Stefan Mittnik

VaR–implied Tail–correlation Matrices

VaR–implied Tail–correlation Matrices

Stefan Mittnik[*]

Chair of Financial Econometrics, Department of Statistics, and

Center for Quantitative Risk Analysis, Ludwig–Maximilians–Universität München

November 10, 2013

Abstract

Empirical evidence suggests that asset returns correlate more strongly in bear markets than conventional correlation estimates imply. We propose a method for determining complete tail–correlation matrices based on Value–at–Risk (VaR) estimates. We demonstrate how to obtain more efficient tail–correlation estimates by use of overidentification strategies and how to guarantee positive semidefiniteness, a property required for valid risk aggregation and Markowitz–type portfolio optimization. An empirical application to a 30–asset universe illustrates the practical applicability and relevance of the approach in portfolio management.

Keywords: Downside risk; estimation efficiency; portfolio optimization; positive semidefiniteness; Solvency II; Value–at–Risk

JEL classification: C1, G11

---

[*]*Address for correspondence:* Chair of Financial Econometrics, Center for Quantitative Risk Analysis, Ludwig Maximilians University Munich, Akademiestrasse 1/I, 80799 Munich, Germany; Phone: +49 (0)89 2180-3224; Email: finmetrics@stat.uni-muenchen.de

1 Introduction

The correlation between financial assets plays also a central role in applied and theoretical finance. A frequent concern is that correlations increase during periods of high market stress.[1] As a consequence, portfolio strategies, risk–management practices and regulation focus increasingly on tail–risk, such as the Value–at–Risk (VaR), and tail–dependence measures. Tail correlations play, for example, a central role in the proposed European *Solvency II* regulation for the insurance industry (European Commission, 2007). The *Standard Formula*, determining insurers' risk–capital requirements, is based on a VaR measure at the 99.5% confidence level and requires that correlations for aggregating risk components should be specified for that tail area.

To derive correlation estimates that are compatible with VaR–type risk measures, Campbell et al. (2002) proposed a VaR–implied correlation estimator, which measures correlational dependence in the VaR–specific tail area of the distribution. Given the VaR estimates for two assets and that of a portfolio built from these two assets (all for the same VaR–confidence level), they derive the correlation coefficient associated with the particular VaR confidence level. To obtain an estimate of the complete VaR–implied tail–correlation matrix for an n–asset universe, coefficient estimates are derived—pair by pair—for each of the $n(n-1)/2$ asset pairs.

This pairwise approach has several drawbacks. In case of n assets, relying exclusively on $n(n-1)/2$ *two*–asset portfolios ignores correlational information contained in multi–asset portfolio–VaRs and is inefficient. More importantly, pairwise derivation does not guarantee that VaR–implied correlations give rise to a proper correlation matrix, as the estimates may lie outside the $[-1, +1]$–interval. Even if there is no interval violation, the resulting matrix may not be positive semidefinite—a requirement for valid risk aggregation and mean–variance portfolio optimization. Whereas interval violations can be fixed via truncation, there is no obvious strategy for imposing positive semidefiniteness when estimating tail–correlation matrices element–by–element.

---

[1]Studies supporting this hypothesis are, for example, Erb et al. (1994), Longin and Solnik (1995), Karolyi and Stulz (1996), Silvapulle and Granger (2001), Longin and Solnik (2001), Ang and Bekaert (2002), Ang and Chen (2002), Butler and Joaquin (2002), Bae et al. (2003), Das and Uppal (2004), Hong et al. (2007), Okimoto (2008), and Haas and Mittnik (2009). Possible explanations are that returns follow non–normal, fat–tailed and asymmetric distributions, so that linear correlation varies across the support of the distribution (Campbell et al., 2008), or that dependence structures are state–dependent (Ang and Chen (2002), Haas et al. (2004), Haas and Mittnik (2009)).

In the following, we summarize the pairwise approach for deriving VaR–implied correlations and outline the new method, discussing exactly and overidentified as well as constrained variants. We present the results of a Monte Carlo study comparing the properties of alternative strategies. A empirical application to the 30–asset universe of DAX stocks illustrates the practical feasibility and relevance of the proposed method for measuring complex dependence structures and portfolio management.

2 Pairwise Approach

Let r_1 and r_2 denote the returns of two assets and $r_p = w_1 r_1 + w_2 r_2$ the return of a portfolio with weights w_1 and w_2, $w_1 + w_2 = 1$. Moreover, let σ_i^2 and $q_{\alpha,i}$, $i = 1, 2, p$, respectively, denote the corresponding return variance and α–quantile, i.e., the (negative) VaR at confidence level $100 \times (1 - \alpha)\%$. If r_1 and r_2 follow an elliptical distribution,[2] we have

$$\sigma_p^2 = w_1^2 \sigma_1^2 + w_2^2 \sigma_2^2 + 2 w_1 w_2 \sigma_1 \sigma_2 \rho_{12}. \tag{1}$$

Assuming, for simplicity sake, that return expectations are zero or that the return data have been de-meaned, then $q_{\alpha,i} = \xi_\alpha \sigma_i$, $i = 1, 2, p$, where ξ_α denotes the α–quantile of the standardized marginal distribution. Substituting, in (1), $\sigma_i = q_{\alpha,i}/\xi_\alpha$ and multiplying both sides by ξ_α^2 gives

$$q_{\alpha,p}^2 = w_1^2 q_{\alpha,1}^2 + w_2^2 q_{\alpha,2}^2 + 2 w_1 w_2 q_{\alpha,1} q_{\alpha,2} \rho_{12}. \tag{2}$$

Campbell et al. (2002) and also Cotter and Longin (2007) use (2) to solve for the VaR–implied correlation via[3]

$$\rho_{\alpha,12} = \frac{q_{\alpha,p}^2 - w_1^2 q_{\alpha,1}^2 - w_2^2 q_{\alpha,2}^2}{2 w_1 w_2 q_{\alpha,1} q_{\alpha,2}}. \tag{3}$$

For elliptical distributions, $\rho_{\alpha,12}$ will be invariant with respect to weights and confidence levels. Otherwise, VaR–implied correlations may vary as weights or confidence levels change. In this case, an estimate derived for a specific weight/confidence–level combination can be viewed as a local elliptical, i.e., correlational, approximation.

