H I E R

Harvard Institute of Economic Research

Discussion Paper Number 2139

Value at Risk Under Dependence and Heavy Tailedness: Models with Comon Shocks*

By Rustam Ibragimov
And
Johan Walden

October 2007

HARVARD UNIVERSITY
Cambridge, Massachusetts

VALUE AT RISK UNDER DEPENDENCE AND HEAVY-TAILEDNESS: MODELS WITH COMMON SHOCKS*

Abbreviated title: Value at risk

Rustam Ibragimov

Department of Economics, Harvard University

Johan Walden

Haas School of Business, University of California at Berkeley

Address for manuscript correspondence:

Rustam Ibragimov; Department of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138; *Email*: ribragim@fas.harvard.edu; Phone: +1-617-496-4795; Fax: +1-617-495-7730

Johan Walden; Haas School of Business, University of California at Berkeley, 545 Student Services Building, #1900, Berkeley, CA 94720-1900; *Email*: walden@haas.berkeley.edu; Phone: +1-510-643-0547; Fax: +1-510-643-1420

*We are grateful to Daniel Ahn, Shashibhushan Borade, John Campbell and Tian Tian Qiu for helpful discussions. The support of the NUS Risk Management Institute is gratefully acknowledged.

2000 Mathematics subject classifications. 91B28, 91B30, 91B70

Keywords and phrases. Portfolio analysis, diversification, value at risk, heavy-tailedness, dependence, majorization, common shocks, factor models, riskiness, power laws, random effects, linear estimators, efficiency

ABSTRACT

This paper presents an analysis of diversification and portfolio value at risk for heavy-tailed dependent risks in models with multiple common shocks. We show that, in the framework of value at risk comparisons, diversification is optimal for moderately heavy-tailed dependent risks with common shocks and finite first moments, provided that the model is balanced, i.e., that all the risks are available for portfolio formation. However, diversification is inferior in balanced extremely heavy-tailed risk models with common factors. Finally, in several unbalanced dependent models, diversification is optimal, even though there is extreme heavy-tailedness in common shocks or in idiosyncratic parts of the risks.

1 Introduction

Numerous studies in economics and finance have argued that many series encountered in these fields are heavy-tailed and can be modeled using risks X with distributions exhibiting power law decline [1]

$$P\Big(|X| > x\Big) \asymp x^{-\alpha} \tag{1}$$

(see, among others, the discussion in Embrechts, Klupperberg & Mikosch 1997, Rachev, Menn & Fabozzi 2005). We mention a sample of estimates of the tail index α for returns on various stocks and stock indices: $3 < \alpha < 5$ (Jansen & de Vries 1991), $2 < \alpha < 4$ (Loretan & Phillips 1994), $1.5 < \alpha < 2$ (McCulloch 1997), $0.9 < \alpha < 2$ (Rachev & Mittnik 2000), $\alpha \approx 3$ (Gabaix, Gopikrishnan, Plerou & Stanley 2003). As discussed by Nešlehova, Embrechts & Chavez-Demoulin (2006), tail indices less than one are observed for empirical loss distributions of a number of operational risks. Scherer, Harhoff & Kukies (2000) and Silverberg & Verspagen (2007) report the tail indices α to be considerably less than one for financial returns from technological innovations. Rachev et al. (2005) discuss and review the vast literature that supports heavy-tailedness and Pareto distributions for equity and bond returns.

Several recent studies in these fields have focused on the analysis of diversification and value at risk theory under heavy-tailedness. Bouchard & Potters (2004), Ch. 12, present a detailed discussion of portfolio choice under various distributional and dependence assumptions and diversification measures, including the asymptotic results in the value at risk framework for heavy-tailed power law distributions. As was shown in Ibragimov (2004, 2005) in a general context based on majorization theory and arbitrary portfolio weights comparisons, diversification may be inferior for heavy-tailed risks whose distributions satisfy power law (1) with $\alpha < 1$. As shown in Ibragimov (2004, 2005), diversification is typically preferable for convolutions of stable heavy-tailed risks that follow (1) with $\alpha > 1$ (see the review in Appendix A1).

Recently, Ibragimov & Walden (2007*a*) showed that, with a value at risk approach with bounded risks concentrated on a sufficiently large interval, diversification may be suboptimal up to a certain number of risks and then become optimal. Ibragimov, Jaffee & Walden (2006) demonstrate how this analysis can be used to explain low levels of reinsurance among insurance

[1] Here and throughout the paper, $f(x) \asymp g(x)$ means that $0 < c \leq f(x)/g(x) \leq C < \infty$ for large x, for constants c and C.

providers in markets for catastrophe reinsurance. Ibragimov & Walden (2007*b*) study portfolio diversification for nonlinear transformations of heavy-tailed risks and for distributions that exhibit local or moderate deviations from power tails (1) in the form of additional slowly varying or exponential factors. Among other results, Ahn (2007) applies the majorization approach in the study of expected utility, optimal investment and diversification over time in futures markets under heavy-tailedness. Several examples that illustrate the phenomenon that diversification is not always preferable are presented in Kaas, Goovaerts & Tang (2004).

While the above works provide several extensions of the value at risk theory for the case of dependence, including the case of multiplicative common shocks, many results in them cover only the settings with uncorrelated risks. Our objective with this paper is to extend the analysis beyond this class of distributions. We obtain general results on portfolio value at risk comparisons for correlated heavy-tailed risks exhibiting additive common shocks structures of the type

$$Y_{ij} = R_i + C_j + U_{ij}, \quad i = 1, ..., r, \; j = 1, ..., c. \tag{2}$$

In (2), the "row effects" common shocks R_i, the "column effects" common shocks C_j and the "error" variables U_{ij} are assumed to be independent of each other and to be independent and identically distributed among themselves.[2] We also present the value at risk analysis for a particular case of (2) given by

$$Y_{ij} = R_i + U_{ij}, \quad i = 1, ..., r, \; j = 1, ..., c. \tag{3}$$

Together with their multiplicative analogues

$$Y_{ij} = R_i U_{ij}, \quad R_i > 0, \tag{4}$$

(which were discussed in Ibragimov 2004, 2005, Ibragimov & Walden 2007*a*,*b*), models (2) and (3) provide a natural framework for modeling risks subject to (additive) common shocks R_i and C_j (such as political or macroeconomic ones, see the discussion in Andrews 2005). The common shocks R_i affect all risks Y_{ij}, $j = 1, ..., c$, in the ith row (say, in the ith country) and the common shocks C_j affect all risks Y_{ij}, $i = 1, ..., r$, in the jth column (say, in the jth industry).[3]

[2]The results in the paper also cover the case of dependence among the risks R_i, C_j and U_{ij} (see Section 7).

[3]The dependence properties of additive common shock models (2) and (3) are more complicated than those in relation (4) and its two-shock analogues. For instance, as already mentioned, while multiplicative common shock models (4) imply uncorrelatedness of the risks Y_{ij}, this, evidently, does not hold for (3).

We notice that, with the index j denoting time t, (3) also includes factor models

$$Y_{it} = C_t + U_{it}, \quad i = 1, ..., r, \ t = 1, ..., T,$$

with the factor C_t. Similarly, multifactor models constitute a subclass of models (2) and (3) and their extensions to the case of multiple common shocks.

The analysis of aggregation properties and portfolio choice for risks of type (3) is needed in many problems in risk management, including models for operational risks that are typically heavy-tailed, as indicated above. For instance, as discussed in Chavez-Demoulin, Nešlehova & Embrechts (2006), the analysis of operational risk measures leads to the study of the total loss $Z = \sum_{i=1}^{r} Z_i$ for a number r (r risk types, business lines or classes) of loss random variables (r.v.'s) Z_i. In turn, the loss variables are often of the form $Z_i = \sum_{j=1}^{n_i} U_j$, $i = 1, ..., r$, where the n_i are frequency variables assumed to be independent of the severity variables U_j. In many real-world frameworks in risk management, the variables U_j in such aggregation problems exhibit dependence. The specifications like (3) with risk type or risk line-specific common shocks R_i and their extensions provide a natural approach for modeling these dependence structures.

The paper is organized as follows. Section 2 introduces the notation and definitions of classes of moderately heavy-tailed and extremely heavy-tailed distributions dealt with throughout the paper.

Section 3 contains the main results of the paper on value at risk analysis and optimal portfolio choice in general common shocks models (2). We show that portfolio value at risk comparisons with the most diversified and the least diversified portfolio weights for dependent risks Y_{ij} in (2) under heavy-tailedness are similar to those in the case of independence (Theorem 1). In particular, under moderate heavy-tailedness with $\alpha > 1$ in the shocks R_i, C_j and U_{ij} in (2), the most diversified portfolio of risks Y_{ij} with equal weights is optimal with respect to portfolio value at risk comparisons. In addition, in such settings, the least diversified portfolio of Y_{ij}'s consisting of only one risk maximizes the value at risk over all portfolio weights. These conclusions are reversed under extreme heavy-tailedness with $\alpha < 1$ in the common shocks R_i and C_j and the idiosyncratic shocks U_{ij} in (2). Under these assumptions, the most diversified portfolio with equal weights has the maximal value at risk among all portfolios of Y_{ij}. The optimal portfolio under extreme heavy-tailedness in the variables in (2) is given by the least diversified portfolio that consists of only one risk. Theorem 2 shows that these results continue to hold separately for the common shock and idiosyncratic components of the returns on the portfolios of risks Y_{ij}. Theorem 2 thus allows one to compare value at risk for the portfolios of common shocks R_i or C_j, or for the portfolios of U_{ij} under heavy-tailedness in these variables.

Section 4 provides extensions of the results in Theorems 1 and 2 to value at risk comparisons between portfolios that are different from the most diversified and the least diversified portfolios (Theorem 3). The results provide value at risk comparisons for the portfolios considered in the literature on efficiency of linear location estimators in models (2). Similar portfolio value at risk comparisons under extreme heavy-tailedness in risks R_i, C_j and U_{ij} are opposite to those in the case of moderate heavy-tailedness. In addition, some of the comparisons have a natural interpretation in terms of the optimal portfolio choice for indices of the risks Y_{ij} in (2).

In Section 5 we show that the majorization approach to value at risk analysis developed in Sections 3 and 4 can be applied in the case of unbalanced models (2) or (3) that have unequal number of rows for each column or unequal number of columns for each row. Building on the interpretation in Section 4, the results in Section 5 are presented in the framework of value at risk analysis for equally weighted indices of heavy-tailed risks in (3) in Theorems 4-6. Theorems 4 and 5 imply optimality of diversification patterns in unbalanced dependent models, even though there is extreme heavy-tailedness in common shocks or in idiosyncratic parts of the risks. These conclusions are in contrast to variance comparisons for portfolio returns implied by the results in the literature (see the discussion in Section 5).

Section 6 discusses econometric and statistical applications of the results obtained in the paper. These applications are the analogues of the value at risk results in the framework of efficiency comparisons of linear estimators of location in random effects models.

To illustrate the main ideas and results, the portfolio value at risk analysis in Sections 3–5 is presented in the framework of risks that are subject to two or less additive common shocks with independence among the variables R_i, C_j and U_{ij}. In Section 7, we discuss how most of the results can be generalized to the case with more than two common factors. In addition, we show how the results in paper can also be extended to the case of dependence within the common shocks and the idiosyncratic risks. These dependence structures include convolutions of $\alpha-$symmetric distributions and important models with multiple multiplicative common shocks. Section 7 further discusses the analogues of the results in the paper for non-identically distributed dependent risks.

Section 8 makes some concluding remarks. Appendix A1 reviews the results on portfolio value at risk under heavy-tailedness and independence in Ibragimov (2004, 2005) that are needed in some of the arguments. Finally, Appendix A2 contains proofs of the results obtained in the paper.

2 Notation and classes of distributions

A r.v. X with density $f : \mathbf{R} \to \mathbf{R}$ and the convex distribution support $\Omega = \{x \in \mathbf{R} : f(x) > 0\}$ is log-concavely distributed if $log\ f(x)$ is concave in $x \in \Omega$, that is, if for all $x_1, x_2 \in \Omega$, and any $\lambda \in [0, 1]$, $f(\lambda x_1 + (1-\lambda)x_2) \geq (f(x_1))^{\lambda}(f(x_2))^{1-\lambda}$ (see An 1998, Bagnoli & Bergstrom 2005). A distribution is said to be log-concave if its density f satisfies the above inequalities. Examples of log-concave distributions include the normal distribution, the uniform density, the exponential density, the Gamma distribution $\Gamma(\alpha, \beta)$ with the shape parameter $\alpha \geq 1$, the Beta distribution $\mathcal{B}(a, b)$ with $a \geq 1$ and $b \geq 1$; and the Weibull distribution $\mathcal{W}(\gamma, \alpha)$ with the shape parameter $\alpha \geq 1$. Log-concave distributions have many appealing properties that have been utilized in a number of works in economics and finance (see the surveys in Karlin 1968, Marshall & Olkin 1979, An 1998, Bagnoli & Bergstrom 2005). However, such distributions cannot be used in the study of heavy-tailedness phenomena since any log-concave density is extremely light-tailed: in particular, if a r.v. X is log-concavely distributed, then its density has at most an exponential tail, that is, $f(x) = o(exp(-\lambda x))$ for some $\lambda > 0$, as $x \to \infty$ and, therefore, all the power moments $E|X|^{\gamma}$, $\gamma > 0$, of the r.v. exist (see Corollary 1 in An 1998). Throughout the paper, $\mathcal{LC}$ denotes the class of symmetric log-concave distributions ($\mathcal{LC}$ stands for "log-concave").

