= \frac{1}{2\pi\sqrt{1-q^2}} \int_{\mathbb{R}^2} (x_1 x_2)^{\langle p \rangle} e^{-\frac{1}{2(1-q^2)}(x_1^2 - 2qx_1 x_2 + x_2^2)} dx_1 dx_2 \end{aligned} \quad (2.3)$$

which can be computed by numerical integration with essentially any accuracy. Figure 1 plots the function $f_{1/2}$ on the interval $[0, 1[$.

Let us consider a simplified case: $d = 2$, $\sigma_1 = \sigma_2 = 1/\sqrt{2}$, and α_1, α_2 given. The assumption $d = 2$ is harmless, as in any case the method works componentwise. The case of unknown α_i and σ_i can be dealt with replacing them with their corresponding consistent estimators, as discussed above.

Let us define

$$\hat{q}_n = f^{\leftarrow} \left(\frac{n^{-1} \sum_{t=1}^n X_1(t)^{\langle p \rangle} X_2(t)^{\langle p \rangle}}{C_{\alpha_1, p} C_{\alpha_2, p}} \right) \quad (2.4)$$

We first prove the following lemma:

Lemma 2.1. *The function $f_p :]-1, 1[\rightarrow \mathbb{R}$ is bounded, continuously differentiable, concave increasing on $] -1, 0[$ and convex increasing on $]0, 1[$.*

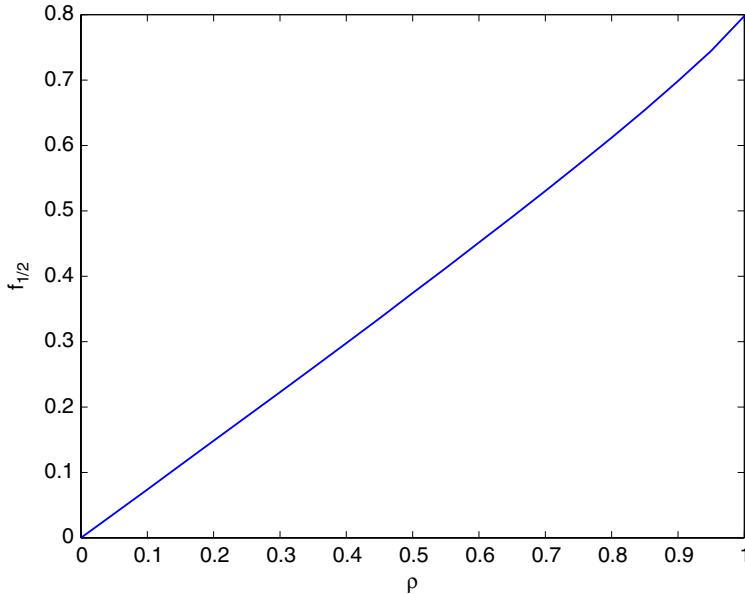


Fig. 1. Plot of the function f_p , with $p = 1/2$.

Proof. Boundedness follows by concavity of the function $x \mapsto |x|^p$ for $p < 1$ and Jensen's inequality, that yield

$$|f_p(q)| = |\mathbb{E}Z_1^{(p)} Z_2^{(p)}| \leq \mathbb{E}|Z_1 Z_2|^p \leq (\mathbb{E}|Z_1 Z_2|)^p \leq 1,$$

where the last inequality follows by Cauchy-Schwarz' inequality and $\mathbb{E}Z_1^2 = \mathbb{E}Z_2^2 = 1$. Continuous differentiability w.r.t. q is immediate by inspection of (2.3). Differentiating (2.3) w.r.t. q twice, one gets (after some cumbersome but elementary calculations) $f_p'(q) > 0$ for all $q \in]-1, 1[$, and $f_p''(q) < 0$ for $q < 0$, $f_p''(0) = 0$, $f_p''(q) > 0$ for $q > 0$. The lemma is thus proved. \square

It is easy to prove that \hat{q}_n is strongly consistent, i.e. that $\hat{q}_n \rightarrow q$ a.s. as $n \rightarrow \infty$. In fact, as above, since $p < (\min_i \alpha_i)/2$, by Kolmogorov's strong law of large numbers one has

$$f_p(\hat{q}_n) = \frac{n^{-1} \sum_{t=1}^n X_1(t)^{(p)} X_2(t)^{(p)}}{C_{\alpha_1, p} C_{\alpha_2, p}} \xrightarrow{n \rightarrow \infty} \mathbb{E}(Z_1 Z_2)^{(p)} = f_p(q) \quad \text{a.s.},$$

from which we can conclude thanks to the continuous mapping theorem and the continuity of f^{\leftarrow} .

We are now going to prove that the estimator (2.4) is asymptotically normal, under a more stringent assumption on the chosen value of p . Let us define the function

$$g_p : \mathbb{R}^2 \rightarrow \mathbb{R} \\ x \mapsto \frac{x_1^{(p)} x_2^{(p)}}{C_{\alpha_1, p} C_{\alpha_2, p}}.$$

It is clear that the estimator (2.4) can be defined as the solution of the equation

$$\mathbb{P}_n g_p := \frac{1}{n} \sum_{k=1}^n g_p(X(k)) = \mathbb{E}_q g_p(X) =: f_p(q), \quad (2.5)$$

where \mathbb{P}_n stands for the (averaged) empirical measure of the sample $X(1), \dots, X(n)$, i.e.

$$\mathbb{P}_n := \frac{1}{n} \sum_{k=1}^n \delta_{X(k)}.$$

Proposition 2.1. *If $p < \min_i(\alpha_i)/4$, then \hat{q}_n is asymptotically normal and satisfies*

$$\sqrt{n}(\hat{q}_n - q) \Rightarrow N(0, f_p'(q)^{-2}(\mathbb{E}_q[g_p^2(X)] - f_p^2(q))), \quad (2.6)$$

where " \Rightarrow " stands for convergence in law.

Proof. We have proved in Lemma 2.1 that $f_p(q) = \mathbb{E}_q g_p(X)$ is a bijection on the open set $] -1, 1[$, it is continuously differentiable on its domain, and $f_p'(x) \neq 0$ for

all $x \in] - 1, 1[$. Moreover, as it follows from (2.4) and (2.5), one can write

$$\sqrt{n}(\hat{q}_n - q) = \sqrt{n}(f_p^-(\mathbb{P}_n g_p) - f_p^-(\mathbb{E}_q g_p(X))). \quad (2.7)$$

We have, by the strong law of large numbers, $\mathbb{P}_n g_p \rightarrow \mathbb{E}_q g_p$ a.s. as $n \rightarrow \infty$. Recalling that by hypothesis $p < \min_i(\alpha_i)/4$, it follows that $\mathbb{E}_q g_p^2(X) < \infty$, hence, by the central limit theorem,

$$\sqrt{n}(\mathbb{P}_n g_p - \mathbb{E}_q g_p(X)) \Rightarrow N(0, \mathbb{E}_q g_p^2(X) - f_p^2(q)).$$

An application of the delta method, taking into account the inverse function theorem, now yields the result. \square

A shortcoming of the asymptotic confidence interval implied by the above proposition is that the asymptotic variance depends on the parameter to be estimated itself. One can overcome this problem by a variance stabilizing transformation: let us define the function $\gamma :] - 1, 1[\rightarrow \mathbb{R}$,

$$\gamma_p(q) = \mathbb{E}_q g_p^2(X) - (\mathbb{E}_q g_p(X))^2 = \frac{\mathbb{E}|X_1 X_2|^{2p}}{C_{\alpha_1, 2p} C_{\alpha_2, 2p}} - \frac{(\mathbb{E}X_1^{(p)} X_2^{(p)})^2}{C_{\alpha_1, p}^2 C_{\alpha_2, p}^2}$$

and

$$\varphi_p(x) = \int_0^x \frac{f_p'(y)}{\gamma_p^{1/2}(y)} dy.$$

Then, again by the delta method, we obtain

$$\sqrt{n}(\varphi_p(\hat{q}_n) - \varphi_p(q)) \Rightarrow N\left(0, \varphi_p'(q)^2 \frac{\gamma_p(q)}{f_p'(q)^2}\right) = N(0, 1),$$

and a corresponding asymptotic confidence interval for q as

$$q \in [\varphi_p^-(\varphi_p(\hat{q}_n) - z_\alpha/\sqrt{n}), \varphi_p^-(\varphi_p(\hat{q}_n) + z_\alpha/\sqrt{n})].$$

This asymptotic normality result for $\varphi_p(\hat{q}_n)$ would of course be better if we had an explicit expression for φ_p , which instead needs to be approximated numerically. However, since both f_p and γ_p are smooth functions (i.e. at least C^2), constructing a numerical approximation of φ_p is a rather simple task.

Remark 2.2. The proof of the previous proposition, as well as the construction of the variance-stabilizing transformation, depend crucially on the assumption that $\sigma_1, \sigma_2, \alpha_1, \alpha_2$ are known (cf. the assumptions stated immediately before formula (2.4)). Therefore, for application purposes, the result should be applied with care, and would do only as a “first approximation”. Nonetheless, it is also quite common in the estimation of multivariate models to separate the estimation of the parameters for the marginals from the estimation of the dependence structure. It would certainly be interesting to obtain asymptotic confidence intervals treating simultaneously $\sigma_i, \alpha_i, i = 1, 2$, and q as parameters to be estimated.

2.3. Simulation

In view of the results of the previous subsection, we assume that the covariance matrix Q is known, hence, with a slight but harmless abuse of notation, we shall write Q instead of \hat{Q} .