---

[2]The multivariate normal and Student's t distributions are prominent members of the elliptical family. For details on elliptical distributions, see, for example, Cambanis et al. (1981).

[3]It is evident from (3) that the estimator only works for α–quantiles away from the center. Otherwise, $q_{\alpha,1}$ and $q_{\alpha,2}$ will be close to zero, so that the ratio becomes unstable or even undefined.

Drawbacks of estimator (3) are that it does not guarantee that $\rho_{\alpha,12}$ satisfies the interval constraint $|\rho_{\alpha,12}| \leq 1$ and that the resulting correlation matrix may fail to be positive semidefinite (PSD). This may be due to VaR not being a coherent risk measure, in the sense of Artzner et al. (1999), potentially lacking subadditivity in the presence of non–elliptical distributions.

As the simulation results below will show, even if the data are drawn from an elliptical distribution, finite–sample variation may easily cause interval violations. In this situation, a truncated version of (3) can be applied, i.e.,[4]

$$\rho_{\alpha,12} = \begin{cases} +1, & \text{if } q_{\alpha,p} \geqslant w_1 q_{\alpha,1} + w_2 q_{\alpha,2} \\ -1, & \text{if } q_{\alpha,p} \leqslant |w_1 q_{\alpha,1} - w_2 q_{\alpha,2}| \\ \dfrac{q_{\alpha,p}^2 - w_1^2 q_{\alpha,1}^2 - w_2^2 q_{\alpha,2}^2}{2 w_1 w_2 q_{\alpha,1} q_{\alpha,2}}, & \text{otherwise.} \end{cases} \tag{4}$$

Being highly susceptible to interval and PSD violations, the practical usefulness of the pairwise estimation is limited. The approach proposed next tackles these deficits by jointly estimating all correlation–matrix elements. It allows to reduce sampling variation and, with that, the frequency and severity of violations by means of overidentification. Although the joint approach will reduce violations, it will not necessarily eliminate them. Strategies to do so will be presented.

3 Joint Estimation

3.1 The Approach

Given an n–asset portfolio with weights w_i, $i = 1, \ldots, n$, $\sum_{i=1}^{n} w_i = 1$, denote the α–quantile of asset i, dropping subscript α, simply by q_i. Then, the α–quantile, q_p, of the portfolio return satisfies

$$q_p^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_i w_j q_i q_j \rho_{ij}, \tag{5}$$

with $\rho_{ii} = 1$, $i = 1, \ldots, n$. Different from the two–asset case, where we can uniquely derive ρ_{ij} from q_i, q_j and q_p, (5) does not allow unique determination of the correlation parameters, as there are altogether $n(n-1)/2$ unknown correlation coefficients. Relationship (5) holds, however, for any hypothetical weight vector, for which we can empirically derive the corresponding portfolio returns and quantiles.

---

[4]Condition $q_{\alpha,p} \geqslant w_1 q_{\alpha,1} + w_2 q_{\alpha,2}$ in (4) implies superadditivity in the sense of Artzner et al. (1999). Analogously, condition $q_{\alpha,p} \leqslant |w_1 q_{\alpha,1} - w_2 q_{\alpha,2}|$ may be referred to as "supersubtractivity."

Let $\boldsymbol{R}$ be the $n \times n$ tail–correlation matrix, $\boldsymbol{q} = (q_1, \ldots, q_n)'$ the $n \times 1$ vector of asset quantiles, and $\boldsymbol{w} = (w_1, \ldots, w_n)'$ the vector of weights. Then, (5) can be written as

$$q_p^2 = (\boldsymbol{q} \odot \boldsymbol{w})' \boldsymbol{R} (\boldsymbol{q} \odot \boldsymbol{w}), \tag{6}$$

where $\odot$ denotes the Schur product.[5] Relationship (6) is linear in $\boldsymbol{R}$, so that, given $n(n-1)/2$ linearly independent analogues, we can uniquely solve for as many unknowns. To set up the system of equations, we bring all $\rho_{ii} = 1$, $i = 1, \ldots, n$, to the left, i.e.,

$$\tilde{q}_p = q_p^2 - \sum_{i=1}^n q_i^2 w_i^2 = (\boldsymbol{q} \odot \boldsymbol{w})'(\boldsymbol{R} - \boldsymbol{I})(\boldsymbol{q} \odot \boldsymbol{w}). \tag{7}$$

Quantity $\tilde{q}_p = q_p^2 - \sum_{i=1}^n q_i^2 w_i^2$ represents the (squared) "correlational excess VaR;" i.e., if, for given weights, the correlation structure is such that positive (negative) correlations "outweigh" the negative (positive) ones, $\tilde{q}_p$ will be positive (negative). If returns are uncorrelated, $\tilde{q}_p = 0$.