For $0 < \alpha \leq 2$, $\sigma > 0$, $\beta \in [-1, 1]$ and $\mu \in \mathbf{R}$, we denote by $S_{\alpha}(\sigma, \beta, \mu)$ the stable distribution with the characteristic exponent (index of stability) α, the scale parameter σ, the symmetry index (skewness parameter) β and the location parameter μ. That is, $S_{\alpha}(\sigma, \beta, \mu)$ is the distribution of a r.v. X with the characteristic function (c.f.)

$$E(e^{ixX}) = \begin{cases} exp\left\{i\mu x - \sigma^{\alpha}|x|^{\alpha}(1 - i\beta sign(x) tan(\pi\alpha/2))\right\}, & \alpha \neq 1, \\ exp\left\{i\mu x - \sigma|x|(1 + (2/\pi)i\beta sign(x) ln|x|\right\}, & \alpha = 1, \end{cases} \quad (5)$$

$x \in \mathbf{R}$, where $i^2 = -1$ and $sign(x)$ is the sign of x defined by $sign(x) = 1$ if $x > 0$, $sign(0) = 0$

and $sign(x) = -1$ otherwise (expression (5) is one of several possible parameterizations of c.f.'s of stable distributions). In what follows, we write $X \sim S_\alpha(\sigma, \beta, \mu)$, if the r.v. X has the stable distribution $S_\alpha(\sigma, \beta, \mu)$.

The index of stability α characterizes the heaviness (the rate of decay) of the tails of stable distributions $S_\alpha(\sigma, \beta, \mu)$. In particular, if $X \sim S_\alpha(\sigma, \beta, \mu)$, then its distribution satisfies power law (1). This implies that the $p-$th absolute moments $E|X|^p$ of a r.v. $X \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2)$ are finite if $p < \alpha$ and are infinite otherwise.

Distributions $S_\alpha(\sigma, \beta, \mu)$ with $\mu = 0$ for $\alpha \neq 1$ and $\beta = 0$ for $\alpha = 1$ are called strictly stable. If $X_i \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2]$, are i.i.d. strictly stable r.v.'s, then, for all $a_i \geq 0$, $i = 1, ..., n$,

$$\sum_{i=1}^{n} a_i X_i / \Big(\sum_{i=1}^{n} a_i^\alpha\Big)^{1/\alpha} \sim S_\alpha(\sigma, \beta, \mu) \tag{6}$$

(see Zolotarev 1986, Embrechts et al. 1997, Rachev & Mittnik 2000, for a detailed review of properties of stable distributions).

We denote by $\overline{\mathcal{CS}}$ the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in [1, 2]$ and $\sigma > 0$ (here and below, $\mathcal{CS}$ stands for "convolutions of stable"; the overline indicates that convolutions of stable distributions with indices of stability not less than 1 are taken). That is, $\overline{\mathcal{CS}}$ consists of distributions of r.v.'s X for which, with some $k \geq 1$, $X = Y_1 + ... + Y_k$, where Y_i, $i = 1, ..., k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in [1, 2]$, $\sigma_i > 0$, $i = 1, ..., k$.

Further, $\underline{\mathcal{CS}}$ stands for the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with indices of stability $\alpha \in (0, 1]$ and $\sigma > 0$ (the underline indicates considering stable distributions with indices of stability not greater than 1). That is, $\underline{\mathcal{CS}}$ consists of distributions of r.v.'s X for which, with some $k \geq 1$, $X = Y_1 + ... + Y_k$, where Y_i, $i = 1, ..., k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, 1]$, $\sigma_i > 0$, $i = 1, ..., k$.

Finally, we denote by $\overline{\mathcal{CSLC}}$ the class of convolutions of distributions from the classes $\mathcal{LC}$ and $\overline{\mathcal{CS}}$. That is, $\overline{\mathcal{CSLC}}$ is the class of convolutions of symmetric distributions which are either log-concave or stable with characteristic exponents not less than one ($\mathcal{CSLC}$ is the abbreviation of "convolutions of stable and log-concave"). In other words, $\overline{\mathcal{CSLC}}$ consists of distributions of r.v.'s X such that $X = Y_1 + Y_2$, where Y_1 and Y_2 are independent r.v.'s with distributions belonging to $\mathcal{LC}$ or $\overline{\mathcal{CS}}$.

All the classes $\mathcal{LC}$, $\overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}}$ and $\underline{\mathcal{CS}}$ are closed under convolutions. In particular, the class $\overline{\mathcal{CSLC}}$ coincides with the class of distributions of r.v.'s X such that, for some $k \geq 1$, $X = Y_1 + ... + Y_k$, where Y_i, $i = 1, ..., k$, are independent r.v.'s with distributions belonging to $\mathcal{LC}$ or $\overline{\mathcal{CS}}$.

In what follows, we write $X \sim \mathcal{LC}$ (resp., $X \sim \overline{\mathcal{CSLC}}$, $X \sim \overline{\mathcal{CS}}$ or $X \sim \underline{\mathcal{CS}}$) if the distribution of the r.v. X belongs to the class $\mathcal{LC}$ (resp., $\overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}}$ or $\underline{\mathcal{CS}}$).

The distributions of r.v.'s X from the class $\underline{\mathcal{CS}}$ are extremely heavy-tailed in the sense that their first moments are infinite: $E|X| = \infty$. In contrast, the distributions of r.v.'s X in $\overline{\mathcal{CSLC}}$ are moderately heavy-tailed in the sense that they have finite moments of order $0 < p < 1$: $E|X|^p < \infty, 0 < p < 1$. For a more extensive discussion on the classes of distributions discussed in this section and their generalizations, see Ibragimov (2004, 2005), Ibragimov & Walden (2007*a*).

3 Portfolio value at risk for models with multiple additive common shocks

Given a loss probability $0 < q < 1/2$ and a r.v. (risk) X, we denote by $VaR_q[X]$ the value at risk (VaR) of X at level q, that is, the $(1-q)-$quantile of X: $VaR_q[X] = \inf\{x \in \mathbf{R} : P(X > x) \leq q\}$ (throughout the paper, we interpret the positive values of risks X as a risk holder's losses). For a risk X with finite second moment, $var[X]$ will stand for its variance $var[X] = E(X - EX)^2$.

In what follows, $\mathbf{R}_+$ stands for $\mathbf{R}_+ = [0, \infty)$. For $N \geq 1$, denote $\mathcal{I}_N = \{w = (w_1, ..., w_N) \in \mathbf{R}_+^N : \sum_{i=1}^N w_i = 1\}$. In addition, given the $w = (w_1, ..., w_N) \in \mathbf{R}_+^N$ and N risks $X_1, ..., X_N$, we denote by $X(w) = \sum_{i=1}^N w_i X_i$ the return on the portfolio of X_i's with weights w.

Let $\underline{w}_N = \underbrace{\Big(1/N, 1/N, ..., 1/N\Big)}_{N} \in \mathcal{I}_N$ stand for the vector of equal portfolio weights and let $\overline{w}_N = \underbrace{\Big(1, 0, ..., 0\Big)}_{N} \in \mathcal{I}_N$ stand for the weights in the portfolio consisting of only one risk.

A vector $a \in \mathbf{R}^N$ is said to be majorized by a vector $b \in \mathbf{R}^N$, written $a \prec b$, if $\sum_{i=1}^k a_{[i]} \leq \sum_{i=1}^k b_{[i]}$, $k = 1, ..., N-1$, and $\sum_{i=1}^N a_{[i]} = \sum_{i=1}^N b_{[i]}$, where $a_{[1]} \geq \ldots \geq a_{[N]}$ and $b_{[1]} \geq \ldots \geq b_{[N]}$ denote the components of a and b in decreasing order. The relation $a \prec b$ implies that the components of the vector a are less diverse than those of b (see Marshall & Olkin 1979). In this context, it is easy to see that the following relations hold:

$$\underline{w}_N \prec w \prec \overline{w}_N \tag{7}$$

for all $w \in \mathcal{I}_N$.

A function $\phi : A \to \mathbf{R}$ defined on $A \subseteq \mathbf{R}^N$ is called *Schur-convex* (resp., *Schur-concave*) on A if $(a \prec b) \implies (\phi(a) \leq \phi(b))$ (resp. $(a \prec b) \implies (\phi(a) \geq \phi(b))$ for all $a, b \in A$. Evidently, if ϕ is Schur-convex or Schur-concave on a symmetric set A (so that $(a_1, ..., a_N) \in A$ implies $(a_{\pi(1)}, ..., a_{\pi(N)}) \in A$ for all permutations π of the set $\{1, ..., N\}$), then ϕ is symmetric on A, that is, $\phi(a_{\pi(1)}, ..., a_{\pi(N)}) = \phi(a_1, ..., a_N)$ for all permutations π.

Suppose that $v = (v_1, ..., v_N) \in \mathbf{R}_+^N$ and $w = (w_1, ..., w_N) \in \mathbf{R}_+^N$, $\sum_{i=1}^N v_i = \sum_{i=1}^N w_i$, are the weights of two portfolios of N risks (indices or assets' returns). If $v \prec w$, it is natural to think about the portfolio with weights v as being more diversified than that with weights w (see the discussion in Ibragimov 2004, 2005). Thus, for example, the portfolio with equal weights $\underline{w}_N$ in (7) is the most diversified among all the portfolios with weights $w \in \mathcal{I}_N$. In contrast, the portfolios with weights given by the components of $\overline{w}_N$ or their permutations consist of one risk and are the least diversified among the portfolios with weights $w \in \mathcal{I}_N$. In this regard, the notion of one portfolio being more or less diversified than another one is, in some sense, the opposite of that for vectors of weights for the portfolio.

For $w = (w_{11}, ..., w_{1c}, w_{21}, ..., w_{2c}, ..., w_{r1}, ..., w_{rc}) \in \mathcal{I}_{rc}$, denote $w_{0j} = \sum_{i=1}^r w_{ij}$, $j = 1, ..., c$, $w_{i0} = \sum_{j=1}^c w_{ij}$, $i = 1, ..., r$. Further, denote

$$w_0^{(row)} = (w_{10}, ..., w_{r0}) \in \mathcal{I}_r, \tag{8}$$

$$w_0^{(col)} = (w_{01}, ..., w_{0c}) \in \mathcal{I}_c. \tag{9}$$

For $w = (w_{11}, ..., w_{1c}, w_{21}, ..., w_{2c}, ..., w_{r1}, ..., w_{rc}) \in \mathcal{I}_{rc}$, the return on the portfolio of risks Y_{ij} in (2) with weights w is given by

$$Y(w) = \sum_{i=1}^r \sum_{j=1}^c w_{ij} Y_{ij} = \sum_{i=1}^r w_{i0} R_i + \sum_{j=1}^c w_{0j} C_j + \sum_{i=1}^r \sum_{j=1}^c w_{ij} U_{ij} = \\ R(w_0^{(row)}) + C(w_0^{(col)}) + U(w), \tag{10}$$

where $R(w_0^{(row)}) = \sum_{i=1}^r w_{i0} R_i$, $C(w_0^{(col)}) = \sum_{j=1}^c w_{0j} C_j$ and $U(w) = \sum_{i=1}^r \sum_{j=1}^c w_{ij} U_{ij}$.

Consider the vector of equal weights $\underline{w}_{rc} = \underbrace{\left(1/(rc), 1/(rc), ..., 1/(rc)\right)}_{rc} \in \mathcal{I}_{rc}$ and the vector $\overline{w}_{rc} = \underbrace{\left(1, 0, ..., 0\right)}_{rc} \in \mathcal{I}_{rc}$ that corresponds to the portfolio of Y_{ij}'s that consists of only one risk.

Observe that the vectors $\underline{w}_0^{(row)}$ and $\underline{w}_0^{(col)}$ that correspond to $\underline{w}_{rc}$ by (8) and (9) consist of equal weights. Namely, $\underline{w}_0^{(row)} = \underbrace{(1/r, ..., 1/r)}_{r} = \underline{w}_r \in \mathcal{I}_r$ and $\underline{w}_0^{(col)} = \underbrace{(1/c, ..., 1/c)}_{c} = \underline{w}_c \in \mathcal{I}_c$.

Similarly, for the weights $\overline{w}_0^{(row)}$ and $\overline{w}_0^{(col)}$ corresponding to $\overline{w}_{rc}$ by (8) and (9) we have $\overline{w}_0^{(row)} = \underbrace{(1, 0, ..., 0)}_{r} = \overline{w}_r \in \mathcal{I}_r$ and $\overline{w}_0^{(col)} = \underbrace{(1, 0, ..., 0)}_{c} = \overline{w}_c \in \mathcal{I}_c$.

The following theorem provides value at risk comparisons for portfolio returns $Y(w)$ in (10) under heavy-tailedness in the risk components R_i, C_j and U_{ij}.[4]

Theorem 1 *Let* $q \in (0, 1/2)$.

(i) If $R_i, C_j, U_{ij} \sim \overline{\mathcal{CSLC}}$, *then* $VaR_q[Y(\underline{w}_{rc})] \leq VaR_q[Y(w)] \leq VaR_q[Y(\overline{w}_{rc})]$ *for all* $w \in \mathcal{I}_{rc}$.

(ii) If $R_i, C_j, U_{ij} \sim \underline{\mathcal{CS}}$, *then* $VaR_q[Y(\underline{w}_{rc})] \geq VaR_q[Y(w)] \geq VaR_q[Y(\overline{w}_{rc})]$ *for all* $w \in \mathcal{I}_{rc}$.

Part (i) of Theorem 1 shows that, similar to the case of independence in Ibragimov (2004, 2005) (see part (ii) of Proposition 1 in Appendix A1), the most diversified portfolio with equal weights $\underline{w}_{rc}$ is preferred to any other portfolio of dependent risks Y_{ij} in (2) under moderate heavy-tailedness. In addition, the least diversified portfolio with weights $\overline{w}_{rc}$ consisting of only one risk is dominated by any other portfolio of risks Y_{ij} with additive common shocks. Part (ii) of Theorem 1 implies that, similar to independence (part (ii) of Proposition 2), the conclusions are reversed under extreme heavy-tailedness. Extreme heavy-tailedness of common shocks R_i, C_j and idiosyncratic shocks U_{ij} in (2) implies optimality of the least diversified portfolio with weights $\overline{w}_{rc}$ with respect to the portfolio value at risk comparisons. In contrast, the portfolio value at risk is maximal for the most diversified portfolio with equal weights $\underline{w}_{rc}$ under such assumptions.