Random vectors from the distribution of X can be simulated by the following simple algorithm:

- (i) generate d independent random variables $Z_i \sim N(0, 1)$, $i = 1, \dots, d$, and form the random vector $Z = (Z_1, \dots, Z_d) \sim N(0, I)$, so that $Q^{1/2}Z \sim N(0, Q)$;
- (ii) independently from Z , generate d independent random variables from the distribution of A_i , $i = 1, \dots, d$, as defined in (2.2);
- (iii) setting $A = \text{diag}(A_1, \dots, A_d)$, one has that $A^{1/2}Q^{1/2}Z$ is a sample from the d -dimensional law of X

Note that the only computational overhead with respect to the simulation of a Gaussian vector is the simulation of the stable subordinators, for which nonetheless efficient algorithms are available (see e.g. [23]).

2.4. Extensions

Let us remark that the model (2.1) for the vector of risk factors can be extended to allow for asymmetries. In particular, setting

$$\tilde{X} = X + B = A^{1/2}G + B,$$

where B is a random vector, independent of X , with independent components B_i distributed according to the law $S_{\alpha_i}(\sigma_{B_i}, 1, 0)$, we have that the i th marginal of the vector \tilde{X} has distribution $S_{\alpha_i}(\tilde{\sigma}_i, \tilde{\beta}_i, 0)$, where

$$\tilde{\sigma}_i = (\sigma_i^{\alpha_i} + \sigma_{B_i}^{\alpha_i})^{1/\alpha_i}, \quad \tilde{\beta}_i = \frac{\sigma_{B_i}^{\alpha_i}}{\sigma_i^{\alpha_i} + \sigma_{B_i}^{\alpha_i}}.$$

One can then estimate the parameters α_i , $\tilde{\sigma}_i$ and $\tilde{\beta}_i$ fitting (e.g. by maximum likelihood estimation) a general Paretian stable law to observed data, and obtain (corresponding estimates of the) values of σ_i , σ_{B_i} :

$$\sigma_i = (1 - \tilde{\beta}_i)^{1/\alpha_i} \tilde{\sigma}_i, \quad \sigma_{B_i} = \tilde{\beta}_i^{1/\alpha_i} \tilde{\sigma}_i.$$

Note also that, since we assume $\alpha_i > 1$ for all $i = 1, \dots, d$, we have $\mathbb{E}\tilde{X}_i = 0$ for all i . Finally, an estimate of Q can be obtained by a rather involved modification of method of fractional moments introduced in Sec. 2.2 above. In particular, assuming $d = 2$ for simplicity and using the notation of Sec. 2.2, let us set

$$f_p(\rho; \alpha_1, \alpha_2, c_1, c_2) = \mathbb{E}(c_1 A_1 G_1^{1/2} + B_1)^{\langle p \rangle} (c_1 A_1 G_1^{1/2} + B_2)^{\langle p \rangle},$$

where c_1 , c_2 are positive constants and B_1 , B_2 are independent with $B_i \sim S_{\alpha_i}(1, 1, 0)$, $i = 1, 2$. Then a moment estimator for ρ can be constructed in a rather

obvious way by matching theoretical and sample fractional moments using the function $f_p(\cdot; \alpha_1, \alpha_2, c_1, c_2)$, treating $\alpha_i, c_i, i = 1, 2$, as “known” (by an immediate scaling argument, the constants c_1, c_2 are uniquely determined by σ_i and $\tilde{\sigma}_i, i = 1, 2$). This method performs unfortunately much slower in comparison to the corresponding procedure in the symmetric case, as the function f_p cannot be computed off-line as in Sec. 2.2.

Moreover, model (2.1) does not allow for tail dependence among different risk factors. As a remedy, one may use the series representation of stable subordinators (see e.g. [23]), setting

$$A_i = \sum_{k=0}^{\infty} \gamma_k^{2/\alpha_i}, \quad i = 1, \dots, n,$$

where $(\gamma_k)_{k \geq 0}$ is a (fixed) sequence of independent standard Gamma random variables. The analysis of this model, however, is considerably more involved, and we plan to elaborate on these issues in a future work.

3. Multivariate t -Like Risk Factors

3.1. Description of the model

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let us consider a d -dimensional random vector of risk factors X such that

$$X_k = \frac{G_k}{\sqrt{V_k/\nu_k}}, \quad k = 1, \dots, d, \tag{3.1}$$

where $G \sim N(0, Q)$ and V_1, \dots, V_d are independent one-dimensional χ^2 -distributed random variables with parameters ν_1, \dots, ν_d , respectively. We also assume that G and (V_1, \dots, V_d) are independent. Then, for each $k = 1, \dots, d$, the k th marginal of X is distributed according to a Student’s t distribution with parameter ν_k , multiplied by $\sigma_k := (\mathbb{E}G_k^2)^{1/2}$. In particular, as in the case of the previous section, risk factors may have different indices of tail thickness (measured by ν_k), and their dependence comes from the Gaussian component G .

3.2. Estimation

Assuming for the time being $\nu_k, k = 1, \dots, d$, to be known, let us estimate the covariance matrix Q by the method of moments. We shall assume from now on that $\nu_k > 2$ for all k , which implies in particular that $\mathbb{E}X_k^2 < \infty$ for all k . One has

$$\begin{aligned} \mathbb{E}X_h X_k &= \sqrt{\nu_h \nu_k} \mathbb{E}G_h G_k \mathbb{E}V_h^{-1/2} \mathbb{E}V_k^{-1/2} \\ &= Q_{hk} \sqrt{\nu_h \nu_k} \mathbb{E}V_h^{-1/2} \mathbb{E}V_k^{-1/2} \end{aligned}$$

for all $h \neq k$, and

$$\mathbb{E}X_k^2 = Q_{kk} \nu_k \mathbb{E}V_k^{-1} = \sigma_k^2 \nu_k \mathbb{E}V_k^{-1}.$$

Denoting, for simplicity, a random variable with $\chi^2(\nu)$ distribution by V , the density of V is given by

$$f_\nu(x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{\nu/2-1} e^{-\frac{x}{2}},$$

so that

$$\begin{aligned} \mathbb{E}V^{-1/2} &= \int_0^\infty x^{-1/2} f_\nu(x) dx = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^\infty x^{\nu/2-3/2} e^{-\frac{x}{2}} dx \\ &= \frac{\Gamma(\nu/2 - 1/2)}{\sqrt{2}\Gamma(\nu/2)} \end{aligned}$$

and, similarly,

$$\mathbb{E}V^{-1} = \int_0^\infty x^{\nu/2-2} e^{-\frac{x}{2}} dx = \frac{1}{2} \frac{\Gamma(\nu/2 - 1)}{\Gamma(\nu/2)} = \frac{1}{\nu - 2}.$$

Here we have used the definition of Gamma function,

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt, \quad z > 0,$$

and its ‘‘factorial’’ property $\Gamma(z + 1) = z\Gamma(z)$. The above calculations yield

$$Q_{hk} = \frac{2}{\sqrt{\nu_h \nu_k}} \frac{\Gamma(\nu_h/2)}{\Gamma(\nu_h/2 - 1)} \frac{\Gamma(\nu_k/2)}{\Gamma(\nu_k/2 - 1)} \mathbb{E}X_h X_k, \quad h \neq k,$$

and

$$Q_{kk} = \sigma_k^2 = \frac{\nu_k - 2}{\nu_k} \mathbb{E}X_k^2.$$

We have thus obtained the following moment estimator for Q :

$$\hat{Q}_{hk} = \frac{2}{\sqrt{\nu_h \nu_k}} \frac{\Gamma(\nu_h/2)}{\Gamma(\nu_h/2 - 1)} \frac{\Gamma(\nu_k/2)}{\Gamma(\nu_k/2 - 1)} \frac{1}{n} \sum_{t=1}^n X_h(t) X_k(t), \quad h \neq k,$$

and

$$\hat{Q}_{kk} = \hat{\sigma}_k^2 = \frac{\nu_k - 2}{\nu_k} \frac{1}{n} \sum_{t=1}^n X_k(t)^2.$$

Note that, for each k , ν_k can be estimate by one-dimensional maximum likelihood on the k th marginal, thus obtaining a family of consistent estimators $\hat{\nu}_k$, $k = 1, \dots, d$. Therefore, the corresponding estimator of Q obtained by substituting in the previous expressions each ν_k with $\hat{\nu}_k$, for each k , is still consistent.

We can now prove that \hat{Q}_{hk} is asymptotically normal. For compactness of notation, we shall set

$$C_\nu := \frac{\sqrt{2}}{\sqrt{\nu}} \frac{\Gamma(\nu/2)}{\Gamma(\nu/2 - 1)},$$

and we shall consider only the case $h \neq k$. The asymptotic normality of the estimators $\hat{\sigma}_k$ can be established analogously (see also Sec. 3.4).

