PORTFOLIO DIVERSIFICATION AND VALUE AT RISK UNDER THICK-TAILEDNESS [1]

Running title: Value at risk

By Rustam Ibragimov[2] [3]
Department of Economics, Harvard University

ABSTRACT

This paper focuses on the study of portfolio diversification and value at risk analysis under heavy-tailedness. We use a notion of diversification based on majorization theory that will be explained in the text. The paper shows that the stylized fact that portfolio diversification is preferable is reversed for extremely heavy-tailed risks or returns. However, the stylized facts on diversification are robust to heavy-tailedness of risks or returns as long as their distributions are moderately heavy-tailed. Extensions of the results to the case of dependence, including convolutions of $\alpha-$symmetric distributions and models with common shocks are provided.

Keywords: value at risk, heavy-tailed risks, portfolios, riskiness, diversification, risk bounds, coherent measures of risk

JEL Classification: G11

---

[1]The results in this paper constitute a part of the author's dissertation "New majorization theory in economics and martingale convergence results in econometrics" presented to the faculty of the Graduate School of Yale University in candidacy for the degree of Doctor of Philosophy in Economics in March, 2005. The results were originally contained in the work circulated in 2003-2005 under the titles "Shifting paradigms: On the robustness of economic models to heavy-tailedness assumptions" and "On the robustness of economic models to heavy-tailedness assumptions".

[2]I am indebted to my advisors, Donald Andrews, Peter Phillips and Herbert Scarf, for all their support and guidance in all stages of the current project. I also thank three anonymous referees, Donald Brown, John Campbell, Aydin Cecen, Gary Chamberlain, Brian Dineen, Darrell Duffie, Paul Embrechts, Frank Fabozzi, Xavier Gabaix, Tilmann Gneiting, Philip Haile, Dwight Jaffee, Samuel Karlin, Benoît Mandelbrot, Alex Maynard, Ingram Olkin, Ben Polak, Stephan Ross, Gustavo Soares, Kevin Song, Johan Walden and the participants at seminars at the Departments of Economics at Yale University, University of British Columbia, the University of California at San Diego, Harvard University, the London School of Economics and Political Science, Massachusetts Institute of Technology, the Université de Montréal, McGill University and New York University, the Division of the Humanities and Social Sciences at California Institute of Technology, Nuffield College, University of Oxford, and the Department of Statistics at Columbia University as well as the participants at the 18th New England Statistics Symposium at Harvard University, April 2004, the International Conference on Stochastic Finance, Lisbon, Portugal, September 2004, and Deutsche Bundesbank Conference on Heavy Tails and Stable Paretian Distributions in Finance and Macroeconomics, Eltville, Germany, November 2005, for helpful comments and discussions.

[3]Department of Economics, Harvard University, Littauer Center, 1805 Cambridge St., Cambridge, MA 02138, USA; Tel.: +1-617-496-4795; Fax: +1-617-495-7730; Email: ribragim@fas.harvard.edu

1 Introduction

1.1 Objectives and key results

Value at risk (VaR) models are frequently used in economics, finance and risk management, providing useful alternatives to the traditional expected utility framework (see, e.g., Embrechts, McNeil & Straumann (2002), Bouchaud & Potters (2004), the papers in Szegö (2004) and Fabozzi, Focardi & Kolm (2006) for a review of the value at risk and related risk measures). Expected utility comparisons are not readily available under heavy-tailedness since moments of the risks or returns in consideration become infinite. The VaR analysis is thus, in many regards, one of the only approaches to portfolio choice and riskiness comparisons that do not impose restrictions on heavy-tailedness of the risks.[4] [5]

VaR models are examples of many models in economics, finance and risk management that have a structure that depends on majorization phenomena for linear combinations of random variables (r.v.'s). The majorization relation is a formalization of the concept of diversity in the components of vectors. Over the past decades, majorization theory, which focuses on the study of majorization pre-ordering and Schur-convex functions that preserve it, has found applications in disciplines ranging from statistics, probability theory and economics to mathematical genetics, linear algebra and geometry. A number of papers in economics used majorization and related concepts in the analysis of income inequality and its effects on the properties of economic models (see, among others, the reviews in Marshall & Olkin (1979) and Ibragimov & Ibragimov (2007) and references therein). Recently, Lapan & Hennessy (2002) and Hennessy & Lapan (2003) applied majorization theory to analyze the portfolio allocation problem.