One can also obtain similar comparisons with the extremal portfolio weights $\underline{w}_{rc}$ and $\overline{w}_{rc}$ for the values at risk of the components $R(w_0^{(row)})$, $C(w_0^{(col)})$ and $U(w)$ in decomposition (10) with weights $w \in \mathcal{I}_{rc}$. Namely, the following conclusions hold.

[4]Throughout the paper, we present the results with non-strict inequalities for the values at risk of portfolios considered. All these results can be easily re-formulated in terms of strict inequalities.

Theorem 2 *Let $q \in (0, 1/2)$.*

(i) If $U_{ij} \sim \overline{\mathcal{CSLC}}$, then $VaR_q[U(\underline{w}_{rc})] \leq VaR_q[U(w)] \leq VaR_q[U(\overline{w}_{rc})]$ for all $w \in \mathcal{I}_{rc}$.

(ii) If $U_{ij} \sim \underline{\mathcal{CS}}$, then $VaR_q[U(\underline{w}_{rc})] \geq VaR_q[U(w)] \geq VaR_q[U(\overline{w}_{rc})]$ for all $w \in \mathcal{I}_{rc}$.

(iii) If $R_i \sim \overline{\mathcal{CSLC}}$, then $VaR_q[R(\underline{w}_r)] \leq VaR_q[R(w_0^{(row)})] \leq VaR_q[R(\overline{w}_r)]$ for all $w \in \mathcal{I}_{rc}$.

(iv) If $R_i \sim \underline{\mathcal{CS}}$, then $VaR_q[R(\underline{w}_r)] \geq VaR_q[R(w_0^{(row)})] \geq VaR_q[R(\overline{w}_r)]$ for all $w \in \mathcal{I}_{rc}$.

(v) If $C_j \sim \overline{\mathcal{CSLC}}$, then $VaR_q[C(\underline{w}_c)] \leq VaR_q[C(w_0^{(col)})] \leq VaR_q[C(\overline{w}_c)]$ for all $w \in \mathcal{I}_{rc}$.

(vi) If $C_j \sim \underline{\mathcal{CS}}$, then $VaR_q[C(\underline{w}_c)] \geq VaR_q[C(w_0^{(col)})] \geq VaR_q[C(\overline{w}_c)]$ for all $w \in \mathcal{I}_{rc}$.

As in the case of independence in parts (i) of Propositions 1 and 2 in Appendix A1, it is of interest to also consider value at risk comparisons for general portfolio weights $v, w \in \mathcal{I}_{rc}$ satisfying $v \prec w$ (so that the portfolio with weights v is more diversified than that with weights w). However, such general comparisons cannot be obtained using majorization on $\mathcal{I}_{rc}$. This is because the values at risk $VaR_q[R(w_0^{(row)})] = VaR_q[\sum_{i=1}^r w_{i0} R_i] = VaR_q[\sum_{i=1}^r (\sum_{j=1}^c w_{ij}) R_i]$ and $VaR_q[C(w_0^{(col)})] = VaR_q[\sum_{j=1}^c w_{0j} C_j] = VaR_q[\sum_{j=1}^c (\sum_{i=1}^r w_{ij}) C_j]$ for the components $R(w_0^{(row)})$ and $C(w_0^{(col)})$ in (10) are not symmetric functions of w_{ij}'s. Thus, these functions are neither Schur-concave nor Schur-convex in $w \in \mathcal{I}_{rc}$).[5]

Nevertheless, the above value at risk comparisons with $v \prec w$ are possible for certain portfolio weights v and w that are different from the most diversified portfolio $\underline{w}_{rc}$ and do not correspond to the least diversification with weights $\overline{w}_{rc}$. Some of these value at risk comparisons have a natural interpretation in terms of value at risk analysis for portfolios of equally weighted indices of risks Y_{ij} in (2). These value at risk orderings and the settings where they arise are considered in the next two sections.

[5]This situation is similar to the majorization-based analysis of variance decompositions for linear estimators of location in two-way classification random effects models in Section 13.B in Marshall & Olkin (1979) (see also the next section). As indicated by Marshall & Olkin (1979), neither Schur-convexity nor Schur-concavity (in $w \in \mathcal{I}_{rc}$) holds for the variances $var[R(w_0^{(row)})] = var[\sum_{i=1}^r w_{i0} R_i]$ and $var[C(w_0^{(col)})] = var[\sum_{j=1}^c w_{0j} C_j]$ because these functions are not symmetric functions of w'_{ij}s.

4 Further applications: Portfolio component value at risk analysis

Let $n_{ij} \in \{0,1\}$, $i = 1, ..., r$, $j = 1, ..., c$, be a set of indicator variables. Denote $n_{i0} = \sum_{j=1}^{c} n_{ij}$, $n_{0j} = \sum_{i=1}^{r} n_{ij}$, $n = \sum_{i=1}^{r} n_{i0} = \sum_{j=1}^{c} n_{0j} = \sum_{i=1}^{r}\sum_{j=1}^{c} n_{ij}$.

Section 13.B in Marshall & Olkin (1979) discusses applications of majorization theory in comparisons of variance components for linear estimators based on observations in specifications (2) referred to as two-way classification random effects models. Marshall & Olkin (1979) consider the equal weights $\underline{w}_{ij} = 1/(rc)$ dealt with in the previous section and also the portfolio weights $\tilde{v}_{ij} = n_{i0}/(nc)$, $\tilde{\tilde{v}}_{ij} = n_{0j}/(nr)$ and $\tilde{w}_{ij} = \frac{(n-n_{i0}-n_{0j}+n_{ij})n_{ij}}{n^2-\sum_{i=1}^{r} n_{i0}^2-\sum_{j=1}^{c} n_{0j}^2+n}$, $i = 1, ..., r$, $j = 1, ..., c$, discussed in Koch (1967*a*,*b*). Denote the weight vectors corresponding to the last three choices by $\tilde{v}$, $\tilde{\tilde{v}}$ and $\tilde{w}$.

In the context of portfolio choice, some of the properties of the portfolios with weights $\underline{w}$, $\tilde{v}$, $\tilde{\tilde{v}}$ and $\tilde{w}$ may be summarized as follows. Consider r equally weighted indices comprised, for $i = 1, ..., r$, of risks Y_{ij}, $j = 1, ..., c$, in the ith row of (2). The return on index i is thus given by $Z_i^{(col)} = R_i + \frac{1}{c}\sum_{j=1}^{c} C_j + \frac{1}{c}\sum_{j=1}^{c} U_{ij}$, $i = 1, ..., r$. For $w = (w_1, ..., w_r) \in \mathcal{I}_r$, the return on the portfolio of the indices $i = 1, ..., r$ with returns $Z_i^{(col)}$ and weights w is given by

$$Z^{(col)}(w) = \sum_{i=1}^{r} w_i Z_i^{(col)} = \sum_{i=1}^{r} w_i R_i + \frac{1}{c}\sum_{j=1}^{c} C_j + \sum_{i=1}^{r} \frac{(U_{i1} + ... + U_{ic})}{c} w_i. \tag{11}$$

Similarly, consider, for $j = 1, ..., c$, the equally weighted indices comprised of risks Y_{ij}, $i = 1, ..., r$, in the jth column of (2). The return on index j is thus given by $Z_j^{(row)} = C_j + \frac{1}{r}\sum_{i=1}^{c} R_i + \frac{1}{r}\sum_{i=1}^{r} U_{ij}$, $j = 1, ..., c$. For $w = (w_1, ..., w_c) \in \mathcal{I}_c$, the return on the portfolio of the indices $j = 1, ..., c$ with returns $Z_j^{(row)}$ and weights w is given by

$$Z^{(row)}(w) = \sum_{j=1}^{c} w_j Z_j^{(row)} = \frac{1}{r}\sum_{i=1}^{r} R_i + \sum_{j=1}^{c} w_j C_j + \sum_{j=1}^{c} \frac{(U_{1j} + ... + U_{rj})}{r} w_j. \tag{12}$$

The risk $Y(\tilde{v})$ obtained using (10) with weights $\tilde{v}$ is the same as the return on the portfolio of the indices $Z_i^{(col)}$, $i = 1, ..., r$, with weights $w = \underbrace{(n_{10}/n, ..., n_{r0}/n)}_{r} \in \mathcal{I}_r$ in (11). The return $Y(\tilde{\tilde{v}})$ in (10) with weights $\tilde{\tilde{v}}$ is the same as the return on the portfolio of the risks $Z_j^{(row)}$, $j = 1, ..., c$, with weights $w = \underbrace{(n_{01}/n, ..., n_{0c}/n)}_{c} \in \mathcal{I}_c$ in (12).

As discussed in the previous section, the equal weights $\underline{w}_{rc} = \underbrace{\Big(1/(rc), 1/(rc), ..., 1/(rc)\Big)}_{rc} \in \mathcal{I}_{rc}$ correspond to the most diversified portfolio of the risks Y_{ij}. The return $Y(\underline{w}_{rc})$ with these weights in (10) is the same as the returns $Z^{(col)}(\underline{w}_r)$ and $Z^{(row)}(\underline{w}_c)$ on the portfolios of indices in (11) and (12) with equal weights $\underline{w}_r = \underbrace{(1/r, ..., 1/r)}_{r} \in \mathcal{I}_r$ and $\underline{w}_c = \underbrace{(1/c, ..., 1/c)}_{c} \in \mathcal{I}_c$.

The weights $\tilde{w}$ correspond to a portfolio of Y_{ij} where, in contrast to $\underline{w}_{rc}$, $\tilde{v}$ and $\tilde{\tilde{v}}$, some of the risks Y_{ij} are taken with zero weights that may be due to the risks' unavailability.

Lemma 13.B.2.a in Marshall & Olkin (1979) and relation (7) with $N = rc$ show that the following majorization comparisons hold for the vectors $\underline{w}_{rc}$, $\tilde{v}$, $\tilde{\tilde{v}}$, $\tilde{w}$ and $\overline{w}_{rc}$:

$$\underline{w}_{rc} \prec \tilde{v} \prec \tilde{w} \prec \overline{w}_{rc}, \tag{13}$$

$$\underline{w}_{rc} \prec \tilde{\tilde{v}} \prec \tilde{w} \prec \overline{w}_{rc}. \tag{14}$$

Theorems 1 and 2 imply value at risk comparisons for the portfolio returns $Y(w)$ and its components $R(w_0^{(row)})$, $C(w_0^{(col)})$ and $U(w)$ in (10) between an arbitrary $w \in \mathcal{I}_{rc}$ (for instance, $w = \tilde{v}, \tilde{\tilde{v}}, \tilde{w}$) and the extremal portfolio weights $\underline{w}_{rc}$ and $\overline{w}_{rc}$. In particular, from parts (iii)-(vi) of Theorem 2 it follows that the following comparisons hold for all $q \in (0, 1/2)$ and all $w \in \mathcal{I}_{rc}$ (e.g., for $w = \tilde{v}, \tilde{\tilde{v}}, \tilde{w}$):

$$VaR_q[R(\tilde{\tilde{v}}_0^{(row)})] = VaR_q[R(\underline{w}_r)] \leq VaR_q[R(w_0^{(row)})] \text{ if } R_i \sim \underline{\mathcal{CSLC}}, \tag{15}$$

$$VaR_q[C(\tilde{v}_0^{(col)})] = VaR_q[C(\underline{w}_c)] \leq VaR_q[C(w_0^{(col)})] \text{ if } C_j \sim \underline{\mathcal{CSLC}}, \tag{16}$$

$$VaR_q[R(\tilde{\tilde{v}}_0^{(row)})] = VaR_q[R(\underline{w}_r)] \geq VaR_q[R(w_0^{(row)})] \text{ if } R_i \sim \underline{\mathcal{CS}}, \tag{17}$$

$$VaR_q[C(\tilde{v}_0^{(col)})] = VaR_q[C(\underline{w}_c)] \geq VaR_q[C(w_0^{(col)})] \text{ if } C_j \sim \underline{\mathcal{CS}}. \tag{18}$$

Comparisons (16) and (18) hold as equalities for $w = \tilde{v}$ and relations (15) and (17) hold as equalities for $w = \tilde{\tilde{v}}$.

Similar value at risk comparisons for $VaR_q[U(w)]$ are also obtained between the portfolio with weights $\tilde{w}$ and those with weights $w = \tilde{v}, \tilde{\tilde{v}}$. Namely, the following result holds.

Theorem 3 *Let $q \in (0, 1/2)$. The following value at risk comparisons hold for the component $U(w)$ of decomposition (10):*

(i) If $U_{ij} \sim \overline{\mathcal{CSLC}}$, then $VaR_q[U(\tilde{w})] \geq VaR_q[U(\tilde{v})]$ and $VaR_q[U(\tilde{w})] \geq VaR_q[U(\tilde{\tilde{v}})]$.

(ii) If $U_{ij} \sim \underline{\mathcal{CS}}$, then $VaR_q[U(\tilde{w})] \leq VaR_q[U(\tilde{v})]$ and $VaR_q[U(\tilde{w})] \leq VaR_q[U(\tilde{\tilde{v}})]$.