Let "vecl" denote the vectorization operator, which stacks all elements *below* the main diagonal of a square matrix into a column vector.[6] There exists a unique duplication matrix, $\boldsymbol{D}$, of dimension $n^2 \times n(n-1)/2$ whose entries consist of zeros and ones, such that $\mathrm{vec}(\boldsymbol{R} - \boldsymbol{I}) = \boldsymbol{D}\mathrm{vecl}(\boldsymbol{R} - \boldsymbol{I}) = \boldsymbol{D}\mathrm{vecl}(\boldsymbol{R})$, where "vec" denotes the conventional vectorization operator. Then, using $\mathrm{vec}(\boldsymbol{ABC}) = (\boldsymbol{C}' \otimes \boldsymbol{A}) \mathrm{vec}(\boldsymbol{B})$, with $\otimes$ denoting the Kronecker product, (7) can be rewritten as

$$\tilde{q}_p = [(\boldsymbol{q} \odot \boldsymbol{w})' \otimes (\boldsymbol{q} \odot \boldsymbol{w})'] \ \mathrm{vec}(\boldsymbol{R} - \boldsymbol{I}) = [(\boldsymbol{q} \odot \boldsymbol{w})' \otimes (\boldsymbol{q} \odot \boldsymbol{w})'] \boldsymbol{D}\boldsymbol{\rho}, \tag{8}$$

where the $n(n-1)/2 \times 1$ vector $\boldsymbol{\rho} = \mathrm{vecl}(\boldsymbol{R})$ collects all unique correlations in $\boldsymbol{R}$.

3.2 Exact Identification

To construct an exactly–identified system of equations, $n(n-1)/2$ linearly independent equations of type (8) are required. They can be established by applying the pairwise approach (3) to each of the $n(n-1)/2$ (i,j)–pairs. Considering, for example, all equal–weight, two–asset portfolios ($k = 2$) in a four–asset universe ($n = 4$), the pairwise approach delivers the necessary number of $m_2 = \binom{n}{k} = n!/(k!\,(n-k)!) = 6$ weight vectors $\boldsymbol{w}_i$, $i = 1, \ldots, 6$, shown in Table 1.

---

[5]I.e., if $m \times n$ matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ have typical elements a_{ij} and b_{ij}, respectively, the $m \times n$ matrix $\boldsymbol{C} = \boldsymbol{A} \odot \boldsymbol{B} = \boldsymbol{B} \odot \boldsymbol{A}$ has typical element $c_{ij} = a_{ij}b_{ij}$.

[6]The "vecl" operator is similar to the more familiar "vech" operator but omits the diagonal elements.

Table 1: Possible weight vectors for two–, three– and four–asset portfolios with equal weights in a four–asset universe.

	$\boldsymbol{w}_1$	$\boldsymbol{w}_2$	$\boldsymbol{w}_3$	$\boldsymbol{w}_4$	$\boldsymbol{w}_5$	$\boldsymbol{w}_6$	$\boldsymbol{w}_7$	$\boldsymbol{w}_8$	$\boldsymbol{w}_9$	$\boldsymbol{w}_{10}$	$\boldsymbol{w}_{11}$
$w_{1.}$	1/2	1/2	1/2	0	0	0	1/3	1/3	1/3	0	1/4
$w_{2.}$	1/2	0	0	1/2	1/2	0	1/3	1/3	0	1/3	1/4
$w_{3.}$	0	1/2	0	1/2	0	1/2	1/3	0	1/3	1/3	1/4
$w_{4.}$	0	0	1/2	0	1/2	1/2	0	1/3	1/3	1/3	1/4

Let $\tilde{q}_{p_i}$ denote the excess VaR of portfolio p_i associated with weight vector $\boldsymbol{w}_i$, i.e.,

$$\tilde{q}_{p_i} = \left[(\boldsymbol{q} \odot \boldsymbol{w}_i)' \otimes (\boldsymbol{q} \odot \boldsymbol{w}_i)' \right] \boldsymbol{D}\boldsymbol{\rho}, \tag{9}$$

and consider portfolios p_i, $i = 1, \ldots, m$, $m = n(n-1)/2$. Defining $\tilde{\boldsymbol{q}} = (\tilde{q}_{p_1}, \ldots, \tilde{q}_{p_m})'$ and the $m \times n^2$ matrix $\boldsymbol{Z} = \left[\boldsymbol{1}_m (\boldsymbol{q} \otimes \boldsymbol{q})' \right] \odot (\boldsymbol{w}_1 \otimes \boldsymbol{w}_1, \ldots, \boldsymbol{w}_m \otimes \boldsymbol{w}_m)'$, with $\boldsymbol{1}_m$ being an $m \times 1$ vector of ones, the $n(n-1)/2$ equations take the matrix form $\tilde{\boldsymbol{q}} = \boldsymbol{X}\boldsymbol{\rho}$ with $\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{D}$. For linearly independent weight vectors, $\boldsymbol{X}$ is a non-singular square matrix, so that the vector of VaR–implied correlation estimates is obtained by

$$\boldsymbol{\rho} = \boldsymbol{X}^{-1}\tilde{\boldsymbol{q}}. \tag{10}$$

Note that the exactly–identified joint estimator, based only on two–asset portfolios, is equivalent to the pairwise estimator (3). Expression (10) provides, however, a compact joint expression for all correlation coefficients in $\boldsymbol{R}$.

3.3 Overidentification

VaR estimates from portfolios consisting of more assets than just i and j also convey information about ρ_{ij} and may help to gain estimation efficiency. Overdetermined systems use more information than exactly–identified ones by taking more risk "measurements" based on additional, linearly independent weight vectors.

Considering, again, a universe of $n = 4$ assets and, for example, all equal–weight, three–asset portfolios ($k = 3$), we can construct the $m_3 = \binom{4}{3} = 4$ weight vectors $\boldsymbol{w}_7$ through $\boldsymbol{w}_{10}$ in listed Table 1. Finally, we can construct one ($m_4 = 1$) additional equal–weight vector, $\boldsymbol{w}_{11}$, from all four assets. Thus, in a four–asset universe, confining ourselves to equal–weight subset–portfolios, we can specify an overdetermined system of altogether $m_{2:4} = m_2 + m_3 + m_4 = 11$ equations to derive the six unknowns. In the general n–asset case, we can construct $m_{2:n} = \sum_{k=2}^{n} \binom{n}{k} = 2^n - n - 1$ different two– to n–asset portfolios with equal weights, to

solve for the $n(n-1)/2$ unknowns.[7]

In an overdetermined system with $m > n(n-1)/2$ equations, (9) will hold only approximately, so that $\tilde{q} = X\rho + u$, where vector u captures the approximation errors. Then, the least–squares estimator of ρ is given by

$$\hat{\rho} = (X'X)^{-1} X'\tilde{q}. \tag{11}$$

Instead of *equal*–weight portfolios, which maximize the "degree of orthogonality" (i.e., minimize $w_i'w_j$), the choice of weights may be motivated by practical consideration. Fund managers, for example, are typically restricted in their asset allocation.[8] Then, it is desirable to obtain good correlation estimates for weights from the permissable region. Clearly, a good fit in regions, where a fund manager is not allowed to operate, is of little use.