Proposition 3.1. *Let $d = 2$,*

$$Q = \begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix},$$

and

$$\hat{q}_n := C_{\nu_1} C_{\nu_2} \frac{1}{n} \sum_{t=1}^n X_1(t) X_2(t).$$

Then one has

$$\sqrt{n}(\hat{q}_n - q) \Rightarrow N(0, v_{\nu_1, \nu_2}(q)),$$

where

$$v_{\nu_1, \nu_2}(q) = \frac{\nu_1 C_{\nu_1}^2}{\nu_1 - 2} \frac{\nu_2 C_{\nu_2}^2}{\nu_2 - 2} (2q^2 + 1) - q^2$$

Proof. We have $\text{Var } q_n = \mathbb{E}q_n^2 - q^2$ and

$$\begin{aligned} \mathbb{E}\hat{q}_n^2 &= C_{\nu_1}^2 C_{\nu_2}^2 \mathbb{E}X_1^2 X_2^2 = \nu_1 \nu_2 C_{\nu_1}^2 C_{\nu_2}^2 \mathbb{E}G_1^2 G_2^2 \mathbb{E}V_1^{-1} \mathbb{E}V_2^{-1} \\ &= \frac{\nu_1 C_{\nu_1}^2}{\nu_1 - 2} \frac{\nu_2 C_{\nu_2}^2}{\nu_2 - 2} \mathbb{E}G_1^2 G_2^2, \end{aligned}$$

where we have used the identity $\mathbb{E}V^{-1} = (\nu - 2)^{-1}$. To compute $\mathbb{E}G_1^2 G_2^2$, let us write

$$\begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \stackrel{d}{=} \begin{bmatrix} 1 & 0 \\ q & \sqrt{1 - q^2} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix},$$

with $(Z_1, Z_2) \sim N(0, I)$. This yields, recalling that the fourth moment of a standard Gaussian measure is equal to 3,

$$\mathbb{E}G_1^2 G_2^2 = q^2 \mathbb{E}Z_1^4 + 1 - q^2 = 2q^2 + 1, \tag{3.2}$$

thus also

$$\text{Var } \hat{q}_n = \frac{\nu_1 C_{\nu_1}^2}{\nu_1 - 2} \frac{\nu_2 C_{\nu_2}^2}{\nu_2 - 2} (2q^2 + 1) - q^2,$$

whence the result follows by the central limit theorem. \square

Remark 3.1. One could derive from this asymptotic normality result an asymptotic confidence interval using a variance stabilizing transformation, as shown in the previous section. The same caveats discussed in Remark 2.2 apply of course in this case as well.

3.3. Simulation

Generating random vectors from the distribution of a multivariate t -like distribution is a straightforward modification of the procedure outlined in Sec. 2.3 above.

3.4. Extensions

Since marginals of the random vector X follow a univariate t distribution, they are symmetric. In order to allow for asymmetric marginals, one may posit $X = (X_1, \dots, X_d)$,

$$X_k := \tilde{X}_k - \eta_k := \frac{G_k + m_k}{\sqrt{V_k/\nu_k}} - \eta_k, \quad k = 1, \dots, d,$$

where $G \sim N(0, Q)$, and $m = (m_1, \dots, m_d)$, $\eta = (\eta_1, \dots, \eta_d) \in \mathbb{R}^d$. Then for the k th marginal one has that $X_k + \eta_k$ follows a noncentral t -distribution. The reason for subtracting the vector η from \tilde{X} is that $\mathbb{E}\tilde{X} \neq 0$, unless $m = 0$, and it is common to assume that risk factors have mean zero. Unfortunately the density of the noncentral t law is expressed in terms of a definite integral depending on parameters (see e.g. [25]), hence maximum likelihood estimation on the marginals becomes numerically quite involved. On the other hand, assuming $\nu_k > 4$ for all k , one can use the method of moments to construct estimators for $\nu = (\nu_1, \dots, \nu_d)$, m , η and Q . In fact, considering k fixed and equal to 1 for the sake of simplicity, the constraint $\mathbb{E}X_1 = 0$ translates into the relation

$$\eta_1 = m_1 \sqrt{\nu_1} \mathbb{E}V_1^{-1/2} = m_1 \sqrt{\nu_1} \frac{\Gamma((\nu_1 - 1)/2)}{\sqrt{2}\Gamma(\nu_1/2)}.$$

Since we need to estimate four parameters, we need other three equations. These can be obtained by matching the second, third, and fourth sample moments to the corresponding theoretical moments, which are known in closed form (see e.g. [17]).

We should also observe that in general it is not necessary to match moments of integer order to obtain consistent and asymptotically normal estimators. One may also use fractional moments, as it has been done in the previous section, thus relaxing the assumptions on the parameters ν_k . For instance, let X be as in (3.1), $d = 2$, $Q = \begin{bmatrix} 1 & q \\ q & 1 \end{bmatrix}$, and consider the problem of estimating q . Setting $g_p(x_1, x_2) = x_1^{(p)} x_2^{(p)}$, we can write

$$\mathbb{E}g_p(X) = (\nu_1 \nu_2)^{p/2} \mathbb{E}V_1^{-p/2} \mathbb{E}V_2^{-p/2} \mathbb{E}(G_1 G_2)^{(p)}.$$

Note that $\mathbb{E}(G_1 G_2)^{(p)} = f_p(q)$, where f_p is the function introduced and studied in Sec. 2.2, and, in analogy to a calculation already encountered in this section,

$$\mathbb{E}V_k^{-p/2} = \frac{\Gamma(\nu_k/2 - p/2)}{2^{p/2} \Gamma(\nu_k/2)}, \quad k = 1, 2. \tag{3.3}$$

This relation can be used as a basis for a moment estimator, as in Sec. 2.2. Choosing p small enough, one does not need to assume $\nu_k > 2$.

4. Multivariate Meta-Stable Risk Factors

4.1. Description of the model

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let $G \sim N(0, Q)$ be a d -dimensional random vector with $\det Q \neq 0$, and

$$A \sim S_{\alpha_0/2} \left(\left(\cos \frac{\pi \alpha_0}{4} \right)^{2/\alpha_0}, 1, 0 \right),$$

with A and G independent. The random vector $X' := A^{1/2}G$ is then symmetric α -stable with characteristic function

$$\mathbb{E} e^{i\langle \xi, X' \rangle} = e^{-|\langle Q\xi, \xi \rangle|^{\alpha_0/2}},$$

in particular X' has an elliptically contoured distribution (see e.g. [23] for the properties of so-called sub-Gaussian α -stable laws, and [6] for elliptically contoured distributions). As is well-known, the marginals of X' are α -stable with index α_0 , i.e. they have all the same index of tail thickness (as measured by α_0). In order to build a model allowing for different tail behavior along different coordinates, one may set $X = f(X')$, with f a deterministic (nonlinear) injective function, for instance to “deform” the marginals of X' (a large part of the literature on the applications of copulas to risk management is centered around this simple idea). A common procedure (see e.g. [16] in a slightly different context) is to define a random vector $X = (X_1, \dots, X_d)$ as

$$X_k = \sigma_k F_{\alpha_k}^{\leftarrow} (F_{\alpha_0}(X'_k)), \quad k = 1, \dots, d, \tag{4.1}$$

with $\sigma_k, k = 1, \dots, d$ positive scaling constants, $\alpha_k \in]1, 2]$ for all $k = 1, \dots, d$, and the diagonal elements of Q are normalized to one. Here and throughout this section $F_\alpha, \alpha \in]0, 2]$, stands for the one-dimensional distribution function of a standard symmetric stable law with index α . It is clear that the law of the k th marginal of X is symmetric α -stable with index α_k . Using the terminology introduced in [14], the random vector X has a meta-elliptical distribution, which we call meta-stable. Note that X' , hence also X , are expected to have nontrivial tail dependence between any two marginals because of the common factor A .

4.2. Estimation

One can estimate the parameters of a meta-stable distribution thanks to the following remarkable relation (see [14, Thm. 3.1] and also [21]): let $X = (X_1, \dots, X_d)$ be a random vector with meta-elliptical distribution, and denote the Kendall’s τ of X_i and $X_j, i, j = 1, \dots, d$, by τ_{ij} . Then we have

$$\tau_{ij} = \frac{2}{\pi} \arcsin Q_{ij},$$

which immediately yields the estimator

$$\hat{Q}_{ij} = \sin \frac{\pi}{2} \hat{\tau}_{ij}.$$

It is worth recalling that Kendall's tau statistic is a U -statistic of order 2 with a bounded kernel, therefore it is asymptotically normal (see e.g. [26, Sec. 12.1]). Unfortunately however there does not seem to be an explicit expression for the asymptotic variance, at least not (to the best of our knowledge) in the cases considered in this paper. One can also infer, by an application of the delta method, that the above estimator of Q_{ij} , $i, j = 1, \dots, d$, is also asymptotically normal.

Moreover, the parameters σ_k and α_k , $k = 1, \dots, d$, can easily be estimated e.g. by maximum likelihood on the marginals of X . Finally, the parameter α_0 can be estimated as follows, where, in view of the above, we treat the parameters δ_k , ν_k and the matrix Q as known (in practice they will have to be replaced by their consistent estimators): defining the \mathbb{R}^d -valued random vector $U = (U_1, \dots, U_d)$ as

$$U_k := F_{\alpha_k}(X_k/\sigma_k), \quad k = 1, \dots, d,$$

since $U_k = F_{\alpha_0}(X'_k)$ for all $k = 1, \dots, d$, by the results in Appendix A.1 we have that the law of U admits the density

$$p_U(u_1, \dots, u_d; \alpha_0) = \frac{h(F_{\alpha_0}^{\leftarrow}(u_1), \dots, F_{\alpha_0}^{\leftarrow}(u_d))}{\prod_{k=1}^d f_{\alpha_0}(F_{\alpha_0}^{\leftarrow}(u_k))},$$

where f_{α_0} is the density of $S_{\alpha_0}(1, 0, 0)$, and h is the function defined in Appendix A.2. The parameter α_0 can then be estimated by maximum likelihood.