The main conclusions from the above value at risk comparisons for the portfolio components $U(w)$, $R(w_0^{(row)})$ and $C(w_0^{(row)})$ in (10) with weights $\underline{w}_{rc}$, $\tilde{v}$, $\tilde{\tilde{v}}$, $\tilde{w}$ and $\overline{w}_{rc}$ are summarized as follows. In the case of moderately heavy-tailed common shocks R_i, C_j and idiosyncratic risks U_{ij}, we conclude that full diversification on the level of underlying Y_{ij} with weights $\underline{w}_{rc}$ is preferred to any other portfolio choice, provided all the rc risks are available (left inequality in part (i) of Theorem 1). In particular, the equal weights $\underline{w}_{rc}$ are preferred to less diversification with weights $\tilde{w}$ where some of the risks may be taken with zero weights. Full diversification with weights $\underline{w}_{rc}$ at Y_{ij} is also preferred to the portfolio of equally weighted indices $Z_i^{(col)}$, $i = 1, ..., r$, with weights $w = \underbrace{(n_{10}/n, ..., n_{r0}/n)}_{r} \in \mathcal{I}_r$ and the portfolio of indices with returns $Z_j^{(row)}$, $j = 1, ..., c$, and the weight vector $w = \underbrace{(n_{01}/n, ..., n_{0c}/n)}_{c} \in \mathcal{I}_c$. In turn, comparisons (15) and (16) for the common shock parts $R(w_0^{(row)})$ and $C(w_0^{(row)})$ in (10) and part (i) of Theorem 3 for $U(w)$ suggest that the weight vectors $\tilde{v}$ and $\tilde{\tilde{v}}$ may be preferred to the vector $\tilde{w}$ where some of the risks are included with zero weights. These conclusions are similar to those in Section 13.B in Marshall & Olkin (1979) for variance comparisons of linear estimators based on observations (2) with $ER_i = \mu$, $EC_j = 0$, $EU_{ij} = 0$, $var(R_i) = \sigma_R^2$, $var(C_j) = \sigma_C^2$, $var(U_{ij}) = \sigma_U^2$, $i = 1, ..., r$, $j = 1, ..., c$.

These results are reversed for extremely heavy-tailed risks R_i, C_j and U_{ij}. In such settings, the equal weights $\underline{w}_{rc}$ are dominated by any other portfolio weights (left inequality in part (ii) of Theorem 1). In contrast, the smallest VaR is achieved at the weights $\overline{w}_{rc}$ and the portfolio consisting of only one risk. In particular, the portfolio of indices with returns $Z_i^{(col)}$, $i = 1, ..., r$, and weights $w = \underbrace{(n_{10}/n, ..., n_{r0}/n)}_{r} \in \mathcal{I}_r$ and the portfolio of indices with returns $Z_j^{(row)}$, $j = 1, ..., c$, and $w = \underbrace{(n_{01}/n, ..., n_{0c}/n)}_{c} \in \mathcal{I}_c$ are preferred to the fully diversified portfolio of Y_{ij}'s with equal weights $\underline{w}_{rc}$. Inequalities (17) and (18) for $R(w_0^{(row)})$ and $C(w_0^{(row)})$ and part (ii) of Theorem 3 for $U(w)$ suggest that the weights $\tilde{w}$ may dominate the weights $\tilde{v}$ and $\tilde{\tilde{v}}$.

The results implied by Theorems 2 and 3 also indicate that, in some cases where heavy-tailedness of the common shocks R_i and C_j and that of the idiosyncratic risks U_{ij} is of different degree (say, in the case of the assumptions $U_{ij} \sim \overline{\mathcal{CSLC}}$ combined with $R_i \sim \underline{\mathcal{CS}}$ and $C_j \sim \underline{\mathcal{CS}}$) the portfolio weights $\tilde{v}$ or $\tilde{\tilde{v}}$ may be preferred to $\tilde{w}$, or vice versa. The optimal choice among the portfolios thus depends crucially on the distributional properties of common shocks R_i and C_j and idiosyncratic risks U_{ij}.

5 When heavy-tailedness helps: value at risk for financial indices

In many real world situations, the sets of available risks in (2) or (3) are unbalanced and include unequal number of rows for each column or unequal number of columns for each row. In this section we show how the approach presented in the paper can be applied in the analysis of such settings. We obtain the results for unbalanced analogues of models (3) with one set of "row effects" common shocks R_i and focus on the framework of value at risk comparisons for indices based on risks Y_{ij} in such models. The results may be extended to more general models, including those corresponding to settings with two sets of common shocks in (2) or multiple common shock models (see the discussion in Section 7).

Let $n_1 \geq ... \geq n_r \geq 1$, $\sum_{i=1}^{r} n_i = n$. Similar to the interpretation of weights $\tilde{v}_{ij}$ and $\tilde{\tilde{v}}_{ij}$ in Section 4, consider r equally weighted indices comprised, for $i = 1, ..., r$, of risks

$$Y_{ij} = R_i + U_{ij}, \quad j = 1, ..., n_i. \tag{19}$$

The return on the index i is given by

$$Z_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij} = R_i + \frac{1}{n_i} \sum_{j=1}^{n_i} U_{ij}. \tag{20}$$

As in Sections 3 and 4, for $w = (w_1, ..., w_r) \in \mathcal{I}_r$, denote by $Z(w)$ the return on the portfolio of the indices $i = 1, ..., r$ with returns Z_i in (20) and weights w:

$$Z(w) = \sum_{i=1}^{r} w_i Z_i = \sum_{i=1}^{r} \frac{(Y_{i1} + ... + Y_{in_i})}{n_i} w_i = \sum_{i=1}^{r} w_i R_i + \sum_{i=1}^{r} \frac{(U_{i1} + ... + U_{in_i})}{n_i} w_i. \tag{21}$$

In what follows, for $N \geq 1$, $e_N = \underbrace{(1, ..., 1)}_{N} \in \mathbf{R}^N$ will denote the $N-$vector of ones. In addition, for m row vectors $x^{(k)} \in \mathbf{R}^{1 \times N_k}$, $N_k \geq 1$, $k = 1, ..., m$, we will denote by $x = (x^{(1)}, ..., x^{(m)}) \in$

$\mathbf{R}^{1\times N}$, $N = N_1 + ... + N_m$, the vector with the first N_1 components equal to those of $x^{(1)}$, the next N_2 components equal to those of $x^{(2)}$, and so on: $x_i = x_i^{(1)}$, $i = 1, ..., N_1$; $x_{N_1+i} = x_i^{(2)}$, $i = 1, ..., N_2$; ..., $x_{N_1+N_2+...+N_{m-1}+i} = x_i^{(m)}$, $i = 1, ..., N_m$.

Decomposition (21) can be written as

$$Z(w) = R(w) + U(\tilde{w}), \tag{22}$$

where $R(w)$ is the return on the portfolio of common shocks R_i with weights $w = (w_1, ..., w_r) \in \mathcal{I}_r$, and $U(\tilde{w}) = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \tilde{w}_{ij} U_{ij}$ is the return on the portfolio of idiosyncratic risks U_{ij} with weights $\tilde{w}_{ij} = w_i/n_i$, $j = 1, ..., n_i$, and the corresponding weight vector

$$\tilde{w} = (\tilde{w}_{11}, ..., \tilde{w}_{1n_1}, ..., \tilde{w}_{r1}, ..., \tilde{w}_{rn_r}) = \big(\frac{w_1}{n_1} e_{n_1}, ..., \frac{w_r}{n_r} e_{n_r}\big) \in \mathcal{I}_n. \tag{23}$$

The return on the portfolio of the indices with equal weights

$$w^{(1)} = \underline{w}_r = \underbrace{(1/r, ..., 1/r)}_{r} \in \mathcal{I}_r \tag{24}$$

is given by the sample mean of the risks Z_i, $i = 1, ..., r$:

$$Z(w^{(1)}) = \frac{1}{r}\sum_{i=1}^{r} Z_i = \frac{1}{r}\sum_{i=1}^{r} \frac{(Y_{i1} + ... + Y_{in_i})}{n_i} = \frac{1}{r}\sum_{i=1}^{r} R_i + \frac{1}{r}\sum_{i=1}^{r} \frac{(U_{i1} + ... + U_{in_i})}{n_i}. \tag{25}$$

Similarly, the choice of portfolio weights

$$w^{(2)} = \big(\frac{n_1}{n}, \frac{n_2}{n}, ..., \frac{n_r}{n}\big) \tag{26}$$

produces the return $Z(w^{(2)})$ equal to the sample mean of the underlying risks Y_{ij} in (19):

$$Z(w^{(2)}) = \sum_{i=1}^{r} \frac{n_i}{n} Z_i = \frac{1}{n}\sum_{i=1}^{r}\sum_{j=1}^{n_i} Y_{ij}, \tag{27}$$

that is, in the notations of Sections 3 and 4, $Z(w^{(2)}) = Y(\underline{w}_n)$, where $\underline{w}_n$ is the portfolio with equal weights $\underline{w}_n = \underbrace{(1/n, ..., 1/n)}_{n} \in \mathcal{I}_n$. In other words, while the weights $w^{(1)}$ reflect diversification on the level of indices with returns Z_i, the weights $w^{(2)}$ correspond to (full) diversification on the level of the underlying risks Y_{ij} that comprise these indices (see the discussion in Sections 3 and 4 for formalization of the notions of diversification in terms of majorization relations for portfolio weights at Y_{ij} and the risks $Z_i^{(col)}$ and $Z_j^{(row)}$ in models (2) and (3)).

A number of works in statistics and its applications have focused on the estimation of location in models (19) that are referred to in the fields as two-stage nested design, random effects location models. Several authors have considered variance decompositions and efficiency comparisons for location estimators in such models (see the discussion and reviews in Weiler & Culpin, 1970; Section 13.B in Marshall & Olkin 1979; Birkes, Seely & Azzam 1981; El-Bassiouni & Abdelhafez 2000, and references therein).

Suppose that $ER_i = \mu$, $var(R_i) = \sigma_R^2$, $EU_{ij} = 0$, $var(U_{ij}) = \sigma_U^2$, $j = 1, ..., n_i$, $i = 1, ..., r$. Evidently, for the variables $Z(w)$ in (21), one has

$$var[Z(w)] = \sigma_R^2 V_R(w) + \sigma_U^2 V_U(w), \tag{28}$$

where $V_R(w) = \sum_{i=1}^r w_i^2$ and $V_U(w) = \sum_{i=1}^r \frac{w_i^2}{n_i}$.

In the framework of inference on the mean μ using linear unbiased estimators, Cochran (1954) recommends using the unweighted ($Z(w^{(1)})$ in (25)) and weighted ($Z(w^{(2)})$ in (27)) averages of group means, for large and small values of the intraclass correlation $\gamma = \sigma_R^2/(\sigma_R^2 + \sigma_U^2)$, respectively. Birkes et al. (1981) show that the minimal complete class of linear unbiased estimators of μ is given by $Z(w(c))$ in (21) with $w(c) = (w_1(c), ..., w_r(c)) \in \mathcal{I}_r$,

$$w_i(c) = \frac{n_i[(n_i - 1)c + 1]^{-1}}{\sum_{i=1}^r n_i[(n_i - 1)c + 1]^{-1}}, \quad i = 1, ..., r, \quad 0 \leq c \leq 1. \tag{29}$$

For these estimators, one has $w(0) = w^{(2)}$ and $w(1) = w^{(1)} = \underline{w}_r$. It is straightforward to show that if the intraclass correlation $\gamma = \sigma_R^2/(\sigma_R^2 + \sigma_U^2)$ is known, then the variance $var[Z(w)]$, $w \in \mathcal{I}_r$, in (28) is minimized under the choice of weights $w(\gamma)$. Birkes et al. (1981) further focus on the analysis of efficiency $\mathit{eff}(c, \gamma)$ for estimators $Z(w(c))$ defined as the ratio $\mathit{eff}(c, \gamma) = var[Z(w(c))]/var[Z(w(\gamma))]$ of the variance of $Z(w(c))$ to the least possible variance $var[Z(w(\gamma))]$ of linear unbiased estimators of μ in (19). The authors identify the maximin efficiency estimator $Z(w(c^*))$ that maximizes (over $c \in [0, 1]$) the minimum possible efficiency $\min_{\gamma\in[0,1]} \mathit{eff}(c, \gamma)$. The value c^* is found from the equation $nV_U(w(c^*)) = rV_R(w(c^*))$, that is, $n \sum_{i=1}^r w_i^2(c^*)/n_i = r \sum_{i=1}^r w_i^2(c^*)$.[6]

[6] If one compares the estimators $Z(w)$ by variances instead of efficiencies, then it is easy to show that, as discussed in Birkes et al. (1981), the unweighted average $Z(\underline{w}_r)$ of group means in (25) has the optimal property of being the "minimax variance" linear unbiased estimator of μ in models (19) with fixed $\sigma_R^2 + \sigma_U^2$. More precisely, $var[Z(\underline{w}_r)] = \min_{w\in\mathcal{I}_r} \max_{\gamma\in[0,1]} var[Z(w)]$, where, from (28), $var[Z(w)] = (\sigma_R^2 + \sigma_U^2)[\gamma V_R(w) + (1 - \gamma)V_U(w)]$, and $\max_{\gamma\in[0,1]} var[Z(w)] = (\sigma_R^2 + \sigma_U^2)V_R(w)$ since $V_R(w) \geq V_U(w)$ for all $w \in \mathcal{I}_r$.

Koch (1967a) discusses variance decompositions (28) for the averages $Z(w^{(1)})$ in (25) and $Z(w^{(2)})$ in (27). He shows that the statistics $Z(w^{(1)})$ and $Z(w^{(2)})$ have the opposite orderings of the contributions to their variances in (28) from the row effects and the idiosyncratic error parts V_R and V_U. More precisely, as shown in Koch (1967a), $V_R(w^{(1)}) \leq V_R(w^{(2)})$ and $V_U(w^{(1)}) \geq V_U(w^{(2)})$.[7] Koch (1967$a$) further conjectures that for the weights $w^{(3)} = (w_1^{(3)}, w_2^{(3)}, ..., w_r^{(3)})$ with

$$w_i^{(3)} = \frac{n_i(n - n_i)}{n^2 - \sum_{s=1}^m n_s^2}, \quad i = 1, ..., r, \tag{30}$$

one has

$$V_R[Z(w^{(1)})] \leq V_R[Z(w^{(3)})] \leq V_R[Z(w^{(2)})] \tag{31}$$

and

$$V_U[Z(w^{(1)})] \geq V_U[Z(w^{(3)})] \geq V_U[Z(w^{(2)})]. \tag{32}$$

This conjecture was proven by Low (1970) using some inequalities implied by majorization theory. An alternative more direct proof of the conjecture is provided in Marshall & Olkin (1979), Section 13.B. As discussed by Birkes et al. (1981), the maximin efficiency of $Z(w(c^*))$ compares favorably with efficiency of $Z(w^{(k)})$, $k = 1, 2, 3$.