This paper focuses on the study of portfolio diversification and value at risk analysis under heavy-tailedness using new majorization results for linear combinations of heavy-tailed r.v.'s.

[4]In addition, as follows from the results in Ibragimov & Walden (2007), the value at risk and expected utility comparisons are closely linked: in particular, non-diversification results for truncations of extremely heavy-tailed distributions implied by the results in this paper also continue to hold in the expected utility framework if investors utility function becomes convex at any point in the domain of large losses (convexity of utility functions in the loss domain is one of the key foundations of Prospect theory, see Kahneman & Tversky (1979); it also effectively arises if there is limited liability).

[5]Several recent papers (see among others, Acerbi & Tasche (2002), Tasche (2002)) recommend to use the expected shortfall as a coherent alternative to the value at risk (see Artzner, Delbaen, Eber & Heath (1999), Embrechts, McNeil & Straumann (2002) and Section 3 in this paper for the definition of coherency for measures of risk). However, the expected shortfall, which is defined as the average of the worst losses of a portfolio, requires existence of first moments of risks to be finite. It is not difficult to see that existence of means of the risks in considerations is also required for finiteness of coherent spectral measures of risk (see Acerbi (2002), Cotter & Dowd (2006)) that generalize the expected shortfall.

We provide a new precise formalization of the concept of portfolio diversification on the basis of majorization pre-ordering (see Section 4). We further show that the stylized fact that portfolio diversification is preferable is reversed for a wide class of distributions of risks (Theorem 4.2). The class of distributions for which this is the case is the class of extremely heavy-tailed distributions: a diversification of a portfolio of extremely heavy-tailed risks leads to an increase in the riskiness of their portfolio. The stylized facts on diversification are robust to heavy-tailedness of risks or returns as long as their distributions are moderately heavy-tailed (Theorem 4.1). We also obtain sharp bounds on the portfolio VaR under heavy-tailedness (Remark 4.1) and discuss implications of the results in the analysis of coherency of the value at risk (Remark 4.4).

1.2 Literature on heavy-tailedness in economics and finance

This paper belongs to a large stream of literature in economics and finance that have focused on the analysis of heavy-tailed phenomena. This stream of literature goes back to Mandelbrot (1963) (see also the papers in Mandelbrot (1997) and Fama (1965*b*)), who pioneered the study of heavy-tailed distributions with tails declining as $x^{-\alpha}$, $\alpha > 0$, in these fields. If a model involves a r.v. X with such heavy-tailed distribution, then[6]

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

The r.v. X for which this is the case has finite moments $E|X|^p$ of order $p < \alpha$. However, the moments are infinite for $p \geq \alpha$. Well-known examples of distributions satisfying (1) are stable laws with $\alpha \in (0, 2)$, that is, distributions that are closed under portfolio formation (see Section 2).

It was documented in numerous studies that the time series encountered in many fields in economics and finance are heavy-tailed (see the discussion in Loretan & Phillips (1994), Gabaix, Gopikrishnan, Plerou & Stanley (2003), Rachev, Menn & Fabozzi (2005) and references therein). Mandelbrot (1963) presented evidence that historical daily changes of cotton prices have the tail index $\alpha \approx 1.7$, and thus have infinite variances. Using different models and statistical techniques, subsequent research reported the following estimates of the tail parameters 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 (1996, 1997)); $0.9 < \alpha < 2$ (Rachev & Mittnik (2000)). As discussed in Borak, Härdle & Weron (2005) and references therein, while stable heavy-tailed distributions with $\alpha \in (1, 2)$ provide a good fit for returns on some financial indices and other variables, many asset returns have the tail index greater than 2 or exhibit semi-heavy

[6]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.