3.4   Constrained Estimation

Joint estimation via (10) or (11) does not guarantee that interval restriction $|\hat{\rho}_{ij}| \leq 1$ holds nor that the fitted correlation matrix is PSD. In this section, we discuss two strategies to overcome this: a direct approach and a two–step procedure.

As in the pairwise approach, the interval restriction can be achieved via truncation. To do so, view the joint estimation as a constrained quadratic programing problem, minimizing $u'u = \tilde{q}'\tilde{q} + \rho'X'X\rho - 2\tilde{q}'X\rho$, with inequality constraints $|\rho| \leq \mathbf{1}_{n(n-1)}$ being imposed. To also guarantee PSDness, a further constraint needs to be imposed. Because the correlation matrix is a symmetric, real matrix, PSDness requires all eigenvalues of $R(\rho)$, collected in $(n \times 1)$ vector λ, to be nonnegative. Then, to directly estimate tail–correlations matrices satisfying interval and PSD constraints, solve

$$\min_{\rho} \frac{1}{2}\rho'X'X\rho - \tilde{q}'X\rho, \quad \text{subject to: } |\rho| \leq \mathbf{1}_{n(n-1)/2} \text{ and } \lambda \geq 0. \tag{12}$$

If strict positive definiteness is required, we specify the last inequality in (12) as $\lambda \geq \varepsilon\mathbf{1}_{n(n-1)/2} > 0$, with ε chosen such that R is "reasonably" well–conditioned to guarantee, for example, stable inversion.

As n grows, direct constrained estimation via (12) becomes impractical, since the number of unknowns, $n(n-1)/2$, quickly becomes too large for iterative numerical optimization. A two–step strategy, based on the spectral decomposition of $R(\rho)$, i.e.,

$$R(\hat{\rho}) = U\Lambda U', \tag{13}$$

---

[7]For large n, $m_{2:n}$ becomes too large, so that one may set up m equations with $m < m_{2:n}$.

[8]E.g., fund managers may be limited to holding no more than a certain percentage of a specific asset type, or have to track a specific benchmark and, thus, to approximately follow its weights.

Table 2: Overview of the estimators investigated in the Monte Carlo study.

Label	Method	Estimator	Constraints	Weight vectors
Pair/J$_{2:2}$NC	pairwise	(3)/(10)	unconstrained	$\boldsymbol{w}_1$–$\boldsymbol{w}_6$
Pair/J$_{2:2}$2S	joint	(3)/(10)+(15)	$\lvert\rho_{ij}\rvert \leq 1$ & PSD	$\boldsymbol{w}_1$–$\boldsymbol{w}_6$
J$_{2:3}$NC	joint	(11)	unconstrained	$\boldsymbol{w}_1$–$\boldsymbol{w}_{10}$
J$_{2:3}$2S	joint	(11)+(15)	$\lvert\rho_{ij}\rvert \leq 1$ & PSD	$\boldsymbol{w}_1$–$\boldsymbol{w}_{10}$
J$_{2:4}$NC	joint	(11)	unconstrained	$\boldsymbol{w}_1$–$\boldsymbol{w}_{11}$
J$_{2:4}$2S	joint	(11)+(15)	$\lvert\rho_{ij}\rvert \leq 1$ & PSD	$\boldsymbol{w}_1$–$\boldsymbol{w}_{11}$

circumvents this drawback. In (13), the $n \times n$ diagonal matrix $\boldsymbol{\Lambda}$, with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, contains the eigenvalues and matrix $\boldsymbol{U}$ the eigenvectors of $\boldsymbol{R}(\boldsymbol{\rho})$. If $\boldsymbol{R}(\boldsymbol{\rho})$ is not PSD, one or more of the eigenvalues are negative. Driessel (2007) shows that by replacing $\boldsymbol{\Lambda}$ with $\tilde{\boldsymbol{\Lambda}}$, which matches $\boldsymbol{\Lambda}$ but has all negative eigenvalues set to zero,[9] we obtain a PSD approximation of the non–PSD matrix $\boldsymbol{R}(\hat{\boldsymbol{\rho}})$,[10] say

$$\tilde{\boldsymbol{R}} = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}\boldsymbol{U}', \tag{14}$$

that is best in terms of the Frobenius norm, $\|\cdot\|_F$, and spectral norm $\|\cdot\|_S$, i.e., $\|\boldsymbol{R} - \tilde{\boldsymbol{R}}\|_F^2 = \mathrm{trace}((\boldsymbol{R} - \tilde{\boldsymbol{R}})^2)$ and $\|\boldsymbol{R} - \tilde{\boldsymbol{R}}\|_S^2 = \lambda_{max}((\boldsymbol{R} - \tilde{\boldsymbol{R}})^2)$.

In general, approximation $\tilde{\boldsymbol{R}}$ will not be a proper correlation matrix, as the diagonal elements will not be exactly one, and needs to be rescaled. Then, the two–step joint estimator is given by

$$\boldsymbol{R}_{2S} = \tilde{\boldsymbol{S}}\tilde{\boldsymbol{R}}\tilde{\boldsymbol{S}}, \tag{15}$$

where the diagonal scaling matrix $\tilde{\boldsymbol{S}}$ contains the reciprocal square roots of the diagonal elements of $\tilde{\boldsymbol{R}}$.