4.3. Simulation

By (4.1) we have $X_k = \sigma_k F_{\alpha_k}^{\leftarrow}(F_{\alpha_0}(A^{1/2}G_k))$, $k = 1, \dots, d$, from which a simulation scheme completely analogous to that outlined in Sec. 2.3 can be devised. From the computational point of view, the main problem is that there is no closed-form expression for the cumulative distribution function and for the inverse distribution function of a one-dimensional stable law. However, there exist representations of them as integrals, which can be implemented numerically. This procedure can become computationally very expensive when simulating large samples. To reduce the simulation time, one can compute off-line $F_{\alpha}(x)$ and $F_{\alpha}^{\leftarrow}(x)$ for sufficiently many values of α and x , and replace the numerical integration by interpolation. Since both $(x, \alpha) \mapsto F_{\alpha}(x)$ and $(x, \alpha) \mapsto F_{\alpha}^{\leftarrow}(x)$ are smooth functions (at least for $\alpha > 1$), interpolated values provide accurate approximations to numerical integrals.

4.4. Extensions

The model can easily be extended to accommodate asymmetric marginals: it is enough to set

$$X_k = H_k^{\leftarrow}(F_{\alpha_0}(A^{1/2}G_k)), \quad k = 1, \dots, d,$$

where H_k is the cumulative distribution function of the general asymmetric centered α -stable law $S_{\alpha_k}(\sigma_k, \beta_k, 0)$. The estimation of this extended model is completely analogous to the symmetric case discussed above, with the only difference

that the parameters β_k , $k = 1, \dots, d$, will also have to be estimated. This can be accomplished again by maximum likelihood estimation on the marginals.

5. Multivariate Meta- t Risk Factors

5.1. Description of the model

On a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let G and V be a d -dimensional random vector with law $N(0, Q)$ and an independent one-dimensional random variable with χ^2 distribution with ν_0 degrees of freedom, respectively. We shall call the law of the random vector $X' = G/\sqrt{V/\nu_0}$ a multivariate t distribution (with parameters ν_0 and Q). There are other possible multivariate generalizations of Student's t distribution (see e.g. [19]), but we shall concentrate exclusively on this definition, which seems to be the most widely used in financial applications.

In complete analogy to the meta-stable model discussed in the previous section, the marginals of X have the same tail thickness (as measured by ν_0), but one expects nontrivial tail dependence between any two marginals because of the common factor V . In order to allow for different tail behavior along different coordinates, one can proceed as in the previous section. In particular (see e.g. [16]), one may define the d -dimensional random vector X as

$$X_k = \delta_k F_{\nu_k}^{\leftarrow}(F_{\nu_0}(X'_k)), \quad k = 1, \dots, d, \quad (5.1)$$

where $\delta_k > 0$ for all $k = 1, \dots, d$, F_ν denotes (here and throughout this section) the distribution function of a one-dimensional t law with ν degrees of freedom, and $Q_{kk} = 1$ for all $k = 1, \dots, d$. Then the k th marginal is t distributed with ν_k degrees of freedom, thus overcoming the problem of having all marginals with the same tail thickness. The distribution of the random vector X belongs to the class of meta-elliptical distributions introduced in [14]. In the latter reference the law of X is called meta- t , terminology which we have borrowed here.

5.2. Estimation

The estimation algorithm for the meta- t model is completely analogous to the one for the meta-stable model. In fact, since X , as just recalled, has a meta-elliptical distribution, the matrix Q can be estimated thanks to its relationship with Kendall's tau already mentioned in Sec. 4.2 above (cf. also [9, 21] about parameter estimation for the t copula). As in the previous Section, Kendall's tau statistics as well as the corresponding estimator of Q_{ij} , $i, j = 1, \dots, d$, are asymptotically normal.

Similarly, the parameters δ_k and ν_k , $k = 1, \dots, d$, can easily be estimated by maximum likelihood on the marginals of X . Finally, the parameter ν_0 can be estimated by maximum likelihood: treating, for the sake of simplicity, the parameters δ_k , ν_k , $k = 1, \dots, d$, and the matrix Q as known, let us define the d -dimensional random vector U as

$$U_k := F_{\nu_k}(X_k/\delta_k), \quad k = 1, \dots, d. \quad (5.2)$$

Recalling the explicit expression for the density of a multivariate t (see e.g. [19]), (5.1) and the result in Appendix A.1 imply that the law of U admits the density

$$p_U(u_1, \dots, u_d; \nu_0) = \frac{1}{(\det Q)^{1/2}} \frac{\Gamma((\nu_0 + d)/2) \Gamma(\nu_0/2)^{d-1}}{\Gamma((\nu_0 + 1)/2)^d} (1 + \nu_0^{-1} \langle Q^{-1} \tilde{u}, \tilde{u} \rangle)^{-\frac{\nu_0+d}{2}} \\ \times \prod_{k=1}^d (1 + \nu_0^{-1} \tilde{u}_k^2)^{\frac{\nu_0+1}{2}},$$

where $\tilde{u}_k := F_{\nu_0}^{\leftarrow}(u_k)$. In practice, of course one needs to replace ν_k , δ_k and Q with the estimates obtained e.g. by the above methods.

5.3. Simulation

Random samples from the distribution of X can be generated by a rather straightforward modification of the procedure outlined in Sec. 4.3. In fact, the distribution function of the univariate t distribution, as well as its inverse, are implemented in several software packages (such as Octave), even though they do not admit a closed-form representation.*

5.4. Extensions

As in the meta-stable case, one can generalize meta- t laws to allow for skewed marginals replacing F_{ν_k} , $k = 1, \dots, d$, in (5.2) with the cumulative distribution functions of noncentral t laws, in analogy to the case discussed in Sec. 3.4.

6. Estimation of Value-at-Risk by Simulation

We shall denote by L the loss of a portfolio depending on the vector of risk factors X . Recall that the Value-at-Risk (VaR) of a portfolio at confidence level β (usually $\beta = 0.95$ or $\beta = 0.99$) is simply the β quantile of the distribution of portfolio losses. Since it is in general very difficult, if not impossible, to obtain analytically tractable expressions for the distribution function of the random variable L (even if the density function, or the characteristic function, of the vector X is known in closed form), one usually estimates quantiles of L by generating random samples from its distribution and computing the corresponding empirical quantiles. We shall exclusively deal with the so-called parametric (estimated) VaR, in the sense that we fit to observed data the parameters of a given family of distributions for the vector X of random factors, and we generate random samples from the law of X . In order to obtain a sample from the law of L we should know the functional relation between L and X . For a linear portfolio (roughly, a portfolio without derivative instruments), one simply has $L = \langle w, X \rangle$, where $w \in \mathbb{R}^d$. In the more interesting

*It might be better to say that they do, but in terms of hypergeometric functions.

case of a portfolio containing derivatives, one has $L = f(X)$, where $f : \mathbb{R}^d \rightarrow \mathbb{R}$ is a nonlinear function. Unless the derivatives in the portfolio are very simple, the function f may not admit a closed-form representation, or could just be obtained by nontrivial numerical procedures, that would have to be carried out for each random sample of X . For this reason one usually relies on approximations of the function f of the form

$$L \approx f(0) + \langle f'(0), X \rangle + \frac{1}{2} \langle f''(0)X, X \rangle,$$

which is obviously motivated by the second-order Taylor expansion of the function $\mathbb{R}^d \ni x \mapsto f(x)$ around zero. The values of $f'(0)$ and $f''(0)$ are in general determined by the so-called greeks (in this case, Delta, Gamma and Theta) of the derivatives in the portfolio. Note that in the above approximation the possible dependence of f on time can be taken into account by including time in the set of risk factors.

The analytic computation, or just approximation, of quantiles of quadratic forms in random vectors (other than Gaussian) is in general a very difficult task. Simulation is hence a viable alternative, as long as one can generate samples from the distribution of X .

We are going to perform a backtesting study on the four classes of parametric models for the distribution of risk factors introduced in Secs. 2–5, to which we refer for the corresponding estimation and simulation procedures. Value-at-Risk is just estimated by empirical quantiles of random samples of L , either obtained by full revaluation, or by the above quadratic approximation. In particular, we do not focus on efficient simulation methods for quantile estimation, but we are rather interested on the relative performance of different distributional hypotheses for risk factors, when tested on real data.

Let us also recall that all parametric families of multivariate laws that we fit to data are symmetric. A detailed comparison of the empirical performance of symmetric models and (some of) their asymmetric counterparts is outside the scope of the present paper, and it is left as an interesting question for future work.

7. Empirical Tests

7.1. The data set

We consider two portfolios of underlyings with quite different characteristics: portfolio A is more diversified, while portfolio B is strongly correlated. In particular, portfolio A is composed of two US stocks from each of four different industries, while portfolio B is composed of eight US stocks from a single industry.[†] While portfolio A is, in some sense, more realistic (e.g. from the point of view of an investor aiming

[†]The selected stocks are Apple, Bank of America, Chevron, Citigroup, Conoco, Microsoft, Johnson and Johnson, and Pfizer for portfolio A and American Express, Banco Santander, Bank of America, Barclays, Citigroup, JP Morgan Chase, U.S. Bancorp, and Wells Fargo for portfolio B .

at holding a reasonably diversified portfolio), portfolio B is constructed as a “stress test” portfolio with high tail thickness and (potentially) high tail dependence.

The raw price series are freely available on the Internet, and the returns are calculated as daily log-differences.[‡] The data set covers the time period from 2-Jan-1991 through 31-Dec-2008.

Let us provide a few descriptive statistics of the data set. Table 1 displays the sample kurtosis for each stock return. Note that all values are (much) larger

Table 1. Descriptive statistics of financial series.