tails with finite moments of any order, like hyperbolic or truncated stable distributions. De Vany & Walls (2004) show that stable distributions with tail indices $1 < \alpha < 2$ provide a good model for distributions of profits in motion pictures. Chapter 11 in Rachev, Menn & Fabozzi (2005) discusses and reviews the vast literature that supports heavy-tailedness and the stable Paretian hypothesis (with $1 < \alpha < 2$) for equity and bond return distributions. Bouchaud & 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 indicated in Borak, Härdle & Weron (2005), although there are a number of approaches to heavy-tailedness modeling available in the literature, stable heavy-tailed distributions exhibit several properties that make them appealing in applications. Most importantly, stable distributions provide natural extensions of the Gaussian law since they are the only possible limits for appropriately normalized and centered sums of i.i.d. r.v.'s. This property is useful in representing heavy-tailed financial data as cumulative outcomes of market agents' decisions in response to information they possess. In addition, stable distributions are flexible to accommodate both heavy-tailedness and skewness in data. Furthermore, their multivariate extensions also allow one to model dependence among the risks or returns in consideration (see the discussion in Sections 1.3 and 5). The multivariate extensions also include frameworks where the risks or returns have finite moments of order greater than 2, like in the case of multivariate $t-$ or many other spherical distributions.[7]

As follows from the results discussed in, e.g., Lux (1996), Guillaume, Dacorogna, Davé, Müller & Olsen (1997), Gabaix, Gopikrishnan, Plerou & Stanley (2003), tail exponents are similar for financial and economic time series in different countries. Gabaix, Gopikrishnan, Plerou & Stanley (2003) propose a model where power laws for stock returns (with tail index estimates $\alpha \approx 3$), trading volume (with $\alpha \approx 1.5$) and the number of trades (with $\alpha \approx 3.4$) are explained by trading of large participants, namely, the largest mutual funds whose sizes are estimated to have the tail index $\alpha \approx 1$. Power law (1) with $\alpha = 1$ referred to as Zipf's law has also been found to hold for firm sizes (see Axtell (2001)) and city sizes (see Gabaix (1999*a*,*b*) for the discussion and explanations of the Zipf law for cities). Some studies have indicated that the tail exponent is close to one or slightly less than one for such financial time series as Bulgarian lev/US dollar exchange spot rates and increments of the market time process for Deutsche Bank price record (see Rachev & Mittnik (2000)). As discussed by Nešlehova, Embrechts & Chavez-Demoulin (2006), tail indices less than one are observed for empirical loss distributions of a

[7] In addition, as discussed in Ibragimov & Walden (2007) and Ibragimov, Jaffee & Walden (2008), truncation arguments imply that (non-)diversification results in value at risk models for stable distributions with unbounded support continue to hold in the framework of bounded risks.

number of operational risks. Furthermore, 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. Ibragimov, Jaffee & Walden (2008) show that standard seismic theory implies that the distributions of economic losses from earthquakes have heavy tails with tail indices $\alpha \in [0.6, 1.5]$ that can thus be significantly less than one. These estimates for the tail indices follow from power laws for moment magnitudes of earthquakes. Similar analysis also holds for economic losses from other natural disasters with heavy-tailed physical characteristics surveyed in Ibragimov, Jaffee & Walden (2008).[8]

The fact that a number of economic and financial time series have the tail exponents of approximately equal to or (slightly or even substantially) less than one is important in the context of the results in this paper: as we demonstrate, the conclusions of portfolio value at risk theory for risk distributions with the tail exponents $\alpha < 1$ with infinite means are the opposites of those for distributions with $\alpha > 1$ for which the first moment is finite.

1.3 Extensions of the results

To illustrate the main ideas of the proof and in order to simplify the presentation of the main results in this paper, we first model heavy-tailedness using the framework of independent stable distributions and their convolutions. More precisely, the class of moderately heavy-tailed distributions is first modelled using convolutions of stable distributions with (different) indices of stability greater than one. Similarly, the results of the paper for extremely heavy-tailed cases are first presented and proven using the framework of convolutions of stable distributions with characteristic exponents less than one.

The proof of the results in the benchmark case of convolutions of independent stable distributions exploits several symmetries in the problem under analysis in these settings. First, the property that i.i.d. stable distributions are closed under convolutions (relation (3)), together with positive homogeneity of the value at risk (relation a3 in Section 3), allows one to reduce the portfolio VaR analysis for i.i.d. stable risks to comparisons of functions of portfolio weights