Theorems 4 and 5 provide value at risk comparisons for equally weighted indices comprised of risks Y_{ij} spanned by heavy-tailed common shocks R_i and idiosyncratic risks U_{ij}. In both of them the degree of heavy-tailedness of common shocks R_i is different from that for idiosyncratic risks U_{ij}.

Theorem 4 concerns the case of extremely heavy-tailed R_i and moderately heavy-tailed U_{ij}.

Theorem 4 *Let $q \in (0, 1/2)$. Suppose that, in (19), $R_i \sim \underline{\mathcal{CS}}$, $i = 1, ...r$, and $U_{ij} \sim \overline{\mathcal{CSLC}}$, $j = 1, ..., n_i$, $i = 1, ..., r$. Then the function $VaR_q[Z(w(c))]$ is non-decreasing in $c \in [0, 1]$. In particular,*

$$VaR_q[Z(w^{(1)})] \geq VaR_q[Z(w(c))] \geq VaR_q[Z(w^{(2)})]$$

for all $c \in [0, 1]$. In addition,

$$VaR_q[Z(w^{(1)})] \geq VaR_q[Z(w^{(3)})] \geq VaR_q[Z(w^{(2)})].$$

[7] Due to a typo, the inequality sign in the second of these two relations is reversed in the review on page 393 in Marshall & Olkin (1979).

Theorem 5 shows that the conclusions of Theorem 4 are reversed in the case where the common shocks R_i are moderately heavy-tailed and the idiosyncratic risks U_{ij} are extremely heavy-tailed.

Theorem 5 *Let $q \in (0, 1/2)$. Suppose that, in (19), $R_i \sim \overline{\mathcal{CSLC}}$, $i = 1, ..., r$, and $U_{ij} \sim \underline{\mathcal{CS}}$, $j = 1, ..., n_i$, $i = 1, ..., r$. Then the function $VaR_q[Z(w(c))]$ is non-increasing in $c \in [0, 1]$. In particular,*

$$VaR_q[Z(w^{(2)})] \geq VaR_q[Z(w(c))] \geq VaR_q[Z(w^{(1)})]$$

for all $c \in [0, 1]$. In addition,

$$VaR_q[Z(w^{(2)})] \geq VaR_q[Z(w^{(3)})] \geq VaR_q[Z(w^{(1)})].$$

Similar to Theorems 1 and 2, the results provided by Theorems 4 and 5 hold for the value at risk comparisons for the components $R(w)$ and $U(\tilde{w})$ in decomposition (22) for the portfolio returns $Z(w)$ (here and below, for weights $w = (w_1, ..., w_r) \in \mathcal{I}_r$ at Z_i's, $\tilde{w} \in \mathcal{I}_n$ is the vector of weights at U_{ij} in (22) that corresponds to w by (23)). These comparisons are provided in Theorem 6.

Theorem 6 *Let $q \in (0, 1/2)$. The following comparisons hold for the components $R(w)$ and $U(\tilde{w})$ in decomposition (22).*

(i) Suppose $R_i \sim \overline{\mathcal{CSLC}}$. Then the function $VaR_q[R(w(c))]$ is non-increasing in $c \in [0, 1]$. In particular,

$$VaR_q[R(w^{(2)})] \geq VaR_q[R(w(c))] \geq VaR_q[R(w^{(1)})]$$

for all $c \in [0, 1]$. In addition, $VaR_q[R(w^{(3)})] \leq VaR_q[R(w^{(2)})]$ and $VaR_q[R(w^{(1)})] \leq VaR_q[R(w)]$ for all $w \in \mathcal{I}_r$.

(ii) Suppose $R_i \sim \underline{\mathcal{CS}}$. Then the function $VaR_q[R(w(c))]$ is non-decreasing in $c \in [0, 1]$. In particular,

$$VaR_q[R(w^{(2)})] \leq VaR_q[R(w(c))] \leq VaR_q[R(w^{(1)})]$$

for all $c \in [0, 1]$. In addition, $VaR_q[R(w^{(3)})] \geq VaR_q[R(w^{(2)})]$ and $VaR_q[R(w^{(1)})] \geq VaR_q[R(w)]$ for all $w \in \mathcal{I}_r$.

(iii) Suppose $U_{ij} \sim \overline{\mathcal{CSLC}}$. *Then the function* $VaR_q[U(\tilde{w}(c))]$ *is non-decreasing in* $c \in [0,1]$. *In particular,*

$$VaR_q[U(\tilde{w}^{(2)})] \leq VaR_q[R(\tilde{w}(c))] \leq VaR_q[R(\tilde{w}^{(1)})].$$

In addition, $VaR_q[U(\tilde{w}^{(2)})] \leq VaR_q[U(\tilde{w}^{(3)})]$ *and* $VaR_q[U(\tilde{w}^{(2)})] = VaR_q[U(\underline{w}_n)] \leq VaR_q[U(\tilde{w})]$ *for all* $w \in \mathcal{I}_r$.

(iv) Suppose $U_{ij} \sim \underline{\mathcal{CS}}$. *Then the function* $VaR_q[U(\tilde{w}(c))]$ *is non-increasing in* $c \in [0,1]$. *In particular,*

$$VaR_q[U(\tilde{w}^{(2)})] \geq VaR_q[R(\tilde{w}(c))] \geq VaR_q[R(\tilde{w}^{(1)})].$$

In addition, $VaR_q[U(\tilde{w}^{(2)})] \geq VaR_q[U(\tilde{w}^{(3)})]$ *and* $VaR_q[U(\tilde{w}^{(2)})] = VaR_q[U(\underline{w}_n)] \geq VaR_q[U(\tilde{w})]$ *for all* $w \in \mathcal{I}_r$.

Let us compare the results provided by Theorems 4-6 with the comparisons for the variances $var[Z(w)]$ in (28) discussed above. If both the common shock variables R_i and the idiosyncratic risks U_{ij} have finite second moments and are thus light-tailed, then solving the optimal portfolio choice problem with minimization of the variance $var[Z(w)]$ is problematic in the following sense. First, the optimal solution is given by the portfolio weights $w_i(\gamma)$ that depend on the value of the intraclass correlation γ which is typically unknown. Second, in (31) and (32), the contributions to the variances of the risks $Z(w)$ from the common shock and the idiosyncratic risk parts V_R and V_U are ordered in the opposite way for the diversified portfolio weights $w^{(1)} = \underline{w}_r$ and $w^{(2)}$. This further holds regardless of the values of σ_R^2 and σ_U^2. Thus, minimization of the variance $var[Z(w)]$ does not point out, even in the case of identically distributed R_i's and U_{ij}'s, to diversification either on the level of indices $i = 1, ..., r$ with returns Z_i (the case of weights $w^{(1)} = \underline{w}_r$) or on the level of underlying risks Y_{ij} (the case of weights $w^{(2)}$).

These conclusions are in sharp contrast with the results for the value at risk portfolio choice under independence discussed in the introduction and reviewed in Appendix A1. They are also in contrast with the results for balanced models (2) presented in Sections 3 and 4. Namely, portfolio value at risk under independence or in balanced models (2) is minimized at equal weights for all moderately heavy-tailed risks with finite first moments (part (ii) of Proposition 1 and part (i) of Theorem 1). Similarly, the solution to the value at risk minimization in such settings is given by the portfolio consisting of one risk within the whole class of extremely heavy-tailed risks with infinite first moments (part (ii) of Proposition 2 and part (ii) of Theorem 1).

In the settings of Theorem 4, the value at risk $VaR_q[Z(w(c))]$ of the portfolios of indices $i = 1, ..., r$, with weights $w(c)$ defined in (29) is non-decreasing in $c \in [0, 1]$. Thus, the choice of portfolio weights $w(0) = w^{(2)}$ in (26) and diversification on the level of underlying risks Y_{ij} is preferred, in terms of value at risk comparisons, to $w(c)$ with any $c \in (0, 1]$. In particular, $w(0) = w^{(2)}$ is preferred to the portfolio of equal weights $w(1) = w^{(1)} = \underline{w}_r$ and the implied diversification on the level of indices. In addition, as shown by Theorem 4, the weight vector $w^{(2)}$ is preferred to $w^{(3)}$ that, in turn, dominates $w^{(1)} = \underline{w}_r$ in terms of the value at risk comparisons for $Z(w)$.

Under the assumptions of Theorem 5, the value at risk $VaR_q[Z(w(c))]$ is non-increasing in $c \in [0, 1]$ for the weights $w(c)$ in (29). Thus, in terms of VaR comparisons, the choice of equal weights $w(1) = w^{(1)} = \underline{w}_r$ and diversification on the level of indices $i = 1, ..., r$, is preferred to $w(c)$ with any $c \in [0, 1)$. The equal portfolio weights are preferred, in particular, to the weights $w(0) = w^{(2)}$ and the implied diversification on the level of individual risks Y_{ij}. Theorem 5 shows that the weight vector $w^{(1)} = \underline{w}_r$ is also preferred to $w^{(3)}$ and $w^{(3)}$ is preferred to $w^{(2)}$.

Parts (i) and (iii) of Theorem 6 show that, if all the variables R_i, U_{ij} (and, thus, the risks Y_{ij} in (19)) are moderately heavy-tailed, then the orderings of the value at risks for the portfolio components $R(w)$ and $U(\tilde{w})$ in (22) with weights $w = w^{(1)}, w^{(2)}, w^{(3)}$ are the same as in the case of variance comparisons in (31) and (32) discussed above. In addition, these comparisons do not point out to optimality of $w^{(1)}$ or $w^{(2)}$ within these three weight vectors or among the weights $w(c)$, $c \in [0, 1]$, in (29). In other words, similar to variance minimization used as the portfolio choice criterion, the value at risk comparisons do not point out to diversification either on the level of indices $i = 1, ..., r$ with returns Z_i (the case of weights $w^{(1)}$) or on the level of underlying risks Y_{ij} (the case of weights $w^{(2)}$).

Parts (ii) and (iv) of Theorem 6 show that the above conclusions are reversed in the case where all the variables R_i, U_{ij} (and, thus, the risks Y_{ij} in (19)) are extremely heavy-tailed. In such setting, the orderings of the value at risks for the portfolio components $R(w)$ and $U(\tilde{w})$ in (22) with weights $w = w^{(1)}, w^{(2)}, w^{(3)}$ are the opposite to those in the case of variance comparisons in (31) and (32) and the case of VaR under moderate heavy-tailedness given by parts (i) and (iii) of Theorem 6. However, again, these orderings do not imply optimality of $w^{(1)}$ or $w^{(2)}$ within these three weight vectors or within the weight vectors $w(c)$, $c \in [0, 1]$. Thus, the orderings do not imply optimality of diversification either on the level of indices $i = 1, ..., r$ with returns Z_i

or the underlying risks Y_{ij}.

6 From risk management to statistics and econometrics: Efficiency of linear estimators in random effects models

As indicated in the introduction, the value at risk results presented in the paper can be reformulated in the framework of the analysis of efficiency of linear estimators in random effects models. As an example, in this section we discuss the implications of the results in Section 5 for efficiency of linear estimators in unbalanced two-stage nested design random effects models (19). Using the results in Sections 3 and 4 and in the next section, similar extensions can be obtained for linear estimators of location in two-way classification random effects models (2) and their analogues with more than two common shocks.

Similar to (19), we consider observations from the model

$$Y_{ij} = \mu + R_i + U_{ij}, \quad j = 1, ..., n_i, \quad i = 1, ..., r, \tag{33}$$

where $\mu \in \mathbf{R}$ and R_i and U_{ij} are symmetric and unimodal r.v.'s. We assume, as before, that the variables R_i and U_{ij} are independent of each other and are independent and identically distributed among themselves.

A natural approach to comparisons of estimators of a population parameter under heavy-tailedness is that based on the likelihood of observing large deviations of these estimators from the true value of the parameter.

Let $\hat{\theta}^{(1)}$ and $\hat{\theta}^{(2)}$ be two estimators of the location parameter μ in model (33). Following the above approach, we say, similar to Ibragimov (2007), that the estimator $\hat{\theta}^{(1)}$ is (weakly) more efficient than $\hat{\theta}^{(2)}$ in the sense of peakedness ($P-$more efficient than $\hat{\theta}^{(2)}$ for short) if $P\big(|\hat{\theta}^{(1)} - \mu| > \epsilon\big) \leq P\big(|\hat{\theta}^{(2)} - \mu| > \epsilon\big)$ for all $\epsilon > 0$.

For $w = (w_1, ..., w_r) \in \mathcal{I}_r$, consider the linear estimators $Z(w)$ of the location parameter μ in form (21), with Z_i, $i = 1, ..., r$, defined in (20): $Z(w) = \sum_{i=1}^{r} w_i Z_i$, $Z_i = \frac{1}{n_i} \sum_{j=1}^{n_i} Y_{ij}$.

As in the Section 5, we deal with the weight vectors $w = w^{(1)}, w^{(2)}, w^{(3)}, w(c)$ defined in (24), (26), (30) and (29). Theorems 7 and 8 below provide $P-$efficiency comparisons for linear estimators $Z(w)$ with the above weights considered, in the context of value at risk analysis, in Section 5.