Exactly–identified joint estimators, Pair/J$_{2:2}$, use only two–asset portfolios, i.e., $\boldsymbol{w}_1$ through $\boldsymbol{w}_6$ in Table 1. The overidentified versions, J$_{2:3}$ and J$_{2:4}$, make use of weight vectors $\boldsymbol{w}_1$ through $\boldsymbol{w}_{10}$ and $\boldsymbol{w}_1$ through $\boldsymbol{w}_{11}$, respectively. Also for the overidentified joint estimators, we investigate unconstrained (labeled "NC") and constrained two–step versions (labeled "2S").

---

[9]As with the direct estimator, setting the negative eigenvalues to zero will produce a *semi*definite tail–correlation matrix. The matrix will be strictly positive definite, if we set the negative eigenvalues to some (small) positive value.

[10]The approximation was also suggested in Rebonato and Jäckel (2000) without, however, discussing or proving its properties. Decomposition–based lower–rank approximations have a long and successful tradition in state space model reduction (see Pernebo and Silverman (1982) and Mittnik (1990)).

Table 3: Correlations used in the Monte Carlo simulation.

	DJIA	DAX	Brazil
DAX	.9		
Brazil	.6	.7	
Russia	.4	.5	.6

Note that we do not report results for the constrained *direct* estimator (12), because it did not produce better fits, measured in terms of mean squared error (MSE), than the two–step estimator. In fact, to reach the accuracy of the two–step estimator, a large number of iterations are typically required, so that the computational burden can be high without gaining precision.

We simulate 10,000 samples of size $1,000$, making iid draws from the joint normal distribution $N(\mathbf{0}, \mathbf{R})$. Hence, dependence is fully described by conventional Pearson correlations, which were estimated from monthly returns (January 2002– July 2010) of the following stock indices: Dow Jones Industrial Average (DJIA), German DAX, MSCI Brazil, and MSCI Russia. The (rounded) correlation estimates are shown in Table 3. The Monte Carlo results for the 90%, 95%, 99%, and 99.5% VaR–implied tail correlations are summarized in Table 4, reporting the estimators' bias and MSE. The columns "Int. Viol." and "PSD Viol." state the percentage of cases, where the estimated correlation matrix violates interval or the PSD condition, respectively.

The simulation results clearly demonstrate that the unconstrained pairwise estimator, $\text{Pair/J}_{2:2}\text{NC}$, is prone to interval violations. The violations tend to increase as one moves into the tail and range from 6.94% of the cases (for the VaR_{90}–implied estimates) to 16.27% ($\text{VaR}_{99.5}$). For the unconstrained overidentified estimators $\text{J}_{2:3}\text{NC}$ and $\text{J}_{2:4}\text{NC}$, interval–violation frequencies diminish as the degree of overidentification grows. For the $\text{J}_{2:4}\text{NC}$ estimator, relative improvements over the unconstrained pairwise estimator range from 15% to 35%, across all confidence levels considered.

Regarding PSD violations, we obtain a similar picture. Their frequency ranges from 13.72% to 30.95% for the pairwise estimator; and there are consistently fewer PSD violations for the overidentified estimators—with relative improvements ranging from 14% to 28% for $\text{J}_{2:4}\text{NC}$. The results in columns "Int. Viol." and "PSD Viol." in Table 4 document that the two–step estimator does, indeed, avoid PSD violations.

With respect to accuracy, we observe that all bias statistics are extremely small, but tend to increase as the VaR–confidence level rises. With 0.29 (after multipli-

Table 4: Monte Carlo evaluation of interval and PSD violations and of the goodness of fit of tail–correlation estimates (multiplied by 100).

Estimator	Int. Viol. (%)	PSD Viol. (%)	Bias	MSE
VaR_{90}				
Pair/$\text{J}_{2:2}$NC	6.94	14.08	0.06	61.93
Pair/$\text{J}_{2:2}$2S	0.00	0.00	-0.06	59.81
$\text{J}_{2:3}$NC	4.92	10.37	0.04	54.68
$\text{J}_{2:3}$2S	0.00	0.00	-0.03	53.44
$\text{J}_{2:4}$NC	4.75	10.21	0.04	54.52
$\text{J}_{2:4}$2S	0.00	0.00	-0.03	53.33
VaR_{95}				
Pair/$\text{J}_{2:2}$NC	6.77	13.72	0.10	65.28
Pair/$\text{J}_{2:2}$2S	0.00	0.00	-0.02	63.20
$\text{J}_{2:3}$NC	4.75	10.20	0.10	57.90
$\text{J}_{2:3}$2S	0.00	0.00	0.02	56.68
$\text{J}_{2:4}$NC	4.42	9.90	0.09	57.69
$\text{J}_{2:4}$2S	0.00	0.00	0.02	56.53
VaR_{99}				
Pair/$\text{J}_{2:2}$NC	12.38	23.96	0.13	119.71
Pair/$\text{J}_{2:2}$2S	0.00	0.00	-0.19	112.53
$\text{J}_{2:3}$NC	9.78	20.32	0.13	107.83
$\text{J}_{2:3}$2S	0.00	0.00	-0.10	103.03
$\text{J}_{2:4}$NC	9.66	20.28	0.15	107.59
$\text{J}_{2:4}$2S	0.00	0.00	-0.08	102.89
$\text{VaR}_{99.5}$				
Pair/$\text{J}_{2:2}$NC	16.27	30.95	0.29	165.98
Pair/$\text{J}_{2:2}$2S	0.00	0.00	-0.20	153.37
$\text{J}_{2:3}$NC	14.04	27.05	0.26	150.75
$\text{J}_{2:3}$2S	0.00	0.00	-0.10	141.98
$\text{J}_{2:4}$NC	13.86	26.77	0.26	150.41
$\text{J}_{2:4}$2S	0.00	0.00	-0.10	141.87

cation by 100), the unconstrained pairwise estimator has the largest bias reported. With one exception, the constrained two–step estimator is always less biased than the unconstrained one. Also for the MSEs a consistent picture arises: pairwise approaches always perform worse; i.e., overidentification consistently improves accuracy. The best results are obtained for $J_{2:4}2S$, the two–step estimator that uses all weight vectors listed in Table 1 and corrects for PSD violations. This suggests that imposing definiteness tends to improve accuracy by enforcing a form of regularization, which gives the estimates less room to stray away from "reasonable" values.