	Kurtosis	Adj. J&B
American Express	9.084	7066 (0.00)
Apple	57.786	573091 (0.00)
Banco Santander	10.040	9452 (0.00)
Bank of America	25.918	99815 (0.00)
Barclays	15.267	28907 (0.00)
Chevron	13.491	20904 (0.00)
Conoco	8.373	5529 (0.00)
Citigroup	38.374	237526 (0.00)
Johnson & Johnson	9.735	8627 (0.00)
JPM Chase	11.255	12946 (0.00)
Microsoft	8.186	5105 (0.00)
Pfizer	5.914	1632 (0.00)
U.S. Bancorp	22.514	72736 (0.00)
Wells Fargo	20.869	60866 (0.00)

This table reports the sample kurtosis and the adjusted Jarque and Bera test for the log-returns of the analyzed time series.

[‡]We restrict ourselves to consider daily data for two reasons: the first and most important is that the industry and regulatory standard is to compute VaR and related risk measures on a daily basis. On the other hand, studying lower frequencies (such as weekly or monthly) would considerably decrease the size of our samples, possibly invalidating the asymptotic properties of the proposed estimators.

than 3, thus providing (rough) empirical evidence of tail-thickness of the underlying distribution. The corresponding adjusted Jarque-Bera test statistic (see e.g. [24]), is reported in the last column of Table 1 (p -values are in parentheses): for each time series the hypothesis of an underlying Gaussian distribution is rejected at 1% level.

7.2. Test portfolios

For each of the two portfolios, we construct three investment strategies adding to the basic linear portfolios containing only the eight underlyings (in equally value-weighted proportions) the following positions in options:

- NLL long 10 calls and 5 puts on each asset (“NonLinear Long”);
- NLS short 5 calls and 10 puts on each asset (“NonLinear Short”);
- NLDC short 10 down-and-out calls with barrier equal to 95% of the asset price, and short 5 cash-or-nothing put with cash payoff equal to the strike price (“NonLinear Down and Cash”).

All options are European, at-the-money, and with time to expiration equal to 6 months. The nonlinear part of the six test portfolios is synthetic, in the sense that option prices, unlike stock prices, are computed on the basis of the information available on the corresponding underlying and on (a proxy for) the risk-free rate, using Black-Scholes formula for standard call and put options, and its variants for barrier and binary options.[§] Even though this procedure is incompatible with the non-Gaussian distributional assumptions we are going to test, this is nonetheless common practice (see e.g. [16] for a more thorough discussion of this issue).

7.3. Backtesting

Let us now turn to the analysis of the performance of the four parametric distributions for risk factors introduced above, when applied to the (predictive, i.e. out-of-sample) estimation of Value-at-Risk. More precisely, we fit each of the multivariate laws to a subset of the time series of stock returns (using a rolling window consisting of 250 observations), and we estimate the 0.95 and 0.99 quantiles of the distribution of losses by simulation, i.e. selecting the corresponding empirical quantiles from a simulated sample. In particular, once a random sample from the distribution of X is obtained, we translate it into a random sample from the distribution of portfolio losses either by a full revaluation of the portfolio value for each sample, or by the usual delta-gamma quadratic approximation (see Sec. 6). Let $[t - \ell, t]$ denote the time period over which the parametric families of distributions are estimated, where ℓ stands for the (fixed) length of the rolling window. Denoting by VaR_t the

[§]We provide formulas for prices and sensitivities of these exotic options in Appendix B.

empirical quantiles of the simulated distribution of losses (with risk factors fitted over $[t - \ell, t]$), we form the statistic

$$\xi_{t+1} = \text{sgn}^+(L_{t+1} - \text{VaR}_t),$$

for all $t \in [\ell, T]$, where T denotes the length of the time series, L_t stands for the observed loss of portfolio value over the period $[t - 1, t]$, and where $\text{sgn}^+ x = 1$ if $x > 0$, and equals zero otherwise. This procedure produces a different set of $(\xi_t)_{\ell \leq t \leq T}$, for each combination of test portfolio, model for risk factors, quantile level (95% and 99%), and portfolio revaluation method (full vs. quadratic approximation).

To assess the accuracy of the VaR estimates, we perform a simple Proportion of Failure (PoF) test (cf. [20]), in analogy to the classical likelihood-ratio test. In particular, setting

$$\zeta = -2 \log \left(\frac{(1 - \beta)^x \beta^{(T - \ell - x)}}{p^x (1 - p)^{(T - \ell - x)}} \right), \tag{7.1}$$

where $\beta \in \{0.95, 0.99\}$,

$$x := \sum_{t=\ell+1}^T \xi_t, \quad p := \frac{x}{T - \ell},$$

Table 2. Value-at-Risk backtesting (full revaluation).

	Violations	Percentage (%)	LR
Panel A: NLL			
t -like _{95%}	204	4.76	0.54
t -like _{99%}	64	1.49	9.13*
Stable-like _{95%}	205	4.78	0.44
Stable-like _{99%}	66	1.54	10.81*
Gaussian _{95%}	191	4.45	2.79
Gaussian _{99%}	61	1.42	6.84*
Panel B: NLS			
t -like _{95%}	225	5.25	0.54
t -like _{99%}	64	1.49	9.13*
Stable-like _{95%}	223	5.20	0.36
Stable-like _{99%}	49	1.14	0.84
Gaussian _{95%}	207	4.83	0.27
Gaussian _{99%}	63	1.47	8.33*
Panel C: NLDC			
t -like _{95%}	201	4.68	0.90
t -like _{99%}	59	1.35	5.48*
Stable-like _{95%}	207	4.83	0.27
Stable-like _{99%}	47	1.10	0.39
Gaussian _{95%}	212	4.94	0.28
Gaussian _{99%}	65	1.52	9.95*

This table reports the results of a Value-at-Risk backtesting with full revaluation of the first portfolio (portfolio A). Panels A and B report the results for the long and short portfolios, respectively, while Panel C reports the results for the down-and-out and cash-or-nothing portfolio. Values marked with an asterisk indicate that the corresponding model is not reliable.

one expects ζ to be asymptotically χ^2 distributed with one degree of freedom. Therefore, the corresponding VaR model can be considered reliable with a 95% confidence level if $\zeta < \zeta_0 \approx 3.84$.

7.3.1. Portfolio A

The results of the backtesting procedure with full revaluation and with quadratic approximation for portfolio A are collected in Tables 2, 3 and Tables 4, 5, respectively, where values of x are in the first column, p in the second column, and ζ in the third column. Note that we included, for comparison, VaR estimates obtained under the assumptions that risk factors are jointly Gaussian.

As one may expect, the benchmark Gaussian approach fails at 99% confidence level for all three test portfolios. On the other hand, as far as VaR estimates at 95% confidence level are concerned, the Gaussian approach is still satisfactory. The same performance is displayed by the multivariate t -like approach. The stable-like approach instead is rejected by the PoF test only once. We may therefore say

Table 3. Value-at-Risk backtesting (full revaluation).

	Violations	Percentage (%)	LR
Panel A: NLL			
Meta- $t_{95\%}$	200	4.66	1.04
Meta- $t_{99\%}$	46	1.07	0.22
Meta-stable $_{95\%}$	223	5.20	0.36
Meta-stable $_{99\%}$	48	1.12	0.59
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	200	4.66	1.04
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	52	1.21	1.83
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	197	4.59	1.53
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	51	1.19	1.46
Panel B: NLS			
Meta- $t_{95\%}$	217	5.06	0.03
Meta- $t_{99\%}$	48	1.12	0.59
Meta-stable $_{95\%}$	233	5.43	1.65
Meta-stable $_{99\%}$	40	0.93	0.20
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	210	4.90	0.10
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	50	1.17	1.13
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	217	5.06	0.03
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	39	0.91	0.37
Panel C: NLDC			
Meta- $t_{95\%}$	219	5.11	0.10
Meta- $t_{99\%}$	48	1.12	0.59
Meta-stable $_{95\%}$	224	5.22	0.45
Meta-stable $_{99\%}$	39	0.91	0.37
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	220	5.13	0.15
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	53	1.24	2.24
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	208	4.85	0.20
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	36	0.84	1.18

This table is a continuation of Table 2. The same notation is used here.

Table 4. Value-at-risk backtesting (quadratic approximation).