Theorem 7 *Let* $\epsilon > 0$. *Suppose that, in (3),* $R_i \sim \underline{\mathcal{CS}}$, $i = 1, ...r$, *and* $U_{ij} \sim \overline{\mathcal{CSLC}}$, $j = 1, ..., n_i$, $i = 1, ..., r$. *Then the function* $\tau(c) = P\Big[\big|Z(w(c)) - \mu\big| > \epsilon\Big]$ *is non-decreasing in* $c \in [0, 1]$. *In particular,*

$$P\Big[\big|Z(w^{(1)}) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w(c)) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w^{(2)}) - \mu\big| > \epsilon\Big].$$

In addition,

$$P\Big[\big|Z(w^{(1)}) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w^{(3)}) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w^{(2)}) - \mu\big| > \epsilon\Big].$$

Theorem 8 *Let* $\epsilon > 0$. *Suppose that, in (3),* $R_i \sim \overline{\mathcal{CSLC}}$, $i = 1, ...r$, *and* $U_{ij} \sim \underline{\mathcal{CS}}$, $j = 1, ..., n_i$, $i = 1, ..., r$. *Then the function* $\tau(c) = P\Big[\big|Z(w(c)) - \mu\big| > \epsilon\Big]$ *is non-increasing in* $c \in [0, 1]$. *In particular,*

$$P\Big[\big|Z(w^{(2)}) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w(c)) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w^{(1)}) - \mu\big| > \epsilon\Big].$$

In addition,

$$P\Big[\big|Z(w^{(2)}) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w^{(3)}) - \mu\big| > \epsilon\Big] \geq P\Big[\big|Z(w^{(1)}) - \mu\big| > \epsilon\Big].$$

Theorems 7 and 8 show that $P-$efficiency comparisons in models (33) under heavy-tailedness are similar to VaR results in risk models (19) dealt with in the previously section. In particular, the results in Theorems 7 and 8 are in contrast to the case of variance comparisons for linear estimators $Z(w)$ in the literature discussed in Section 5. Namely, in contrast to the results for variances, under extreme heavy-tailedness in the common shocks R_i and moderate heavy-tailedness in the idiosyncratic risks U_{ij}, $P-$efficiency comparisons point out to optimality of $Z(w^{(2)}) = Z(w(0))$ among the estimators $Z(w(c))$, $c \in [0, 1]$ (Theorem 7). $P-$efficiency of this estimator is also maximal among those of $Z(w^{(k)})$, $k = 1, 2, 3$.

Similarly, in the case of moderate heavy-tailedness in R_i and extreme heavy-tailedness in U_{ij}, $P-$efficiency of the estimators $Z(w(c))$, $c = [0, 1]$, is maximal under equal weights $\underline{w}_r = w(1)$ (Theorem 8). $P-$efficiency of this estimator is also greater than that of $Z(w^{(2)})$ and $Z(w^{(3)})$.

7 Extensions: Multiple additive and multiplicative common shocks

The analysis presented in this paper can be extended to the case where the underlying risks Y in the portfolios exhibit dependence with more than two common shocks. For instance, let $m \geq 1$, and let $N_1, N_2, ..., N_m \in \mathbf{N}$. Denote $L = \prod_{s=1}^{m} N_s$. One can show that the analogues of the results in Section 3 also hold for the multiple shock extensions of (2) given by

$$Y_{i_1,i_2,...,i_m} = \sum_{s=1}^{m} \sum_{1 \leq j_1 < ... < j_s \leq m} U^{(j_1,...,j_s)}_{i_{j_1},...,i_{j_s}}, \tag{34}$$

$1 \leq i_k \leq N_k$, $k = 1, ..., m$, where the variables $U^{(j_1,...,j_s)}_{i_{j_1},...,i_{j_s}}$ are independent over all the indices $1 \leq j_1 < ... < j_s \leq m$, $s = 1, ..., m$, $1 \leq i_k \leq N_k$, $k = 1, ..., m$. The underlying risks Y in (34) thus have dependence structures exhibited by sums of $U-$statistics. Such dependence structures and a number of probabilistic and statistical results for them are discussed, among others, in de la Peña, Ibragimov & Sharakhmetov (2002, 2003) and references therein.

A particular case of models (34) with $m = 3$ is given by the risks

$$Y_{i,j,k} = U^{(1)}_i + U^{(2)}_j + U^{(3)}_k + U^{(4)}_{ij} + U^{(5)}_{ik} + U^{(6)}_{jk} + U_{ijk}, \tag{35}$$

$1 \leq i \leq N_1$, $1 \leq j \leq N_2$, $1 \leq k \leq N_3$, where the summands are independent over all the indices. Specifications (34) also include, for instance, multi-stage nested design random effects models with $U^{(j_1,...,j_s)}_{i_{j_1},...,i_{j_s}} = 0$ for $(j_1, ..., j_s) \neq (1, ..., s)$: $Y_{i_1,i_2,...,i_m} = U^{(1)}_{i_1} + U^{(2)}_{i_1,i_2} + ... + U^{(m)}_{i_1,i_2,...,i_m}$, and multi-way classification random effects models and their analogues (see Koch 1967*a*,*b*).

Consider the portfolios of risks $Y_{i_1,i_2,...,i_m}$ in (34) with weights $w_{i_1,i_2,...,i_m} \in \mathbf{R}_+$ such that $\sum_{i_1=1}^{N_1} ... \sum_{i_m=1}^{N_m} w_{i_1,i_2,...,i_m} = 1$. Let $w \in \mathcal{I}_L$ denote the corresponding weight vector with components $w_{i_1,i_2,...,i_m}$. The return on the portfolio with weights $w_{i_1,i_2,...,i_m}$ is given by $Y(w) = \sum_{i_1=1}^{N_1} ... \sum_{i_m=1}^{N_m} w_{i_1,i_2,...,i_m} Y_{i_1,i_2,...,i_m}$. As in Section 3, one concludes that, in the case of moderately heavy-tailed $U^{(j_1,...,j_s)}_{i_{j_1},...,i_{j_s}} \sim \overline{\mathcal{CSLC}}$, the value at risk $VaR_q[Y(w)]$ of $Y(w)$, $w \in \mathcal{I}_L$ is minimized in the case of the most diversified portfolio with equal weights $\underline{w}_{i_1,i_2,...,i_m} = \underbrace{(1/L, ..., 1/L)}_{L} \in \mathcal{I}_L$. In such settings, the value at risk $VaR_q[Y(w)]$, $w \in \mathcal{I}_L$, is maximized in the case of the least diversified portfolio with weights $\overline{w}_{i_1,i_2,...,i_m} = \underbrace{(1, 0, ..., 0)}_{L} \in \mathcal{I}_L$ that consists of only one risk.

These comparisons are reversed for extremely heavy-tailed $U^{(j_1,...,j_s)}_{i_{j_1},...,i_{j_s}} \sim \underline{\mathcal{CS}}$. Under extreme heavy-tailedness, the equal weights $\underline{w}_{i_1,i_2,...,i_m} = (1/L, ..., 1/L) \in \mathcal{I}_L$ maximize the portfolio value at risk $VaR_q[Y(w)]$ over $w \in \mathcal{I}_L$. In contrast, the minimal portfolio value at risk over $w \in \mathcal{I}_L$ is achieved for the least diversified portfolio with weights $\overline{w}_{i_1,i_2,...,i_m} = (1, 0, ..., 0) \in \mathcal{I}_L$.

The analysis in the paper can also be generalized to the settings where the summands R_i, C_j and U_{ij} in (2) and its analogues, including the case of multiple additive common shocks $U^{(j_1,...,j_s)}_{i_{j_1},...,i_{j_s}}$ in (34), exhibit dependence. For instance, using in the proof the extensions of Propositions 1 and 2 to the case of dependence discussed in Ibragimov (2004, 2005) and Ibragimov & Walden (2007a), one obtains that all the results in the paper also hold in settings where the risks R_i, C_j and U_{ij} in (2) and (3) are dependent among themselves or are bounded. These generalizations include models (2) and (3) in which the vectors of common shocks $(R_1, ..., R_r)$ and $(C_1, ..., C_c)$ and the vector of idiosyncratic errors $(U_{11}, ..., U_{1c}, ..., U_{r1}, ..., U_{rc})$ have distributions which are convolutions of $\alpha-$symmetric distributions (see Ibragimov (2004, 2005) and Ibragimov & Walden (2007a) for the definition and discussion of properties of $\alpha-$symmetric distributions). The class of $\alpha-$symmetric distributions contains, as a subclass, spherical distributions corresponding to the case $\alpha = 2$ (see Fang, Kotz & Ng 1990, p. 184). Spherical distributions, in turn, include such examples as Kotz type, multinormal, multivariate t and multivariate spherically symmetric $\alpha-$stable distributions. In addition, vectors with $\alpha-$symmetric distributions contain important examples of models with multiplicative common shocks. The latter specifications provide generalizations of the results in the paper for the risks

$$Y_{ij} = R'_i + C'_j + U'_{ij}, \quad i = 1, ..., r, \quad j = 1, ..., c, \tag{36}$$

where $R'_i = \sum_{s=1}^{m_1} F_s R_{is}$, $C'_j = \sum_{s=1}^{m_2} G_s C_{js}$, $U'_{ij} = \sum_{s=1}^{m_3} H_s U_{ijs}$, and the risks $F_s, G_s, H_s > 0$ and R_{is}, C_{js}, U_{ijs} are independent of each other and among themselves (the proof of extensions to the case of models (36) can be obtained using conditioning arguments). In models (36), the risks Y_{ij} are thus affected by two additive common shocks R and C and by $m_1 + m_2 + m_3$ common multiplicative shocks F, G and H.

Models (36) with R_{is}, C_{js} and U_{ijs} in the classes $\overline{\mathcal{CSLC}}$ and $\underline{\mathcal{CS}}$ exhibit both heavy-tailedness in R'_i, $C'_j s$, U'_{ij} and dependence among these variables. For instance, the variables FR, GC, HU with $R, C, U \sim S_\alpha(\sigma, 0, 0)$, $\alpha < 1$, have extremely heavy-tailed distributions with infinite means. On the other hand, the products FR, GC, HU with $R, C, U \sim S_\alpha(\sigma, 0, 0)$, $\alpha > 1$, can have marginals with power moments finite up to a certain positive order (or finite exponential moments) depending on the choice of the common multiplicative shocks F, G, H. For instance, these products with $R, C, U \sim S_\alpha(\sigma, 0, 0)$, $1 < \alpha < 2$ and $E[F], E[G], E[H] < \infty$ have finite means but infinite variances. However, marginals of such convolutions have infinite means if the first moments $E[F], E[G], E[H]$ are infinite for the common shocks F, G, H. The moments $E|FR|^p$, $E|GC|^p$, $E|HU|^p$, $p > 0$, of the products FR, GC and HU with Gaussian R, C, U are finite if and only if $E|F|^p, E|G|^p, E|H|^p < \infty$. In particular, all power moments $E|FR|^p$, $E|GC|^p$, $E|HU|^p$, $p > 0$, with Gaussian R, C, U are finite if and only if $E|F|^p, E|G|^p, E|H|^p < \infty$ for all $p > 0$.

Using the discussed extensions, we get that the results in the paper continue to hold for Y_{ij} in (36) if the shocks R_{is}, C_{js} and U_{ijs} satisfy the assumptions in the theorems presented. These generalizations also hold for models with more than two additive shocks, like those in (34).

The results in the paper can also be extended to portfolio choice problems for non-identically distributed risks. These extensions are obtained similar to the arguments for Theorems 1-6, using the results in Proposition 3 in Appendix A1. As an example of such extensions, we provide the analogues of Theorems 4 and 5 for non-identically distributed risks Y_{ij} in (19).

Similar to Section 3, for a weight vector $w = (w_1, ..., w_r) \in \mathcal{I}_r$, we denote by $w_{(1)} \leq \ldots \leq w_{(r)}$ the components of w in increasing order. Theorem 9 provides VaR comparisons in the non-identically distributed case for portfolios of risks Z_i in (20) with weights $(w_{(1)}, \ldots, w_{(r)})$ and the corresponding returns $\tilde{Z}(w) = \sum_{i=1}^r w_{(i)} Z_i$.[8] As follows from Theorem 9, value at risk orderings for portfolios of risks Z_i with ordered weights $w_{(i)}$ and the returns $\tilde{Z}(w)$ are similar to the case of independence in Proposition 3. For the portfolios with ordered weights $w_{(i)}$, in contrast to the unordered weights w_i, diversification is preferable under moderate heavy-tailedness in R_i and U_{ij}. Diversification becomes suboptimal if the risks R_i and U_{ij} are extremely heavy-tailed.

Theorem 9 *Let $\alpha, \alpha' \in (0, 2]$, and let $R_i \sim S_\alpha(\sigma_i, 0, 0)$, $U_{ij} \sim S_{\alpha'}(\sigma'_i, 0, 0)$, $\sigma_i, \sigma'_i > 0$, $j = 1, ..., n_i$, $i = 1, ..., r$, be independent, not necessarily identically distributed, stable risks. Theorem 4 continues to hold for the portfolio returns $\tilde{Z}(w) = \sum_{i=1}^r w_{(i)} Z_i$ if $\alpha, \alpha' \geq 1$, $\sigma_1 \leq ... \leq \sigma_r$*

[8]A certain ordering in the components of the portfolio weight vector $w \in \mathcal{I}_r$ is necessary in the analogues of the majorization results in the non-identically distributed case. This is because one needs to guarantee symmetry of the value at risk $VaR_q[\tilde{Z}(w)]$ of the considered returns $\tilde{Z}(w)$ in the components w_i of w.

and $\sigma_1' \leq ... \leq \sigma_r'$. *Theorem 5 continues to hold for the portfolio returns* $\tilde{Z}(w) = \sum_{i=1}^r w_{(i)} Z_i$ *if* $\alpha', \alpha' \leq 1$, $\sigma_1 \geq ... \geq \sigma_r$ *and* $\sigma_1' \geq ... \geq \sigma_r'$.

8 Conclusion

Our analysis illustrates the generality of the majorization-based approach to the study of portfolio diversification and value at risk. In particular, the results in this paper show that the approach can be used in a wide range of dependent models, including those with multiple additive or multiplicative common shocks.