4   Empirical Illustration

To assess the applicability of the two–step estimator to larger sets of assets, we estimate tail–correlation matrices for the 30 stocks belonging to the German DAX index. Using daily returns over the period March 2003 to April 2013, we estimate left– and right–tail correlations for quantiles ranging from 1% to 25% and 75% to 99%.

With a total of 435 correlation coefficients, the degree of overidentification, as outlined in Section 3.3, can become excessively large. We obtain, for example, 435 two–, 4,060 three– and 27,405 four–asset portfolios. Overidentification using all possible two– through n–asset portfolios—as done in the Monto Carlo simulations reported above—would produce close to $2^{30} \approx 10^9$ equations. Below, we confine ourselves to specifying only equal–weight portfolios made up of all possible two–, three– and $(n-3)$–asset combinations. This amounts to a total of 8,555 (=435+4060+4060) linearly independent portfolios for determining the 435 tail–correlation coefficients.

The results for both tails are summarized in Figure 1, displaying the average of the 435 estimated tail correlations (marked by "+") associated with the respective quantiles. The horizontal line at 0.444 indicates the average of the 435 Pearson correlation estimates. The averages of the left–tail correlations start at the 25%–quantile with 0.400, i.e., well below the Pearson average, but increases as we move further into the loss tail, reaching 0.534 at the 1%–quantile. The right–tail correlations behave quite differently, starting with 0.463 at the 75%–quantile and falling monotonically to 0.349 at the 99%–quantile.

To check, we also estimate tail correlations from simulated iid draws from the multivariate normal distribution N($\mathbf{0}, \hat{\boldsymbol{R}}$), with $\hat{\boldsymbol{R}}$ being the Pearson correlation matrix estimated from the 30 stock–return series. As they should, the averages of the tail–correlation estimates (in Figure 1 marked by "o") are, indeed, about
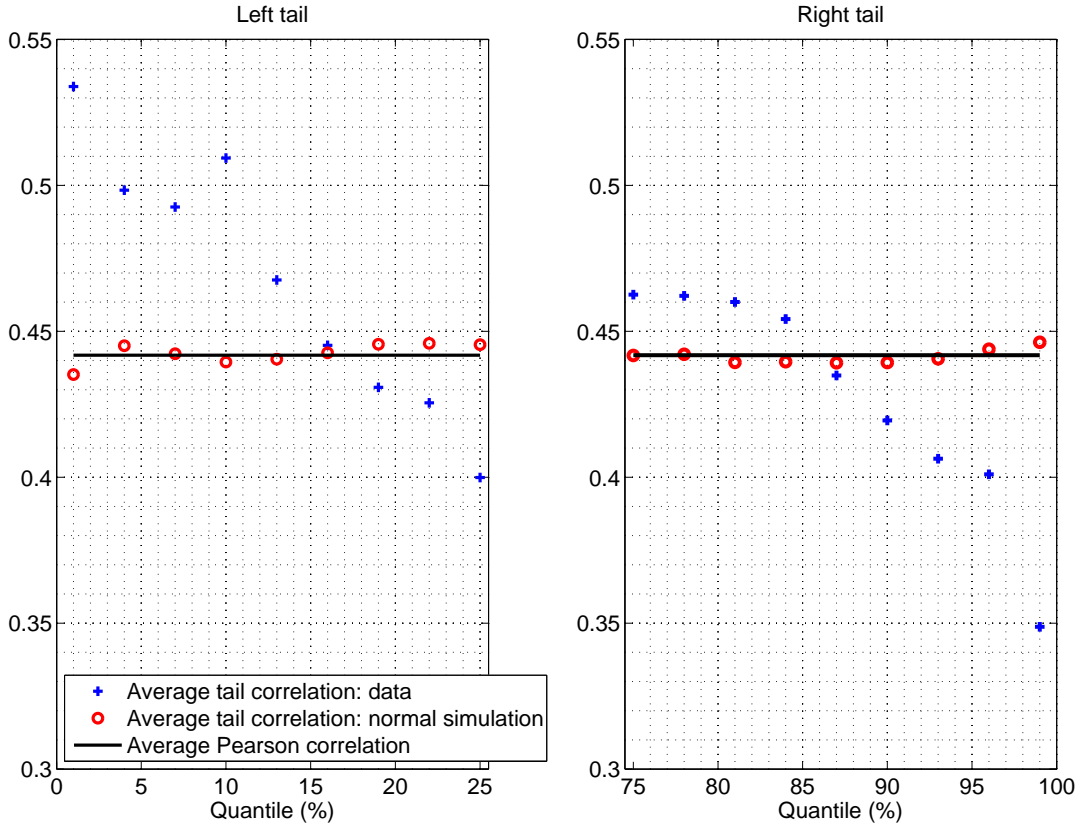
Figure 1: Average tail and Pearson correlation estimates for the 30 DAX stocks.

constant across both tails and very close to the Pearson value.[11] This exercise demonstrates that the correlational dependence of the DAX returns varies distinctly as we move into the tails and that it is not compatible with an elliptical data–generating process.