[8]One should note here that commonly used approaches to inference on the tail indices, such as OLS log-log rank-size regression estimators and Hill's estimator, are strongly biased in small samples and are very sensitive to deviations from power laws (1) in the form of regularly varying tails (see, among others, the discussion in Embrechts, Klüppelberg & Mikosch (1997), Weron (2001), Borak, Härdle & Weron (2005) and Gabaix & Ibragimov (2006)). In particular, these procedures tend to overestimate the tail index for observations from infinite variance stable distributions with $\alpha < 2$ and sample sizes typical in applications (see McCulloch (1997) and Borak, Härdle & Weron (2005)). Therefore, point estimates of the tail index greater than 1 do not necessarily exclude heavy-tailedness with infinite means and true values $\alpha < 1$ in the same way as point estimates of the tail exponent greater than 2 do not necessarily exclude stable regimes with infinite variances, as discussed in McCulloch (1997) and Weron (2001).

that are Schur-convex or Schur-concave and are thus symmetric in their arguments (see Section 4 for definition of Schur-convex and Schur-concave functions and their symmetry property in (6)). Convolution and VaR properties of symmetric and unimodal distributions reviewed in Appendix A1 then allow us to transfer the value at risk comparisons in the i.i.d. stable case to convolutions of independent stable risks.

In Section 5 we show that the results obtained in the paper also hold for convolutions of risks with joint $\alpha-$symmetric distributions that exhibit both heavy-tailedness and dependence.[9] Convolutions of $\alpha-$symmetric distributions contain, as subclasses, convolutions of several models with common shocks affecting all heavy-tailed risks as well as spherical distributions which are $\alpha-$symmetric with $\alpha = 2$. Spherical distributions, in turn, include such examples as Kotz type, multinormal, logistic and multivariate $\alpha-$stable distributions. In addition, they include a subclass of mixtures of normal distributions as well as multivariate $t-$distributions that were used in the literature to model heavy-tailedness phenomena with dependence and finite moments up to a certain order.

Similar to the framework based on stable distributions, the stylized facts on portfolio diversification hold for convolutions of $\alpha-$symmetric distributions with $\alpha > 1$. The stylized facts are reversed in the case of convolutions of $\alpha-$symmetric distributions with $\alpha < 1$.

As discussed in Appendix A3 available on the author's website, the results in the paper also hold for heterogenous risks and skewed heavy-tailed risks.[10] The results in Appendix A3 further suggest that extensions of the majorization pre-ordering, such as $p-$majorization, may be helpful in formalizations of portfolio diversification under heterogeneity.

Besides the analysis of portfolio diversification under heavy-tailedness, the results on portfolio VaR comparisons and their analogues on majorization properties of tail probabilities of linear combinations of r.v.'s have a number of other applications. These applications, presented, for the most part, in Ibragimov (2005), include the study of robustness of efficiency of linear estimators, firm growth theory for firms that can invest in information about their markets and optimal strategies for a multiproduct monopolist, as well that of inheritance models in mathematical evolutionary theory. The results in this paper and those in Ibragimov (2005) demonstrate that

[9] An $n-$dimensional distribution is called $\alpha-$symmetric if its characteristic function can be written as $\phi((\sum_{i=1}^n |t_i|^\alpha)^{1/\alpha})$, where ϕ is a continuous function and $\alpha > 0$. Such distributions should not be confused with multivariate spherically symmetric stable distributions, which have characteristic functions $\exp\left[-\lambda\left(\sum_{i=1}^n t_i^2\right)^{\beta/2}\right]$, $0 < \beta \leq 2$. Obviously, spherically symmetric stable distributions are particular examples of $\alpha-$symmetric distributions with $\alpha = 2$ (that is, of spherical distributions) and $\phi(x) = \exp(-x^\beta)$.

[10] As recently demonstrated in Ibragimov & Walden (2007), the conclusions in this paper also continue to hold for a wide class of bounded r.v.'s concentrated on a sufficiently large interval with distributions given by truncations of stable and $\alpha-$symmetric ones.

many economic models are robust to heavy-tailedness assumptions as long as the distributions entering these assumptions are moderately heavy-tailed. But the implications of these models are reversed for distributions with extremely heavy tails.