	Violations	Percentage (%)	LR
Panel A: NLL			
<i>t</i> -like _{95%}	203	4.73	0.65
<i>t</i> -like _{99%}	63	1.47	8.33*
Stable-like _{95%}	204	4.75	0.54
Stable-like _{99%}	65	1.52	9.95*
Gaussian _{95%}	188	4.38	3.56
Gaussian _{99%}	61	1.42	6.84*
Panel B: NLS			
<i>t</i> -like _{95%}	228	5.32	0.89
<i>t</i> -like _{99%}	66	1.54	10.81*
Stable-like _{95%}	232	5.41	1.48
Stable-like _{99%}	49	1.14	0.84
Gaussian _{95%}	215	5.01	0.00
Gaussian _{99%}	63	1.47	8.33*
Panel C: NLDC			
<i>t</i> -like _{95%}	185	4.31	4.44*
<i>t</i> -like _{99%}	58	1.35	4.85*
Stable-like _{95%}	204	4.76	0.54
Stable-like _{99%}	45	1.05	0.10
Gaussian _{95%}	208	4.85	0.20
Gaussian _{99%}	64	1.49	9.13*

This table reports the results of a Value-at-Risk backtesting with quadratic approximation of the first portfolio (portfolio A). Panels A and B report the results for the long and short portfolios, respectively, while Panel C reports the results for the down-and-out and cash-or-nothing portfolio. Values marked with an asterisk indicate that the corresponding model is not reliable.

that our tests suggest that, between the two models constructed by multiplying the marginals of a Gaussian vector by a set of independent random variables, the stable-like approach might be preferable. It may also be tempting to infer that models with trivial tail dependence cannot adequately be used to estimate the probability of large losses of financial portfolios. As we shall see below, this conjecture does not seem to be supported by other empirical results. Meta-*t* and meta-stable both perform well, as the corresponding VaR estimates cannot be rejected for any one of the test portfolios. Since the estimated values of ν_0 and α_0 were very large for long portions of the time series, we tested also the performance of “degenerate” meta-*t* and meta-stable models, corresponding to the limiting cases $\nu_0 = \infty$ and $\alpha_0 = 2$, respectively. As it is well-known, these models correspond to certain nonlinear deterministic transformations of Gaussian laws (or, equivalently, to laws with *t*-distributed and α -stable distributed marginals, respectively, and a Gaussian copula). Somewhat surprisingly, these “degenerate” meta-*t* and meta-stable models give very accurate results on our test portfolios. Since these models do have trivial tail dependence, it is impossible to conclude, at least not with the data at hand, that models with nontrivial tail dependence should be preferable. In other words,

Table 5. Value-at-Risk backtesting (quadratic approximation).

	Violations	Percentage (%)	LR
Panel A: NLL			
Meta- $t_{95\%}$	200	4.66	1.04
Meta- $t_{99\%}$	44	1.03	0.03
Meta-stable $_{95\%}$	219	5.11	0.10
Meta-stable $_{99\%}$	45	1.05	0.10
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	200	4.66	1.04
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	52	1.21	1.83
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	193	4.50	2.32
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	50	1.17	1.13
Panel B: NLS			
Meta- $t_{95\%}$	225	5.25	0.54
Meta- $t_{99\%}$	50	1.17	1.13
Meta-stable $_{95\%}$	236	5.50	2.22
Meta-stable $_{99\%}$	42	0.98	0.02
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	220	5.13	0.15
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	52	1.21	1.83
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	227	5.29	0.77
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	44	1.03	0.03
Panel C: NLDC			
Meta- $t_{95\%}$	216	5.04	0.01
Meta- $t_{99\%}$	46	1.07	0.22
Meta-stable $_{95\%}$	223	5.20	0.36
Meta-stable $_{99\%}$	41	0.96	0.08
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	212	4.94	0.03
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	49	1.14	0.84
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	200	4.66	1.04
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	35	0.82	1.56

This table is a continuation of Table 4. The same notation is used here.

our empirical result seem to imply that, for portfolios whose underlyings are not (jointly) strongly dependent, the sophistication of models allowing for both heavy tails and tail dependence might not be indispensable.

Completely analogous observations could be made for the estimates of VaR obtained by the delta-gamma quadratic approximation of portfolio losses, for which we refer to Tables 4 and 5. As in the case of full revaluation, the Gaussian approach performs remarkably well at the 95% confidence level. In this respect, it is probably worth recalling that obtaining the quantiles of a quadratic form in Gaussian vectors is particularly simple and can be done with very little computational effort. In this sense, the classical quadratic approximation with Gaussian risk factors could still be regarded as a useful tool.

7.3.2. Portfolio B

The empirical results in the previous subsection, as already observed, do not offer a decisive argument in favor of models featuring both heavy tails and non-trivial tail dependence. For this reason it is interesting to perform a back-testing analysis on

Table 6. Value-at-Risk backtesting (full revaluation).

	Violations	Percentage (%)	LR
Panel A: NLL			
t -like _{95%}	217	5.06	0.03
t -like _{99%}	72	1.68	16.59*
Stable-like _{95%}	232	5.41	1.48
Stable-like _{99%}	88	2.05	36.77*
Gaussian _{95%}	208	4.85	0.20
Gaussian _{99%}	70	1.63	14.55*
Panel B: NLS			
t -like _{95%}	217	5.06	0.03
t -like _{99%}	59	1.37	5.48*
Stable-like _{95%}	230	5.36	1.17
Stable-like _{99%}	507	1.33	4.26*
Gaussian _{95%}	203	4.73	0.65
Gaussian _{99%}	64	1.63	9.13*
Panel C: NLDC			
t -like _{95%}	253	5.90	6.93*
t -like _{99%}	78	1.82	23.37*
Stable-like _{95%}	250	5.83	5.92*
Stable-like _{99%}	68	1.59	12.62*
Gaussian _{95%}	250	5.83	5.92*
Gaussian _{99%}	85	1.98	32.50*

This table reports the results of a Value-at-Risk backtesting with full revaluation of the second portfolio (portfolio B). Panels A and B report the results for the long and short portfolios, respectively, while Panel C reports the results for the down-and-out and cash-or-nothing portfolio. Values marked with an aster.

the “stress test” portfolio B , whose underlyings are (presumably) heavy tailed and strongly dependent.

The results with full revaluation and with quadratic approximation for portfolio B are collected in Tables 6–9, respectively, where we included again, for comparison purposes, VaR estimates obtained under the assumptions that risk factors are jointly Gaussian.

One can see immediately (cf. Tables 6 and 8) that both the t -like and the stable-like models, as well as the standard Gaussian model, are unreliable at the 99% confidence level for all portfolios, and even at the 95% confidence level for the test portfolio NLDC containing exotic options. This could be interpreted as empirical evidence that these classes of models, all of which have trivial tail dependence, are not adequate to estimate the probability of large losses, especially for highly nonlinear portfolios. It does not seem possible, however, to assert that the meta- t and meta-stable models, both of which feature heavy tails and non-trivial tail dependence, are superior in terms of their empirical performance to their degenerate counterparts (i.e. the meta- t and meta-stable models with $\nu_0 = \infty$ and $\alpha_0 = 2$, respectively), which allow heavy tails but no tail dependence. In fact (cf. Tables 7 and 9), the performance of all meta- t and meta-stable models is comparable for the

Table 7. Value-at-Risk backtesting (full revaluation).

	Violations	Percentage (%)	LR
Panel A: NLL			
Meta- $t_{95\%}$	227	5.50	2.22
Meta- $t_{99\%}$	51	1.05	0.10
Meta-stable $_{95\%}$	236	5.17	0.28
Meta-stable $_{99\%}$	45	0.96	0.08
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	223	5.20	0.36
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	54	1.26	2.69
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	221	5.15	0.21
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	46	1.07	0.22
Panel B: NLS			
Meta- $t_{95\%}$	223	5.20	0.36
Meta- $t_{99\%}$	44	1.03	0.03
Meta-stable $_{95\%}$	241	5.62	3.34
Meta-stable $_{99\%}$	41	0.99	0.08
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	219	5.10	0.10
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	49	1.14	0.84
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	223	5.20	0.36
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	44	1.03	0.03
Panel C: NLDC			
Meta- $t_{95\%}$	259	6.04	9.18*
Meta- $t_{99\%}$	68	1.59	12.62*
Meta-stable $_{95\%}$	266	6.20	12.18*
Meta-stable $_{99\%}$	59	1.38	5.48*
Meta- $t_{95\%}$ ($\nu_0 = \infty$)	255	5.95	7.65*
Meta- $t_{99\%}$ ($\nu_0 = \infty$)	69	1.61	13.57*
Meta-stable $_{95\%}$ ($\alpha_0 = 2$)	252	5.88	6.59*
Meta-stable $_{99\%}$ ($\alpha_0 = 2$)	49	1.14	0.84

This table is a continuation of Table 6. The same notation is used here.

Table 8. Value-at-Risk backtesting (quadratic approximation).

	Violations	Percentage (%)	LR
Panel A: NLL			
t -like $_{95\%}$	214	4.99	0.00
t -like $_{99\%}$	71	1.66	15.55*
Stable-like $_{95\%}$	229	5.33	1.02
Stable-like $_{99\%}$	84	1.96	31.12*
Gaussian $_{95\%}$	207	4.83	0.27
Gaussian $_{99\%}$	69	1.61	13.57*
Panel B: NLS			
t -like $_{95\%}$	221	5.15	0.21
t -like $_{99\%}$	65	1.52	9.95*
Stable-like $_{95\%}$	242	5.64	3.60
Stable-like $_{99\%}$	57	1.33	4.26*
Gaussian $_{95\%}$	206	4.80	0.35
Gaussian $_{99\%}$	68	1.59	12.62*

Table 8. (Continued)

	Violations	Percentage (%)	LR
Panel C: NLDC			
t -like _{95%}	244	5.69	4.13*
t -like _{99%}	76	1.77	21.01*
Stable-like _{95%}	243	5.67	3.86*
Stable-like _{99%}	65	1.52	9.95*
Gaussian _{95%}	243	5.87	3.86*
Gaussian _{99%}	83	1.94	29.77*

This table reports the results of a Value-at-Risk backtesting with quadratic approximation of the second portfolio (portfolio B). Panels A and B report the results for the long and short portfolios, respectively, while Panel C reports the results for the down-and-out and cash-or-nothing portfolio. Values marked with an asterisque indicate that the corresponding model is not reliable.

Table 9. Value-at-Risk backtesting (quadratic approximation).