The behavior of the empirical tail–correlation estimates is in line with the literature cited in Footnote 1: during severe market downturns, DAX stocks tend to be more in sync than in sideways or upward markets. This finding does have direct implications for portfolio construction. Assume, for example, a portfolio manager pursues a so–called risk–parity strategy, where the portfolio weights are such that each asset contributes the same amount of volatility to the portfolio. Then, the weights satisfy $w_i \sigma_i \sim 1/n$, $i = 1, \ldots, n$. In case of equality, the portfolio variance is simply given by $\sigma_p^2 = \boldsymbol{w}'\boldsymbol{\Sigma}\boldsymbol{w} = \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij}$, where $\boldsymbol{\Sigma}$ denotes the covariance matrix. In other words, the portfolio variance is directly related to the average correlation, $\bar{\rho}$ reported in Figure 1, since $\bar{\rho} = \frac{2}{n(n-1)} \sum_{i=2}^{n} \sum_{j=1}^{i-1} \rho_{ij}$. The

---

[11]The plotted estimates are the means from 20 replications with the sample size matching that of the underlying stock data.
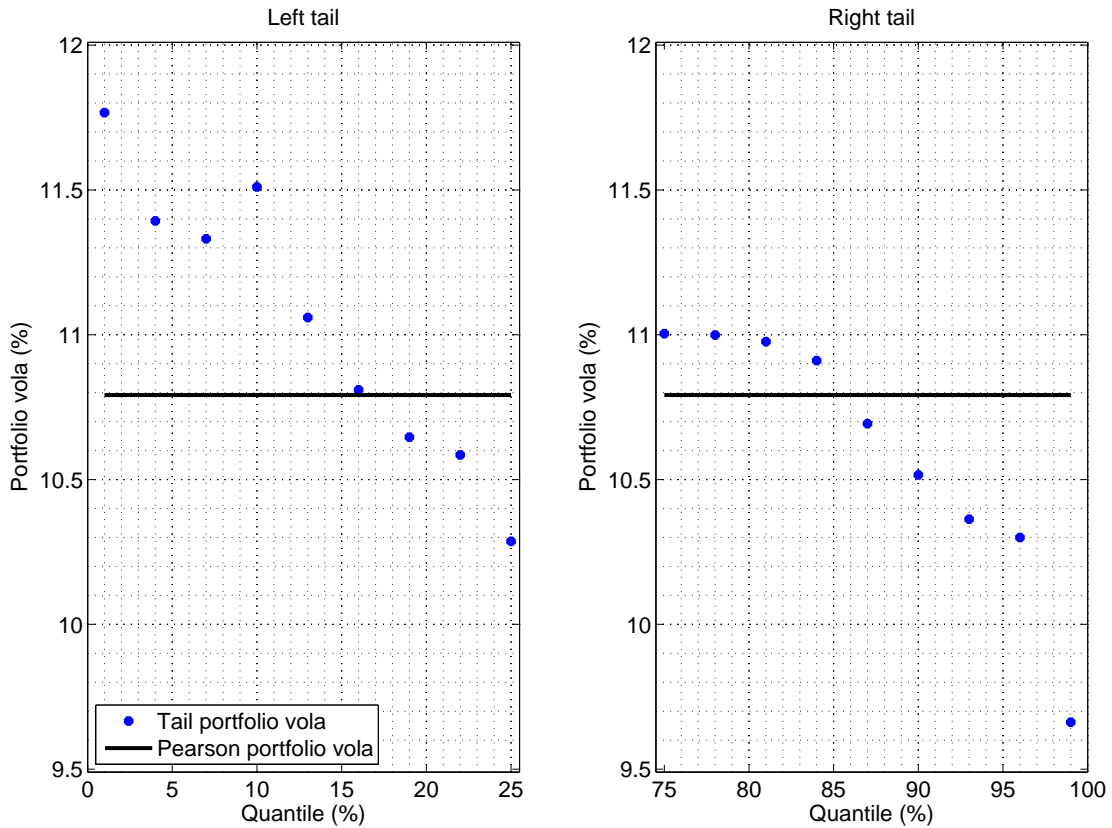
Figure 2: Annualized portfolio volatilities based on tail– and Pearson–correlation estimates for the 30 DAX stocks under a risk–parity strategy.

annualized (i.e., multiplied by $\sqrt{252}$) portfolio volatilities based on the above tail–and Pearson–correlation estimates are shown in Figure 2. The Pearson estimate for the portfolio volatility is 10.79%, whereas, for example, the tail–correlation–based estimate at the 1%–quantile amounts to 11.76%.

Furthermore, assume that, with a confidence level of 99%, the portfolio manager wants to limit the annualized portfolio volatility to 10% by holding an appropriate risk–free cash position. This can be accomplished by setting the weight of the cash component, w_{cash}, such that $(1 - w_{cash})\sigma_p = 10$ or $w_{cash} = 1 - 10/\sigma_p$. Then, regardless of the confidence level chosen, the "Pearson manager's" cash position would be $1-10/10.79$ or 7.34%, whereas the "tail–correlation manager" would hold more than twice as much cash, namely, $1-10/11.76$ or 15.02%. This demonstrates that tail–correlation analysis can be a valuable tool for portfolio management when trying to control downside risk.

5 Concluding Remarks

We have proposed a method for jointly estimating the elements of VaR–implied tail–correlation matrices which simply requires the solution of a system of linear equations. Monte Carlo simulations show that overidentified versions of the estimator improve efficiency. Two variants, guaranteeing positive semidefiniteness of the estimated matrix, have been presented: a direct and a two–step approach. Both are similarly accurate, but the latter is computationally more appealing, as it does not involve complex iterative numerical optimization. An application to 30 German DAX stocks has demonstrated that the two–step estimator is straightforwardly applicable to "larger–than–textbook" asset universes. The resulting tail–correlation estimates strongly suggest that the DAX stocks' dependence structure varies systematically and distinctly across left and right tails. Knowledge about such properties is useful when pursuing, for example, downside–risk–based portfolio optimization.