1.4 Organization of the paper

The paper is organized as follows: Section 2 contains notation and definitions of classes of heavy-tailed distributions used throughout the paper and reviews their properties. Section 3 discusses the definition of the value at risk and coherent risk measures and summarizes some relevant properties of the VaR needed for the analysis. Section 4 discusses the definition of majorization pre-ordering and introduces the formalization of the concept of portfolio diversification on its basis. It further presents the main results of the paper on the effects of diversification of a portfolio on its riskiness. Section 5 discusses extensions of the results in the paper to the case of dependence, including convolutions of α-symmetric and spherical distributions and models with common shocks. Section 6 makes some concluding remarks. Appendix A1 summarizes auxiliary results on VaR comparisons and unimodality properties for log-concave and stable distributions needed in the arguments for the main results. Appendix A2 contains the proofs of the results obtained in the paper. Appendix A3 available on the author's website discusses extensions in the case of skewness and heterogeneity.

2 Notation and classes of distributions

In what follows, a univariate density $f(x)$, $x \in \mathbf{R}$, will be referred to as symmetric (about zero) if $f(x) = f(-x)$ for all $x > 0$. In addition, as usual, an absolutely continuous distribution or a r.v. X with the density $f(x)$ will be called symmetric if $f(x)$ is symmetric (about zero).[11]

Throughout the paper, for two r.v.'s X and Y, we write $X =^d Y$ if X and Y have the same distribution.

A r.v. X with density $f(x)$, $x \in \mathbf{R}$, 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) and Bagnoli & Bergstrom (2005)). 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 $\alpha \geq 1$,

[11]This concept of (univariate) symmetry is not to be confused with joint $\alpha-$symmetric or spherical distributions discussed in Sections 1.3 and 5, which provide models of dependence among the components of random vectors.

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 $\alpha \geq 1$. The class of log-concave distributions is closed under convolution. Log-concave distributions have many other 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) and 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 all the power moments $E|X|^\gamma$, $\gamma > 0$, of the r.v. are finite (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.)

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

$x \in \mathbf{R}$, 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 (2) 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 $S_\alpha(\sigma,\beta,\mu)$.

A closed form expression for the density $f(x)$ of the distribution $S_\alpha(\sigma,\beta,\mu)$ is available in the following cases (and only in those cases): $\alpha = 2$ (Gaussian distributions); $\alpha = 1$ and $\beta = 0$ (Cauchy distributions with densities $f(x) = \sigma/(\pi(\sigma^2 + (x-\mu)^2)))$; $\alpha = 1/2$ and $\beta = \pm 1$ (Lévy distributions that have densities $f(x) = (\sigma/(2\pi))^{1/2} exp(-\sigma/(2x)) x^{-3/2}$, $x \geq 0$; $f(x) = 0$, $x < 0$, where $\sigma > 0$, and their shifted versions). Degenerate distributions correspond to the limiting case $\alpha = 0$.

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 $X \sim S_\alpha(\sigma,\beta,\mu)$, $0 < \alpha < 2$, then the distribution of X 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.

The symmetry index β characterizes the skewness of the distribution. The densities $f(x)$ of stable distributions with $\beta = 0$ are symmetric about the location parameter $\mu : f(\mu+x) = f(\mu-x)$ for all $x > 0$. The stable distributions with $\beta = \pm 1$ and $\alpha \in (0,1)$ (and only they) are one-sided, the support of these distributions is the semi-axis $[\mu, \infty)$ for $\beta = 1$ and is $(-\infty, \mu]$ for $\beta = -1$ (in particular, the Lévy distribution with $\mu = 0$ is concentrated on the positive semi-axis for $\beta = 1$ and on the negative semi-axis for $\beta = -1$). In the case $\alpha > 1$ the location parameter μ is the mean of the distribution $S_\alpha(\sigma, \beta, \mu)$. The scale parameter σ is a generalization of the concept of standard deviation; it coincides with the standard deviation in the special case of Gaussian distributions ($\alpha = 2$). 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 $c_i \geq 0$, $i = 1, ..., n$, such that $\sum_{i=1}^n c_i \neq 0$,

$$\sum_{i=1}^n c_i X_i / \Big(\sum_{i=1}^n c_i^\alpha \Big)^{1/\alpha} =^d X_1 \tag{3}$$

(convolution property (3) is implied by defining relation (2) and the product decomposition of the c.f. of linear combinations of stable r.v.'s under independence).

For a detailed review of properties of stable and power-law distributions the reader is referred to Zolotarev (1986), Uchaikin & Zolotarev (1999), Bouchaud & Potters (2004) and Borak, Härdle & Weron (2005).