	Violations	Percentage (%)	LR
Panel A: NLL			
Meta- t _{95%}	224	5.22	0.45
Meta- t _{99%}	51	1.19	1.46
Meta-stable _{95%}	233	5.43	1.65
Meta-stable _{99%}	45	1.05	0.10
Meta- t _{95%} ($\nu_0 = \infty$)	220	5.13	0.15
Meta- t _{99%} ($\nu_0 = \infty$)	53	1.24	2.24
Meta-stable _{95%} ($\alpha_0 = 2$)	221	5.15	0.21
Meta-stable _{99%} ($\alpha_0 = 2$)	44	1.03	0.03
Panel B: NLS			
Meta- t _{95%}	228	5.32	0.89
Meta- t _{99%}	48	1.12	0.59
Meta-stable _{95%}	249	5.80	5.60*
Meta-stable _{99%}	45	1.05	0.10
Meta- t _{95%} ($\nu_0 = \infty$)	227	5.29	0.77
Meta- t _{99%} ($\nu_0 = \infty$)	54	1.26	2.69
Meta-stable _{95%} ($\alpha_0 = 2$)	233	5.43	1.65
Meta-stable _{99%} ($\alpha_0 = 2$)	45	1.05	0.10
Panel C: NLDC			
Meta- t _{95%}	253	5.90	6.93*
Meta- t _{99%}	66	1.54	10.81*
Meta-stable _{95%}	262	6.11	10.42*
Meta-stable _{99%}	56	1.31	3.70
Meta- t _{95%} ($\nu_0 = \infty$)	250	5.83	5.92*
Meta- t _{99%} ($\nu_0 = \infty$)	66	1.54	10.81*
Meta-stable _{95%} ($\alpha_0 = 2$)	244	5.69	4.13*
Meta-stable _{99%} ($\alpha_0 = 2$)	46	1.07	0.22

This table is a continuation of Table 8. The same notation is used here.

vanilla portfolios NLL and NLS, independently of having non-trivial dependence or not (with the exception of the meta-stable model which is rejected in one case, see Table 9). The picture changes drastically for the exotic portfolio NLDC, for which the meta- t , meta-stable and degenerate meta- t models behave poorly. On the other hand, surprisingly, the degenerate meta-stable model displays a good performance at the 99% level. It appears to be very difficult, if not impossible, to give an explanation for this observation.

8. Concluding Remarks

Let $G \sim N(0, Q)$, and consider the random vector X obtained from G by variance-mixture where S is a positive random variable independent of G , that is $X := S^{1/2}G$. If S has the distribution of the reciprocal of a χ^2 -distributed random variable resp. of an $\alpha/2$ -stable subordinator, we obtain the class of multivariate t resp. symmetric α -stable sub-Gaussian laws. Plenty of other distributions obtained by variance mixing of a Gaussian measure on \mathbb{R}^d that have appeared in the literature for different purposes, including of course the modeling of financial risk factors. Similarly, many generalizations of variance mixing have appeared in the literature, and it is evident that endless variations on the theme are possible. In this article we have limited ourselves to two special cases of two possible extensions. Namely, variance mixture models could be generalized setting $X = T^{1/2}G$, where T is a positive-definite random matrix (cf. [3] for related classes of distributions), or one could consider nonlinear images of (the law of) $S^{1/2}G$, such as $X = \phi(S^{1/2}G)$, with ϕ e.g. a (deterministic) injective function. In particular, if T is a diagonal random matrix with independent entries and, for each $i = 1, \dots, d$, T_{ii} is distributed like the reciprocal of a χ^2 random variable with ν_i degrees of freedom, we obtain the class of t -like laws of Sec. 3. A completely analogous observation holds for the stable-like laws of Sec. 2. Similarly, for particular choices of functions ϕ , we obtain meta-elliptical distributions, of which meta- t and meta-stable are just special cases.

We have considered t -like and stable-like laws because of their simplicity and ease of estimation and simulation, whereas the two specific instances of meta-elliptical distribution have been considered for their seemingly widespread use (at least as regards the meta- t law), especially in connection with applications of copula methods.

Let us mention other possible generalizations of the classical variance-mixture approach that have recently appeared in the literature, without any claim of completeness (which, as already mentioned above, would not be possible). Assume $X = T^{1/2}G$, with the same notation as above, where T is diagonal, $T_{ii} = F_i^{\leftarrow}(U)$ for each $i = 1, \dots, d$, U is a uniformly distributed random variable independent of G , and F_i , $i = 1, \dots, d$ are cumulative distribution functions. In the particular case in which each F_i is the cumulative distribution function of the reciprocal of a (rescaled) χ^2 -distributed random variable with ν_i degrees of freedom, we obtain the class of grouped t -distributions (see e.g. [1, 2, 9]). Of course nothing prevents

us from considering arbitrary distributions functions as F_i 's. By combining this construction with a nonlinear map, so that $X = \phi(T^{1/2}G)$, one could for instance construct laws with the dependence structure of a grouped- t law and with arbitrary marginals (see e.g. [8] for the so-called grouped- t copula). The reader clearly understand that the possibilities for constructing multivariate laws by any of these methods, or a combination thereof, are endless.

Our empirical results suggest that, among the infinitely many possible models for risk factors with non-trivial tail dependence, both classes of meta- t and meta-stable laws offer good performance, at least on reasonably diversified portfolios. Nevertheless, there is weaker evidence that, under "extreme" conditions such as those characterizing our portfolio B , these classes of models can perform well in highly non-linear situations. We believe that the most important message of our paper is that it is indeed worth taking into account that "classical" multivariate Gaussian laws with changed marginals (in particular α -stable) might perform surprisingly well. Such (relatively) simple models are undoubtedly attractive from the viewpoint of practical implementation, as their estimation, simulation and quadratic approximation are very easy and "light" in terms of computational resources.

Appendix A. Densities of Some Random Vectors

A.1. Densities of a class of images of random vectors

Let $\phi \in C^1(\mathbb{R}^d \rightarrow \mathbb{R}^d)$ be an injective function of the type

$$\phi : (x_1, x_2, \dots, x_d) \mapsto (\phi_1(x_1), \phi_2(x_2), \dots, \phi_d(x_d)),$$

for functions $\phi_k : \mathbb{R} \rightarrow \mathbb{R}$, $k = 1, \dots, d$.

Proposition A.1. *Let X be a d -dimensional random vector with density p . Then the density of $\phi(X)$ is the function*

$$(y_1, \dots, y_k) \mapsto \frac{p(\phi_1^{-1}(y_1), \dots, \phi_d^{-1}(y_d))}{\phi_1'(\phi_1^{-1}(y_1)) \cdots \phi_d'(\phi_d^{-1}(y_d))}.$$

Proof. By the multidimensional change of variable formula for Lebesgue integrals and by the inverse function theorem we have, for any measurable set A ,

$$\begin{aligned} \mathbb{P}(\phi(X) \in A) &= \mathbb{P}(X \in \phi^{-1}(A)) = \int_{\phi^{-1}(A)} p(x) dx \\ &= \int_A p(\phi^{-1}(y)) |\det D\phi^{-1}(y)| dy = \int_A p(\phi^{-1}(y)) \frac{1}{|\det D\phi(\phi^{-1}(y))|} dy \\ &= \int_A \frac{p(\phi_1^{-1}(y_1), \dots, \phi_d^{-1}(y_d))}{\phi_1'(\phi_1^{-1}(y_1)) \cdots \phi_d'(\phi_d^{-1}(y_d))} dy_1 \cdots dy_d, \end{aligned}$$

thus proving the claim. □

A.2. Densities of sub-Gaussian α -stable vectors

Let A and G be as in Sec. 4, and set $Y = A^{1/2}G$. Let us recall that the law of G admits the density

$$\gamma_Q(x) = \frac{1}{(2\pi)^{d/2}} \frac{1}{(\det Q)^{1/2}} \exp\left(-\frac{1}{2}\langle Q^{-1}x, x \rangle\right).$$

Therefore, for any constant $c \in \mathbb{R}$, we have

$$\gamma_{cQ}(x) = \frac{1}{(2\pi)^{d/2}} \frac{1}{(\det Q)^{1/2}} e^{-d/2} \exp\left(-\frac{1}{2c}\langle Q^{-1}x, x \rangle\right).$$

For any measurable set B , recalling that A and G are independent, we have

$$\mathbb{P}(Y \in B) = \int_0^\infty \mathbb{P}(a^{1/2}G \in B) p_A(a) da = \int_0^\infty \int_B \gamma_{aQ}(x) dx p_A(a) da,$$

where p_A denotes the density of the law of A . Therefore the law of Y admits a density $h(x) = g(\langle Q^{-1}x, x \rangle)$, where

$$g(z) := \frac{1}{(2\pi)^{d/2}} \frac{1}{(\det Q)^{1/2}} \int_0^\infty a^{-d/2} e^{-z/(2a)} p_A(a) da, \quad z > 0.$$

In fact g can be extended by continuity at $z = 0$, since the above integral with the exponential term suppressed is well-defined, e.g. using the series expansion at zero for the density of $S_{\alpha/2}(1, 1, 0)$ of [27, p. 99].

Appendix B. Prices and Sensitivities of Some Exotic Options

Throughout this appendix we place ourselves in a standard Black-Scholes model with one “underlying” stock, whose price process is denoted by S_t , $0 \leq t \leq T$, and whose (constant) volatility is denoted by σ . The risk-free rate will be denoted by r . We shall consider options written on the stock, denoting the exercise time by T , the strike price by K , and the barrier by H .

In the following table we collect the definitions, in terms of their payoff, of some barrier and binary options.