Similar to the case of independence, the tail index threshold $\alpha = 1$ and finiteness of first moments of some of the risk components is the boundary between the robustness and reversals of the standard results in the variance minimization framework. Usually, these reversals under extreme heavy-tailedness point away from diversification, like the results in Sections 3 and 4.

Surprisingly, however, for some important problems — including the optimal portfolio choice for indices of dependent heavy-tailed risks — the implications are reversed and diversification is optimal, as discussed in Section 5. The value of diversification thus depends crucially on the interplay between the dependence and tail behavior of the risks involved.

Appendix A1. Value at risk comparisons for portfolios of heavy-tailed independent risks

This appendix reviews majorization results and value at risk comparisons for linear portfolios of heavy-tailed risks obtained in Ibragimov (2004, 2005) and, in the context of closely related econometric problems for linear location estimators, in Ibragimov (2007).

Consider N risks $X_1, ..., X_N$. As in Sections 3-5, $X(w) = \sum_{i=1}^N w_i X_i$ denotes the return on the portfolio of X_i's with weights $w = (w_1, ..., w_N) \in \mathbf{R}_+^N$. As before, we denote $\underline{w}_N = \underbrace{(1/N, 1/N, ..., 1/N)}_{N} \in \mathcal{I}_N$ and $\overline{w}_N = \underbrace{(1, 0, ..., 0)}_{N} \in \mathcal{I}_N$. In the notation of Sections 3-5, the expressions $VaR_q[X(\underline{w}_N)]$ and $VaR_q[X(\overline{w}_N)]$ are, thus, the values at risk of the portfolio of X_i, $i = 1, ..., N$, with equal weights and of the portfolio consisting of only one risk.

The following result, implied by Corollary 1.2.2 in Ibragimov (2005) and Theorem 3.1 in

Ibragimov (2007), shows that diversification of a portfolio of independent moderately heavy-tailed risks X_i, $i = 1, ..., N$, with weights $w = (w_1, ..., w_N) \in \mathbf{R}_+^N$, leads to a decrease in the riskiness of its return $X(w) = \sum_{i=1}^N w_i X_i$, as measured by VaR.

Proposition 1 *Let $q \in (0, 1/2)$ and let X_i, $i = 1, ..., N$, be i.i.d. risks such that $X_i \sim \overline{\mathcal{CSLC}}$, $i = 1, ..., N$. Then*

(i) $VaR_q[X(v)] \leq VaR_q[X(w)]$ if $v \prec w$ (in other words, the function $\varphi(w, q) = VaR_q[X(w)]$ is Schur-convex in $w \in \mathbf{R}_+^N$).

(ii) In particular, $VaR_q[X(\underline{w}_N)] \leq VaR_q[X(w)] \leq VaR_q[X(\overline{w}_N)]$ for all $w \in \mathcal{I}_N$.

From Corollary 1.2.3 in Ibragimov (2005) and Theorem 3.2 in Ibragimov (2007) it follows that the results for the portfolio VaR given by Proposition 1 are reversed under the assumption that the distributions of the risks $X_1, ..., X_N$ are extremely heavy-tailed. In such settings, diversification of a portfolio of the risks increases the value at risk of its return. More precisely, the following proposition holds.

Proposition 2 *Let $q \in (0, 1/2)$ and let X_i, $i = 1, ..., N$, be i.i.d. risks such that $X_i \sim \underline{\mathcal{CS}}$, $i = 1, ..., N$. Then*

(i) $VaR_q[X(v)] \geq VaR_q[X(w)]$ if $v \prec w$ (in other words, the function $\varphi(w, q) = VaR_q[X(w)]$, is Schur-concave in $w \in \mathbf{R}_+^N$).

(ii) In particular, $VaR_q[X(\overline{w}_N)] \leq VaR_q[X(w)] \leq VaR_q[X(\underline{w}_N)]$ for all $w \in \mathcal{I}_N$.

Proposition 3 below is implied by Proposition 3.1 in Ibragimov (2007). It provides analogues of the results in Proposition 1 and 2 in the case of not necessarily identically distributed independent heavy-tailed risks $X_1, ..., X_N$.

As in Section 7, for a weight vector $w = (w_1, ..., w_N) \in \mathcal{I}_N$, denote by $w_{(1)} \leq \ldots \leq w_{(N)}$ the components of w in increasing order. Proposition 3 provides VaR comparisons for the portfolios with ordered weights $(w_{(1)}, \ldots, w_{(N)})$ and the corresponding returns $\tilde{X}(w) = \sum_{i=1}^N w_{(i)} X_i$.

Proposition 3 *Let $X_i \sim S_\alpha(\sigma_i, 0, 0)$, $\alpha \in (0, 2]$, be independent, not necessarily identically distributed, stable risks.*

(i) Proposition 1 holds for the portfolio returns $\tilde{X}(w) = \sum_{i=1}^N w_{(i)} X_i$ if $\alpha \geq 1$, $0 < \sigma_1 \leq ... \leq \sigma_n$.

(ii) Proposition 2 holds for the portfolio returns $\tilde{X}(w) = \sum_{i=1}^N w_{(i)} X_i$ if $\alpha \leq 1$, $\sigma_1 \geq ... \geq \sigma_n > 0$.

Appendix A2. Proofs

The proof of Theorems 1, 4 and 5 is based on the results in Theorems 2 and 6; this explains the order of the arguments presented below.

Proof of Theorem 2. Parts (i) and (ii) of Theorem 2 follow from Propositions 1 and 2 and majorization comparisons $\underline{w}_{rc} \prec w \prec \overline{w}_{rc}$ for all $w \in \mathcal{I}_{rc}$ implied by (7) with $N = rc$.

Using (7) with $N = r$ and $N = c$ we conclude that, for the vectors $w_0^{(row)}$ and $w_0^{(col)}$ in (8) and (9) one has

$$\underline{w}_0^{(row)} \prec w_0^{(row)} \prec \overline{w}_0^{(row)}, \tag{37}$$

$$\underline{w}_0^{(col)} \prec w_0^{(col)} \prec \overline{w}_0^{(col)}, \tag{38}$$

where, as in Section 3, $\underline{w}_0^{(row)} = \underline{w}_r = \underbrace{(1/r, 1/r, ..., 1/r)}_{r} \in \mathcal{I}_r$, $\underline{w}_0^{(col)} = \underline{w}_c = \underbrace{(1/c, 1/c, ..., 1/c)}_{c} \in \mathcal{I}_c$, $\overline{w}_0^{(row)} = \underbrace{(1, 0, ..., 0)}_{r} = \overline{w}_r \in \mathcal{I}_r$ and $\overline{w}_0^{(col)} = \underbrace{(1, 0, ..., 0)}_{c} = \overline{w}_c \in \mathcal{I}_c$ are the vectors that correspond to $\underline{w}_{rc}$ and $\overline{w}_{rc}$ by (8) and (9). Majorization comparisons (37) and (38), together with Propositions 1 and 2, imply parts (iii)-(vi) of Theorem 2. ■

Proof of Theorem 1. Let $R_i, C_j, U_{ij} \sim \underline{\mathcal{CS}}$, $i = 1, ..., r$, $j = 1, ..., c$, and let $w \in \mathcal{I}_{rc}$. From part (ii) of Theorem 2 it follows that the risks $U(w)$ in decomposition (10) satisfy

$$VaR_q[U(\underline{w}_{rc})] \geq VaR_q[U(w)] \geq VaR_q[U(\overline{w}_{rc})], \quad q \in (0, 1/2). \tag{39}$$

In addition, from parts (iv) and (vi) of Theorem 2 we conclude that the following value at risk comparisons hold for the components $R(w_0^{(row)})$ and $C(w_0^{(col)})$ in decomposition (10):

$$VaR_q[R(\underline{w}_0^{(row)})] \geq VaR_q[R(w_0^{(row)})] \geq VaR_q[R(\overline{w}_0^{(row)})] \quad q \in (0, 1/2), \tag{40}$$

$$VaR_q[C(\underline{w}_0^{(col)})] \geq VaR_q[C(w_0^{(col)})] \geq VaR_q[C(\overline{w}_0^{(col)})] \quad q \in (0, 1/2), \tag{41}$$

where $\underline{w}_0^{(row)} = \underline{w}_r \in \mathcal{I}_r$, $\overline{w}_0^{(row)} = \overline{w}_r \in \mathcal{I}_r$, $\underline{w}_0^{(col)} = \underline{w}_c \in \mathcal{I}_c$ and $\overline{w}_0^{(col)} = \overline{w}_c \in \mathcal{I}_c$ are the vectors that correspond to $\underline{w}_{rc}$ and $\overline{w}_{rc}$ by (8) and (9).

From Theorem 2.7.6 in Zolotarev (1986), p. 134, and Theorems 1.6 and 1.10 in Dharmadhikari & Joag-Dev (1988), pp. 13 and 20, by induction it follows that the densities of the r.v.'s

$R(w_0^{(row)})$, $C(w_0^{(col)})$ and $U(w)$ are symmetric and unimodal if the assumptions of Theorem 1 hold. From Lemma in Birnbaum (1948) (see also Theorem 3.D.4 on p. 173 in Shaked & Shanthikumar 2007) it follows that if $X_1, ..., X_n$ and $Y_1, ..., Y_n$ are independent absolutely continuous symmetric unimodal r.v.'s such that, for $i = 1, 2, ..., n$, and all $q \in (0, 1/2)$, $VaR_q(X_i) \leq VaR_q(Y_i)$, then $VaR_q(\sum_{i=1}^n X_i) \leq VaR_q(\sum_{i=1}^n Y_i)$ for all $q \in (0, 1/2)$.

This, together with inequalities (39)-(41) implies that, for all $q \in (0, 1/2)$,

$$VaR_q[R(\underline{w}_0^{(row)}) + C(\underline{w}_0^{(col)}) + U(\underline{w}_{rc})] \geq VaR_q[R(w_0^{(row)}) + C(w_0^{(col)}) + U(w)] \geq VaR_q[R(\overline{w}_0^{(row)}) + C(\overline{w}_0^{(col)}) + U(\overline{w}_{rc})].$$

Consequently,

$$VaR_q[Y(\underline{w}_{rc})] \geq VaR_q[Y(w)] \geq VaR_q[Y(\overline{w}_{rc})] \tag{42}$$

for all $q \in (0, 1/2)$. Thus, part (ii) of Theorem 1 holds. Part (i) of Theorem 1 may be proven in a similar way, with the use of parts (i), (iii) and (v) of Theorem 2 instead of parts (ii), (iv) and (vi) of the theorem. ■

Proof of Theorem 3. The theorem follows from parts (i) of Propositions 1 and 2 and the majorization comparisons between $\tilde{v}$ and $\tilde{w}$ and between $\tilde{v}$ and $\tilde{\tilde{w}}$ given by (13) and (14). ■

For the proof of Theorems 4-6 we need a lemma follows from Proposition 5.B.1 in Section 5.B in Marshall & Olkin (1979) applied with condition (a') in that section.

Lemma 1 *(Marshall & Olkin 1979, Proposition 5.B.1) If* $a_1 \geq ... \geq a_r > 0$, $b_1 \geq ... \geq b_r > 0$ *and* b_i/a_i *is non-increasing in* $i = 1, ..., r$, *then*

$$\Big(\frac{a_1}{\sum_{i=1}^r a_i}, ..., \frac{a_r}{\sum_{i=1}^r a_i}\Big) \prec \Big(\frac{b_1}{\sum_{i=1}^r b_i}, ..., \frac{b_r}{\sum_{i=1}^r b_i}\Big). \tag{43}$$

If $0 < a_1 \leq ... \leq a_r$, $0 < b_1 \leq ... \leq b_r$ *and* b_i/a_i *is non-decreasing in* $i = 1, ..., r$, *then*

$$\Big(\frac{a_1}{\sum_{i=1}^r n_i a_i} e_{n_1}, ..., \frac{a_r}{\sum_{i=1}^r n_i a_i} e_{n_r}\Big) \prec \Big(\frac{b_1}{\sum_{i=1}^r n_i b_i} e_{n_1}, ..., \frac{b_r}{\sum_{i=1}^r n_i b_i} e_{n_r}\Big), \tag{44}$$

where, as in Section 5, for $N \geq 1$, $e_N = \underbrace{(1, ..., 1)}_{N} \in \mathbf{R}^N$ denotes the $N-$vector of ones.

Proof of Theorem 6. Consider the vectors $w^{(1)} = \underline{w}_r = \underbrace{(1/r, ..., 1/r)}_{r} \in \mathcal{I}_r$, $w^{(2)}$, $w^{(3)}$ and $w(c)$, $0 \le c \le 1$, defined in (24), (26), (30) and (29). From the left majorization comparison in (7) it follows that

$$w^{(1)} \prec w \tag{45}$$

for all $w \in \mathcal{I}_r$ and, since $\left(\frac{w_1^{(2)}}{n_1} e_{n_1}, ..., \frac{w_r^{(2)}}{n_r} e_{n_r}\right) = \underline{w}_n = \underbrace{(1/n, ..., 1/n)}_{n} \in \mathcal{I}_n$,

$$\left(\frac{w_1^{(2)}}{n_1} e_{n_1}, ..., \frac{w_r^{(2)}}{n_r} e_{n_r}\right) \prec w \tag{46}$$

for all $w \in \mathcal{I}_n$. Let us show, using Lemma 1, that the following majorization relations hold:

$$w^{(3)} \prec w^{(2)} \tag{47}$$

(relation (47) is a part of Lemma 13.B.1.a in Marshall & Olkin 1979);

$$\left(\frac{w_1^{(3)}}{n_1} e_{n_1}, ..., \frac{w_r^{(3)}}{n_r} e_{n_r}\right) \prec \left(\frac{w_1^{(1)}}{n_1} e_{n_1}, ..., \frac{w_r^{(1)}}{n_r} e_{n_r}\right), \tag{48}$$

and

$$w(c') \prec w(c), \tag{49}$$

$$\left(\frac{w_1(c)}{n_1} e_{n_1}, ..., \frac{w_r(c)}{n_r} e_{n_r}\right) \prec \left(\frac{w_1(c')}{n_1} e_{n_1}, ..., \frac{w_r(c')}{n_r} e_{n_r}\right), \tag{50}$$

if $0 \le c < c' \le 1$.