For $0 < r < 2$, we denote by $\overline{\mathcal{CS}}(r)$ the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in (r, 2]$ 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 *greater* than the threshold value r are taken). That is, $\overline{\mathcal{CS}}(r)$ consists 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 such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (r, 2]$, $\sigma_i > 0$, $i = 1, ..., k$.

Further, for $0 < r \leq 2$, $\underline{\mathcal{CS}}(r)$ stands for the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with indices of stability $\alpha \in (0, r)$ and $\sigma > 0$ (the underline indicates considering stable distributions with indices of stability *less* than the threshold value r). That is, $\underline{\mathcal{CS}}(r)$ consists 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 such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, r)$, $\sigma_i > 0$, $i = 1, ..., k$.

Finally, we denote by $\overline{\mathcal{CSLC}}$ the class of convolutions of distributions from the classes $\mathcal{LC}$ and $\overline{\mathcal{CS}}(1)$. That is, $\overline{\mathcal{CSLC}}$ is the class of convolutions of symmetric distributions which are either log-concave or stable with characteristic exponents greater 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 $\overline{\mathcal{CS}}(1)$.[12]

[12]Remark 4.2 and Appendix A1 discuss the properties that explain, in particular, why one can pool the risks

All the classes $\mathcal{LC}$, $\overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}}(r)$ and $\underline{\mathcal{CS}}(r)$ 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 $\overline{\mathcal{CS}}(1)$.

A linear combination of independent stable r.v.'s with the *same* characteristic exponent α also has a stable distribution with the same α. However, in general, this does not hold in the case of convolutions of stable distributions with *different* indices of stability. Therefore, the class $\overline{\mathcal{CS}}(r)$ of *convolutions* of symmetric stable distributions with *different* indices of stability $\alpha \in (r, 2]$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (r, 2]$ and $\sigma > 0$. Similarly, the class $\underline{\mathcal{CS}}(r)$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (0, r)$ and $\sigma > 0$.

Clearly, $\overline{\mathcal{CS}}(1) \subset \overline{\mathcal{CSLC}}$ and $\mathcal{LC} \subset \overline{\mathcal{CSLC}}$. It should also be noted that the class $\overline{\mathcal{CSLC}}$ is wider than the class of (two-fold) convolutions of log-concave distributions with stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (1, 2]$ and $\sigma > 0$.

By definition, for $0 < r_1 < r_2 \leq 2$, the following inclusions hold: $\overline{\mathcal{CS}}(r_2) \subset \overline{\mathcal{CS}}(r_1)$ and $\underline{\mathcal{CS}}(r_1) \subset \underline{\mathcal{CS}}(r_2)$. Cauchy distributions $S_1(\sigma, 0, 0)$ are at the dividing boundary between the classes $\underline{\mathcal{CS}}(1)$ and $\overline{\mathcal{CS}}(1)$ (and between the classes $\underline{\mathcal{CS}}(1)$ and $\overline{\mathcal{CSLC}}$). Similarly, for $r \in (0, 2)$, stable distributions $S_r(\sigma, 0, 0)$ with the characteristic exponent $\alpha = r$ are at the dividing boundary between the classes $\underline{\mathcal{CS}}(r)$ and $\overline{\mathcal{CS}}(r)$. Further, normal distributions $S_2(\sigma, 0, 0)$ are at the dividing boundary between the class $\mathcal{LC}$ of log-concave distributions and the class $\underline{\mathcal{CS}}(2)$ of convolutions of symmetric stable distributions with indices of stability $\alpha < 2$. More precisely, the Cauchy distributions $S_1(\sigma, 0, 0)$ are the only ones that belong to all the classes $\underline{\mathcal{CS}}(r)$ with $r > 1$ and all the classes $\overline{\mathcal{CS}}(r)$ with $r < 1$. Stable distributions $S_r(\sigma, 0, 0)$ are the only ones that belong to all the classes $\underline{\mathcal{CS}}(r')$ with $r' > r$ and all the classes $\overline{\mathcal{CS}}(r')$ with $r' < r$. Normal distributions are the only distributions belonging to the class $\mathcal{LC}$ and all the classes $\overline{\mathcal{CS}}(r)$ with $r \in (0, 2)$.