The conventional Pearson–correlation concept assumes that the joint distribution is elliptical. Given that any distributional assumption represents a—more or less accurate—approximation of the true data–generating process, we do not expect ellipticity to hold exactly in practice. In this case, VaR–implied correlation estimates can be viewed as local elliptical approximations, with the location being determined by both the VaR–confidence level and the portfolio weights specified. If a portfolio manager needs to operate in a particular subspace of the investment–opportunity set, the proposed estimation strategy enables the manager to obtain a best local correlational approximation in that portfolio–weight region which matters most. Similarly, in situations where assets do not adhere to idealizing distributional assumptions and a portfolio manager pursues VaR–based strategies for downside–risk protection, he or she can obtain correlation estimates that are relevant for the particular VaR confidence level implied by the strategy.

Note that the computational cost for the two–step estimator is rather modest. In the 30–asset DAX case, the estimation of a tail–correlation matrix took about 0.63 seconds (using Matlab on a laptop with an Intel i7Q740 CPU). Obtaining the set of empirical quantiles used in Figure 1, involving altogether 8,585 (individual and Portfolio) return series with 2,099 observation each, took about another 2 seconds. Thus, computational burden is no argument against estimating VaR–implied tail correlation matrices.[12]

Throughout the analysis, we have assumed that the assets' VaRs are constant

---

[12]Still, with about 0.073 seconds, the computation of a 30×30 Pearson correlation matrix from 2,099 observations is almost ten times faster.

over time. Dynamic extensions are currently under investigation. One strategy is to adopt the Conditional Autoregressive Value at Risk (CAViaR) framework suggested by Engle and Manganelli (2004), which is based on quantile regressions and, as, for example, shown in Kuester et al. (2006), well capable of capturing GARCH–type conditional heteroskedasticity in asset returns.

References

Ang, A., Bekaert, G., 2002. International Asset Allocation with Regime Shifts. Review of Financial Studies 15, 1137–1187.

Ang, A., Chen, J., 2002. Asymmetric Correlations of Equity Portfolios. Journal of Finacial Economics 63, 443–494.

Artzner, P., Delbaen, F., Eber, J. M., Heath, D., 1999. Coherent Measures of Risk. Mathematical Finance 9, 203–228.

Bae, K.-H., Karolyi, A., Stulz, R., 2003. A new approach to measuring financial contagion. Review of Financial Studies 16, 717–763.

Butler, K. C., Joaquin, D. C., 2002. Are the gains from international portfolio diversification exaggerated? The influence of downside risk in bear markets. Journal of International Money and Finance 21, 981–1011.

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

Campbell, R., Forbes, C. S., Koedijk, K. G., Kofman, P., 2008. Increasing correlations or just fat tails? Journal of Empirical Finance 15, 287–309.

Campbell, R., Koedijk, K., Kofman, P., 2002. Increased Correlation in Bear Markets. Financial Analysts Journal 58, 87–94.

Cotter, J., Longin, F., 2007. Implied Correlations from VaR. Working paper, University College Dublin.

Das, S. R., Uppal, R., 2004. The Effect of Systemic Risk on International Portfolio Choice. Journal of Finance 59, 2809–2834.

Driessel, K. R., 2007. Computing the Best Positive Semi-definite Approximation of a Symmetric Matrix Using a Flow. http://www.ima.umn.edu/AlgGeom/W3.5-9.07/activities/Driessel-Kenneth/poster.pdf, Institute for Mathematics and its Applications, University of Minnesota, March 2007.

Engle, R. F., Manganelli, S., 2004. CAViAR: Conditional Autoregressive Value at Risk by Regression Quantiles. Journal of Business and Economics Statistics 22, 367–381.

Erb, C. B., Harvey, C. R., Viskanta, T. E., 1994. Forecasting International Equity Correlations. Financial Analysts Journal 50, 32–45.

European Commission, July 2007. Proposal for a Directive of the European Parliament and of the Council on the taking–up and pursuit of the business of Insurance and Reinsurance – Solvency II. 2007/0143(COD).

Haas, M., Mittnik, S., 2009. Portfolio Selection with Common Correlation Mixture Models. In: Bol, G., Rachev, S. T., Würth, R. (Eds.), Risk Assessment: Decisions in Banking and Finance. Physika Verlag.

Haas, M., Mittnik, S., Paolella, M. S., 2004. Mixed Normal Conditional Heteroskedasticity. Journal of Financial Econometrics 2, 211–250.

Hong, Y., Tu, J., Zhou, G., 2007. Asymmetries in in Stock Returns: Statistical Tests and Economic Evaluation. Review of Financial Studies 20, 1547–1581.

Karolyi, G. A., Stulz, R. M., 1996. Why Do Markets Move Together? An Investigation of U.S.–Japan Stock Return Comovement. Journal of Finance 51, 951–986.

Kuester, K., Mittnik, S., Paolella, M., 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. Journal of Financial Econometrics 4, 53–89.

Longin, F. M., Solnik, B. H., 1995. Is the Correlation in International Equity Returns Constant: 1960–1990? Journal of International Money and Finance 13, 3–26.

Longin, F. M., Solnik, B. H., 2001. Extreme Correlation of International Equity Markets. Journal of Finance 56, 651–678.

Mittnik, S., 1990. Macroeconomic Forecasting Experience with Balanced State Space Models. International Journal of Forecasting 6, 337–348.

Okimoto, T., 2008. New Evidence of Asymmetric Dependence Structures in International Equity Markets. Journal of Financial and Quantitative Analysis 43, 787–815.

Pernebo, L., Silverman, L. M., 1982. Model Reduction via Balanced State Space Representation. IEEE Transactions on Automatic Control AC-27, 382.

Rebonato, R., Jäckel, P., 2000. The Most General Methodology to Create a Valid Correlation Matrix for Risk Management and Option Pricing Purposes. Journal of Risk 2, 17–26.

Silvapulle, P., Granger, C. W. J., 2001. Large returns, conditional correlation and portfolio diversification: a value-at-risk approach. Quantitative Finance 1, 542–551.