To obtain (47), take $a_i = n_i(n - n_i)$ and $b_i = n_i$, $i = 1, ..., r$, in Lemma 1. Under the assumptions of the theorem, $b_1 \ge ... \ge b_r$. As indicated in the proof of Lemma 13.B.1.a in Marshall & Olkin (1979), because $z_1 \ge z_2$ and $z_1 + z_2 \le 1$ together imply $z_1(1 - z_1) \ge z_2(1 - z_2)$, one also has $a_1 \ge ... \ge a_r$. In addition, evidently, $b_i/a_i = 1/(n - n_i)$ is non-increasing in $i = 1, ..., r$. Consequently, by (43), (47) indeed holds.

To establish (48), take $a_i = n - n_i$ and $b_i = 1/(rn_i)$. Then $a_1 \le ... \le a_r$, $b_1 \le ... \le b_r$ and $b_i/a_i = 1/(rn_i(n - n_i))$ is non-decreasing in $i = 1, ..., r$. Consequently, (48) holds by (44).

Relation (49) is a consequence of (43) applied to $a_i = n_i/((n_i - 1)c' + 1)$ and $b_i = n_i/((n_i - 1)c + 1)$.

Majorization (50) follows from (44) applied to $a_i = 1/((n_i - 1)c + 1)$ and $b_i = 1/((n_i - 1)c' + 1)$.

Theorem 6 now follows from parts (i) of Propositions 1 and 2 and majorization comparisons (45)-(50). ■

Proof of Theorems 4 and 5. Suppose that, in (19), $R_i \sim \overline{\mathcal{CSLC}}$, $i = 1, ..., r$, and $U_{ij} \sim \underline{\mathcal{CS}}$, $j = 1, ..., n_i$, $i = 1, ..., r$. Let $0 \le c < c' \le 1$. Using parts (i) and (ii) of Theorem 6, we obtain

$$VaR_q[R(w(c'))] \le VaR_q[R(w(c))], \tag{51}$$

$$VaR_q[U(\tilde{w}(c'))] \le VaR_q[U(\tilde{w}(c))], \tag{52}$$

$$VaR_q[R(w^{(1)})] \le VaR_q[R(w^{(3)})] \le VaR_q[R(w^{(2)})], \tag{53}$$

$$VaR_q[U(\tilde{w}^{(1)})] \le VaR_q[U(\tilde{w}^{(3)})] \le VaR_q[U(\tilde{w}^{(2)})] \tag{54}$$

for all $q \in (0, 1/2)$.

Similar to the proof of Theorem 1, from Theorem 2.7.6 in Zolotarev (1986), p. 134, and Theorems 1.6 and 1.10 in Dharmadhikari & Joag-Dev (1988), pp. 13 and 20, we conclude that the densities of the r.v.'s $R(w)$ and $U(\tilde{w})$ are symmetric and unimodal under the assumptions of Theorems 4-6. As in the proof of Theorem 1, inequalities (51)-(54), together with Lemma in Birnbaum (1948) and Theorem 3.D.4 on p. 173 in Shaked & Shanthikumar (2007), imply

$$VaR_q[R(w(c')) + U(w(c'))] \le VaR_q[R(w(c)) + U(w(c))], \tag{55}$$

$$VaR_q[R(w^{(1)}) + U(\tilde{w}^{(1)})] \le VaR_q[R(w^{(3)}) + U(\tilde{w}^{(3)})] \le VaR_q[R(w^{(2)}) + U(\tilde{w}^{(2)})] \tag{56}$$

for all $q \in (0, 1/2)$. That is, $VaR_q[Z(w(c'))] \le VaR_q[Z(w(c))]$ and $VaR_q[Z(w^{(1)})] \le VaR_q[Z(w^{(3)})] \le VaR_q[Z(w^{(2)})]$ for all $q \in (0, 1/2)$. This proves Theorem 5.

Theorem 4 for $R_i \sim \underline{\mathcal{CS}}$ and $U_{ij} \sim \overline{\mathcal{CSLC}}$ may be proven in a similar way, with the reversals of the inequality signs in (51)-(56) implied by parts (ii) and (iv) of Theorem 6. ■

Proof of Theorems 7 and 8. As is easy to see, for symmetric r.v.'s X_1 and X_2, $P(|X_1| > \epsilon) \le P(|X_2| > \epsilon)$ for all $\epsilon > 0$ if and only if $VaR_q(X_1) \le VaR_q(X_2)$ for all $q \in (0, 1/2)$. Therefore, Theorems 7 and 8 follow from the value at risk comparisons in Theorems 4 and 5. ■

Proof of Theorem 9. Using (21), we obtain $\tilde{Z}(w) = \sum_{i=1}^{r} w_{(i)} Z_i = \tilde{R}(w) + \tilde{U}(w)$, where $\tilde{R}(w) = \sum_{i=1}^{r} w_{(i)} R_i$, $\tilde{U}(w) = \sum_{i=1}^{r} \frac{w_{(i)}}{n_i} \tilde{U}_i$, $\tilde{U}_i = \sum_{j=1}^{n_i} U_{ij}$. By the assumptions of the theorem and (6), $R_i \sim S_\alpha(\sigma_i, 0, 0)$ and $\tilde{U}_i / n_i = \sum_{j=1}^{n_i} U_{ij}/n_i \sim S_{\alpha'}(\sigma_i' n_i^{1/\alpha' - 1}, 0, 0)$.

Suppose that $\alpha, \alpha' \leq 1$, $\sigma_1 \geq ... \geq \sigma_r$ and $\sigma_1' \geq ... \geq \sigma_r'$. Then $\sigma_1' n_1^{1/\alpha'-1} \geq ... \geq \sigma_r' n_r^{1/\alpha'-1}$, $w_{(1)}/n_1 \leq ... \leq w_{(r)}/n_r$. Using part (ii) of Proposition 3 and majorization comparisons (45), (47) and (49), we obtain $VaR_q[\tilde{R}(w(c'))] \geq VaR_q[\tilde{R}(w(c))]$, $VaR_q[\tilde{R}(w^{(1)})] \geq VaR_q[\tilde{R}(w^{(3)})] \geq VaR_q[\tilde{R}(w^{(2)})]$, $VaR_q[\tilde{U}(w(c'))] \geq VaR_q[\tilde{U}(w(c))]$, $VaR_q[\tilde{U}(w^{(1)})] \geq VaR_q[U(w^{(3)})] \geq VaR_q[U(w^{(2)})]$ for $0 \leq c < c' \leq 1$ and all $q \in (0, 1/2)$. As in the proof of Theorems 4 and 5, from the above inequalities it follows that $VaR_q[Z(w(c'))] \geq VaR_q[Z(w(c))]$ and $VaR_q[Z(w^{(1)})] \geq VaR_q[Z(w^{(3)})] \geq VaR_q[Z(w^{(2)})]$ for all $q \in (0, 1/2)$. The case $\alpha, \alpha' \geq 1$, $\sigma_1 \leq ... \leq \sigma_r$ and $\sigma_1' \leq ... \leq \sigma_r'$ follows in a similar way, with the use of part (i) of Proposition 3. ■

References

Ahn, D. (2007), 'Fat tails and futures market liquidity: Theory and evidence from crude oil and natural gas', *Working paper, Harvard University* .

An, M. Y. (1998), 'Logconcavity versus logconvexity: A complete characterization', *Journal of Economic Theory* **80**, 350–369.

Andrews, D. W. K. (2005), 'Cross-section regression with common shocks', *Econometrica* **73**(5), 1551–1585.

Bagnoli, M. & Bergstrom, T. (2005), 'Log-concave probability and its applications', *Economic theory* **26**, 445–469.

Birkes, D., Seely, J. & Azzam, A.-M. (1981), 'An efficient estimator of the mean in a two-stage nested model', *Technometrics* **23**, 143–148.

Birnbaum, Z. W. (1948), 'On random variables with comparable peakedness', *Annals of Mathematical Statistics* **19**, 76–81.

Bouchard, J.-P. & Potters, M. (2004), *Theory of financial risk and derivative pricing: From statistical physics to risk management*, 2nd edn, Cambridge University Press.

Chavez-Demoulin, V., Nešlehova, J. & Embrechts, P. (2006), 'Quantitative models for operational risk: Extremes, dependence and aggregation', *Journal of Banking and Finance* **30**, 2635–2658.

Cochran, W. G. (1954), 'The combination of estimates from different experiments', *Biometrics* **10**, 101–129.

de la Peña, V. H., Ibragimov, R. & Sharakhmetov, S. (2002), 'On sharp Burkholder-Rosenthal-type inequalities for infinite-degree U-statistics', *Annales de l'Institut Henri Poincaré. Probabilités et Statistiques* **38**, 973–990. En l'honneur de J. Bretagnolle, D. Dacunha-Castelle, I. Ibragimov.

de la Peña, V. H., Ibragimov, R. & Sharakhmetov, S. (2003), 'On extremal distributions and sharp L_p-bounds for sums of multilinear forms', *Annals of Probability* **31**, 630–675.

Dharmadhikari, S. W. & Joag-Dev, K. (1988), *Unimodality, convexity and applications*, Academic Press, Boston.

El-Bassiouni, M. Y. & Abdelhafez, M. E. M. (2000), 'Interval estimation of the mean in a two-stage nested model', *Journal of Statistical Computation and Simulation* **67**, 333–350.

Embrechts, P., Klupperberg, C. & Mikosch, T. (1997), *Modelling extremal events for insurance and finance*, Springer, New York.

Fang, K. T., Kotz, S. & Ng, K. W. (1990), *Symmetric multivariate and related distributions*, Vol. 36 of *Monographs on Statistics and Applied Probability*, Chapman and Hall Ltd., London.

Gabaix, X., Gopikrishnan, P., Plerou, V. & Stanley, H. E. (2003), 'A theory of power-law distributions in financial market fluctuations', *Nature* **423**, 267–270.

Ibragimov, R. (2004), 'Portfolio diversification and value at risk under thick-tailedness', *Harvard University Research Discussion Paper 2086* . Available at http://post.economics.harvard.edu/hier/2005papers/HIER2086.pdf.

Ibragimov, R. (2005), *New majorization theory in economics and martingale convergence results in econometrics.* Ph.D. dissertation, Yale University.

Ibragimov, R. (2007), 'Efficiency of linear estimators under heavy-tailedness: Convolutions of $\alpha-$symmetric distributions', *Econometric Theory* **23**, 501–517.

Ibragimov, R., Jaffee, D. & Walden, J. (2006), 'Non-diversification traps in markets for catastrophic risk', *Review of Financial Studies* . Forthcoming.

Ibragimov, R. & Walden, J. (2007*a*), 'The limits of diversification when losses may be large', *Journal of Banking and Finance* **31**, 2551–2569. http://dx.doi.org/10.1016/j.jbankfin.2006.11.014.

Ibragimov, R. & Walden, J. (2007*b*), 'Portfolio diversification under local, moderate and global deviations from power laws', *Insurance: Mathematics and Economics* . In press. http://dx.doi.org/10.1016/j.insmatheco.2007.06.006.

Jansen, D. W. & de Vries, C. G. (1991), 'On the frequency of large stock returns: Putting booms and busts into perspective', *Review of Economics and Statistics* **73**, 18–32.

Kaas, R., Goovaerts, M. & Tang, Q. (2004), 'Some useful counterexamples regarding comonotonicity', *Belgian Actuarial Bulletin* **4**, 1–4.

Karlin, S. (1968), *Total positivity. Vol. I*, Stanford University Press, Stanford, Calif.

Koch, G. G. (1967*a*), 'A general approach to the estimation of variance components', *Technometrics* **9**, 93–118.

Koch, G. G. (1967*b*), 'A procedure to estimate the population mean in random effects models', *Technometrics* **9**, 577–585.

Loretan, M. & Phillips, P. C. B. (1994), 'Testing the covariance stationarity of heavy-tailed time series', *Journal of Empirical Finance* **1**, 211–248.

Low, L. (1970), 'An application of majorization to comparison of variances', *Technometrics* **12**, 141–145.

Marshall, A. W. & Olkin, I. (1979), *Inequalities: Theory of majorization and its applications*, Academic Press, New York.

McCulloch, J. H. (1997), 'Measuring tail thickness to estimate the stable index alpha: A critique', *Journal of Business and Economic Statistics* **15**, 74–81.

Nešlehova, J., Embrechts, P. & Chavez-Demoulin, V. (2006), 'Infinite mean models and the LDA for operational risk', *Journal of Operational Risk* **1**, 3–25.

Rachev, S. T., Menn, C. & Fabozzi, F. J. (2005), *Fat-tailed and skewed asset return distributions: Implications for risk management, portfolio selection, and option pricing*, Wiley, Hoboken, NJ.

Rachev, S. T. & Mittnik, S. (2000), *Stable Paretian Models in Finance*, Wiley, New York.

Scherer, F. M., Harhoff, D. & Kukies, J. (2000), 'Uncertainty and the size distribution of rewards from innovation', *Journal of Evolutionary Economics* **10**, 175–200.

Shaked, M. & Shanthikumar, J. G. (2007), *Stochastic orders*, Springer, New York.

Silverberg, G. & Verspagen, B. (2007), 'The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance', *Journal of Econometrics* **139**, 318–339.

Weiler, H. & Culpin, D. (1970), 'Variance of weighted means', *Technometrics* **12**, 757–773.

Zolotarev, V. M. (1986), *One-dimensional stable distributions*, American Mathematical Society, Providence.