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

The properties of stable distributions discussed above imply that the $p-$th absolute moments $E|X|^p$ of a r.v. $X \sim \overline{\mathcal{CS}}(r)$, $r \in (0, 2)$, are finite if $p \leq r$. On the other hand, the r.v.'s $X \sim \underline{\mathcal{CS}}(r)$, $r \in (0, 2]$ have infinite moments of order r: $E|X|^r = \infty$. In particular, the distributions of r.v.'s X from the class $\underline{\mathcal{CS}}(1)$ 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 means: $E|X| < \infty$.

from the classes $\mathcal{LC}$ and $\overline{\mathcal{CS}}(1)$ in the results obtained in the paper.

3 Value at risk: definition and the main properties

In this section, we discuss some properties of the value at risk needed in the analysis throughout the paper.

In what follows, given a loss probability $q \in (0,1)$ and a r.v. (risk) X we denote by $VaR_q(X)$ the value at risk (VaR) of X at level q, that is, its $(1-q)-$quantile: $VaR_q(X) = \inf\{z \in \mathbf{R} : P(X > z) \le q\}$ (throughout the paper, we interpret the positive values of risks X as a risk holder's losses).

Let $\mathcal{X}$ be a certain linear space of r.v.'s X defined on a probability space $(\Omega, \Im, P)$. We assume that $\mathcal{X}$ contains all degenerate r.v.'s $X \equiv a \in \mathbf{R}$. According to the definition in Artzner, Delbaen, Eber & Heath (1999) (see also Embrechts, McNeil & Straumann (2002), Frittelli & Gianin (2002)), a functional $\mathcal{R} : \mathcal{X} \to R$ is said to be a *coherent* measure of risk if it satisfies the following axioms:

a1. (Monotonicity) $\mathcal{R}(X) \ge \mathcal{R}(Y)$ for all $X, Y \in \mathcal{X}$ such that $Y \le X$ (a.s.), that is, $P(X \le Y) = 1$.

a2. (Translation invariance) $\mathcal{R}(X + a) = \mathcal{R}(X) + a$ for all $X \in \mathcal{X}$ and any $a \in \mathbf{R}$.

a3. (Positive homogeneity) $\mathcal{R}(\lambda X) = \lambda \mathcal{R}(X)$ for all $X \in \mathcal{X}$ and any $\lambda \ge 0$.

a4. (Subadditivity) $\mathcal{R}(X + Y) \le \mathcal{R}(X) + \mathcal{R}(Y)$ for all $X, Y \in \mathcal{X}$.

In some papers (see Fölmer & Schied (2002), Frittelli & Gianin (2002)), the axioms of positive homogeneity and subadditivity are replaced by the following weaker axiom of convexity:

a5. (Convexity) $\mathcal{R}(\lambda X + (1-\lambda)Y) \le \lambda \mathcal{R}(X) + (1-\lambda)\mathcal{R}(Y)$ for all $X, Y \in \mathcal{X}$ and any $\lambda \in [0,1]$

(clearly, a5 follows from a3 and a4).

It is easy to verify that the value at risk $VaR_q(X)$ satisfies the axioms of monotonicity, translation invariance and positive homogeneity a1, a2 and a3. However, as follows from the counterexamples constructed by Artzner, Delbaen, Eber & Heath (1999) and Embrechts, McNeil & Straumann (2002), in general, it fails to satisfy the subadditivity and convexity properties a4 and a5 (see Remarks 4.3 and 4.4 for implications of the results in this paper for coherency of the value at risk and the asymptotic analysis in the case of distributions with regularly varying tails).

4 Main results: portfolio diversification and (non-)coherency of the value at risk under heavy-tailedness

This paper demonstrates that majorization theory provides powerful tools for portfolio value at risk analysis.

Throughout the paper, $\mathbf{R}_+$ stands for $\mathbf{R}_+ = [0, \infty)$. 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:

$$\Big(\sum_{i=1}^n a_i/n, ..., \sum_{i=1}^n a_i/n\Big) \prec (a_1, ..., a_n) \prec \Big(\sum_{i=1}^n a_i, 0, ..., 0\Big), \quad a \in \mathbf{R}_+^n, \tag{4}$$

for all $a \in \mathbf{R}_+^n$. In particular,

$$(1/(n+1), ..., 1/(n+1), 1/(n+1)) \prec (1/n, ..., 1/n, 0), \quad n \geq 1. \tag{5}$$

It is also immediate that if $a = (a_1, ..., a_n) \prec b = (b_1, ..., b_n)$, then $(a_{\pi(1)}, ..., a_{\pi(n)}) \prec (b_{\pi(1)}, ..., b_{\pi(n)})$ for all permutations π of the set $\{1, ..., n\}$.