Name	Payoff	
Down-and-In call	$\max(S_T - K, 0)$	if $\min_{0 \leq t \leq T} S_t \leq H$
Down-and-In put	$\max(K - S_T, 0)$	if $\min_{0 \leq t \leq T} S_t \leq H$
Down-and-Out call	$\max(S_T - K, 0)$	if $\min_{0 \leq t \leq T} S_t \geq H$
Down-and-Out put	$\max(K - S_T, 0)$	if $\min_{0 \leq t \leq T} S_t \geq H$
Up-and-In call	$\max(S_T - K, 0)$	if $\max_{0 \leq t \leq T} S_t \geq H$
Up-and-In put	$\max(K - S_T, 0)$	if $\max_{0 \leq t \leq T} S_t \geq H$
Up-and-Out call	$\max(S_T - K, 0)$	if $\max_{0 \leq t \leq T} S_t \leq H$
Up-and-Out put	$\max(K - S_T, 0)$	if $\max_{0 \leq t \leq T} S_t \leq H$
Cash-or-Nothing call	1	if $S_T \geq K$
Cash-or-Nothing put	1	if $S_T \leq K$

We shall use C_{di} and P_{di} to denote the price (at time zero) of a down-and-in call and a down-and-in put, respectively. Completely analogous notation will be used for the remaining options, replacing the subscripts accordingly. The price of plain European call and put options will be denoted by C_{BS} and P_{BS} , respectively. The price at time zero of a European call option with strike K and exercise time T , written on an underlying whose price at time zero is S_0 , will be denoted by $C_{BS}(S_0, K, T)$. The corresponding notation will be also used for European put options.

Setting

$$\lambda = \frac{2r}{\sigma^2} - 1, \quad m = \frac{r}{\sigma^2} + \frac{1}{2}$$

and assuming $H < K$, one has (see e.g. [7]),

$$\begin{aligned} C_{di} &= H^\lambda S_0^{-\lambda} C_{BS}(H^2 S_0^{-1}, K, T), \\ P_{di} &= C_{di} + KH^{-1} P_{BS}(S_0, H, T) - (HS_0^{-1})^{2m-2} HK^{-1} \\ &\quad \times C_{BS}(KHS_0^{-1}, K^2 H^{-1}, T), \\ C_{ui} &= C_{BS} \\ P_{ui} &= H^\lambda S_0^{-\lambda} P_{BS}(H^2 S_0^{-1}, K, T). \end{aligned}$$

By the obvious identities

$$C_{di} + C_{do} = C_{BS}, \quad C_{ui} + C_{uo} = C_{BS},$$

and the corresponding ones for put options (i.e. those obtained replacing C with P), we obtain pricing formulas for all barrier options listed in the above table. By the well-known formulas for sensitivities of European call and put options, elementary calculus yields

$$\begin{aligned} \frac{\partial C_{di}}{\partial S_0} &= -\lambda H^\lambda S_0^{-\lambda-1} C_{BS}(H^2 S_0^{-1}, K, T) - H^{\lambda+2} S_0^{-\lambda-2} \Delta_{BS}(H^2 S_0^{-1}, K, T), \\ \frac{\partial^2 C_{di}}{\partial S_0^2} &= \lambda(\lambda + 1) H^\lambda S_0^{-\lambda-2} C_{BS}(H^2 S_0^{-1}, K, T) + 2(\lambda + 1) H^{\lambda+2} S_0^{-\lambda-3} \\ &\quad \times \Delta_{BS}(H^2 S_0^{-1}, K, T) + H^{\lambda+4} S_0^{-\lambda-4} \Gamma_{BS}(H^2 S_0^{-1}, K, T), \\ \frac{\partial C_{di}}{\partial T} &= H^\lambda S_0^{-\lambda} \Theta_{BS}(H^2 S_0^{-1}, K, T). \end{aligned}$$

Similar expressions can be derived for the sensitivities of the other binary options.

Setting

$$d_1 := \frac{\log(S_0/K) + (r + \sigma^2)T}{\sigma\sqrt{T}}, \quad d_2 := d_1 - \sigma\sqrt{T},$$

we have (see e.g. [18])

$$C_{cn} = e^{-rT} \Phi(d_2), \quad P_{cn} = e^{-rT} \Phi(-d_2),$$

where $\Phi(\cdot)$ stands for the distribution function of the Gaussian law on \mathbb{R} with mean zero and unit variance. The sensitivities of binary options are just an exercise in elementary calculus. Let us include, for the sake of completeness, the sensitivities of the cash-or-nothing put, which is used in our portfolios:

$$\begin{aligned} \frac{\partial P_{cn}}{\partial S_0} &= \frac{-e^{-rT} \Phi'(-d_2)}{\sigma S_0 \sqrt{T}}, \\ \frac{\partial^2 P_{cn}}{\partial S_0^2} &= \frac{e^{-rT} \Phi'(-d_2)}{\sigma S_0^2 \sqrt{T}} + \frac{-d_2 e^{-rT-d_2^2/2}}{\sigma^2 T S_0^2 \sqrt{2\pi}}, \\ \frac{\partial P_{cn}}{\partial T} &= -r e^{-rT} \Phi(-d_2) + \frac{r e^{-rT} \Phi'(-d_2) \log(S_0/K)}{2\sigma T^{3/2}} - \frac{r - \sigma^2/2}{\sigma \sqrt{T}}. \end{aligned}$$

Acknowledgments

We are very grateful to two anonymous referees for a careful reading of previous versions of this paper. Their corrections and useful suggestions led to an improved version and a better presentation of our results.

References

- [1] K. Banachewicz and A. van der Vaart, Tail dependence of skewed grouped t -distributions, *Statist. Probab. Lett.* **78**(15) (2008) 2388–2399.
- [2] K. Banachewicz and A. van der Vaart, Corrigendum to: “Tail dependence of skewed grouped t -distributions,” *Statist. Probab. Lett.* **79**(15) (2009) 1731.
- [3] O. E. Barndorff-Nielsen and V. Pérez-Abreu, Extensions of type G and marginal infinite divisibility, *Teor. Veroyatnost. i Primenen.* **47**(2) (2002) 301–319.
- [4] J. Berkowitz and J. O’Brien, How accurate are value-at-risk models at commercial banks?, *J. Finance* **57**(3) (2002) 1093–1111.
- [5] R. Blattberg and N. Gonedes, A comparison of the stable and student distributions as statistical models of stock prices, *J. Business* **47** (1974) 244–280.
- [6] S. Cambanis, S. Huang and G. Simons, On the theory of elliptically contoured distributions, *J. Multivariate Anal.* **11**(3) (1981) 368–385.
- [7] P. Carr, K. Ellis and V. Gupta, Static hedging of exotic options, *J. of Finance* **53**(3) (1998) 1165–1190.
- [8] S. Daul, E. De Giorgi, F. Lindskog and A. McNeil, The grouped t -copula with an application to credit risk, *Risk* **16**(11) (2003) 73–76.
- [9] S. Demarta and A. J. de McNeil, The t copula and related copulas, *Internat. Statist. Rev.* **73** (2005) 111–129.
- [10] D. Duffie and J. Pan, An overview of value at risk, *J. Deriv.* **4** (1997) 7–72.
- [11] D. Duffie and J. Pan, Analytical value-at-risk with jumps and credit risk, *Finance Stoch.* **5**(2) (2001) 155–180.
- [12] E. Fama, The behaviour of stock market prices, *J. of Business* **38**(1) (1965) 34–105.
- [13] E. Fama, Portfolio analysis in a stable Paretian market, *Management Sci.* **3** (1965) 404–419.

- [14] H.-B. Fang, K.-T. Fang and S. Kotz, The meta-elliptical distributions with given marginals, *J. Multivariate Anal.* **82**(1) (2002) 1–16.
- [15] J. C. Frain, *Value at Risk (var) and the Alpha-Stable Distribution*, Trinity economics Papers tep0308 (Trinity College Dublin, Department of Economics, May 2008).
- [16] P. Glasserman, Ph. Heidelberger and P. Shahabuddin, Portfolio value-at-risk with heavy-tailed risk factors, *Math. Finance* **12**(3) (2002) 239–269.
- [17] D. Hogben, R. S. Pinkham and M. B. Wilk, The moments of the non-central t -distribution, *Biometrika* **48** (1961) 465–468.
- [18] J. Hull, *Options, Futures, and Other Derivatives* (Prentice Hall, 2008).
- [19] S. Kotz and S. Nadarajah, *Multivariate t Distributions and Their Applications* (Cambridge University Press, Cambridge, 2004).
- [20] P. H. Kupiec, Techniques for verifying the accuracy of risk management models, *J. Deriv.* **3** (1995) 73–84.
- [21] F. Lindskog, A. McNeil and U. Schmock, *Kendall's Tau for Elliptical Distributions*, Credit risk: Measurement, evaluation and management (Physica Verlag, 2003), pp. 149–156.
- [22] S. T. Rachev and S. Mittnik, *Stable Paretian Models in Finance* (John Wiley and Sons, NY, 2000).
- [23] G. Samorodnitsky and M. S. Taqqu, *Stable Non-Gaussian Random Processes* (Chapman & Hall, New York, 1994).
- [24] C. M. Urzua, On the correct use of omnibus tests for normality, *Economics Letters* **53**(3) (1996) 247–251.
- [25] A. van Aubel and W. Gawronski, Analytic properties of noncentral distributions, *Appl. Math. Comput.* **141**(1) (2003) 3–12.
- [26] A. W. van der Vaart, *Asymptotic Statistics* (Cambridge UP, Cambridge, 1998).
- [27] V. M. Zolotarev, *One-Dimensional Stable Distributions*, Translations of Mathematical Monographs (American Mathematical Society, Providence, RI, 1986).