A function $\phi : \mathbf{R}_+^n \to \mathbf{R}$ is called *Schur-convex* (resp., *Schur-concave*) if $(a \prec b) \Longrightarrow (\phi(a) \leq \phi(b))$ (resp. $(a \prec b) \Longrightarrow (\phi(a) \geq \phi(b))$ for all $a, b \in \mathbf{R}_+^n$. If, in addition, $\phi(a) < \phi(b)$ (resp., $\phi(a) > \phi(b)$) whenever $a \prec b$ and a is not a permutation of b, then ϕ is said to be *strictly* Schur-convex (resp., *strictly* Schur-concave) on A. Evidently, if $\phi : \mathbf{R}_+^n \to \mathbf{R}$ is Schur-convex or Schur-concave, then

$$\phi(a_1, ..., a_n) = \phi(a_{\pi(1)}, ..., a_{\pi(n)}) \tag{6}$$

for all $(a_1, ..., a_n) \in \mathbf{R}_+^n$ and all permutations π of the set $\{1, ..., n\}$.

For $w = (w_1, ..., w_n) \in \mathbf{R}_+^n$, denote by Z_w the return on the portfolio of risks $X_1, ..., X_n$ with weights w. Several results in the paper do not require the assumption that $\sum_{i=1}^n w_i = 1$ for the portfolio weights w_i, $i = 1, ..., n$. If this is the case, we write that w belongs to the simplex $\mathcal{I}_n = \{w = (w_1, ..., w_n) : w_i \geq 0, i = 1, ..., n, \sum_{i=1}^n w_i = 1\} : w \in \mathcal{I}_n$.

Denote $\underline{w} = (1/n, 1/n, ..., 1/n) \in \mathcal{I}_n$ and $\overline{w} = (1, 0, ..., 0) \in \mathcal{I}_n$. The expressions $VaR_q(Z_{\underline{w}})$ and $VaR_q(Z_{\overline{w}})$ are, thus, the values at risk of the portfolio with equal weights and of the portfolio consisting of only one return (risk).

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 risks (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 so that, for example, the portfolio with equal weights $\underline{w}$ is the most diversified and the portfolio with weights $\overline{w}$ consisting of one risk is the least diversified among all 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).[13]

Observe that the above formalization of diversification is applicable in the case where some components of v or w are zero. Thus, it allows one to compare portfolios with possibly unequal number of (non-redundant) risks. For instance, portfolio diversification orderings implied by majorization relations (5) are intuitive. It is natural to think about the portfolio of X_i, $i = 1, ..., n+1$, with weights $(1/(n+1), ..., 1/(n+1), 1/(n+1)) \in \mathcal{I}_{n+1}$ as being more diversified than the portfolio of X_i, $i = 1, ..., n$, with weights $(1/n, ..., 1/n) \in \mathcal{I}_n$ since the former portfolio contains an additional non-redundant risk X_{n+1}.

Examples of strictly Schur-convex functions $\phi : \mathbf{R}^n_+ \to \mathbf{R}$ are given by $\phi(w_1, ..., w_n) = \sum_{i=1}^n w_i^2$ and, more generally, by $\phi_\alpha(w_1, ..., w_n) = \sum_{i=1}^n w_i^\alpha$ for $\alpha > 1$. The functions $\phi_\alpha(w_1, ..., w_n)$ are strictly Schur-concave for $\alpha < 1$ (see Proposition 3.C.1.a in Marshall & Olkin (1979)).[14]

A simple example where diversification is preferable is provided by the standard case with normal risks. Let $n \geq 2$, $q \in (0, 1/2)$, and let $X_1, ..., X_n \sim S_2(\sigma, 0, 0)$ be i.i.d. symmetric normal r.v.'s. Then, for the portfolio of X_i's with the equal weights $\underline{w} = (1/n, 1/n, ..., 1/n)$ we have $Z_{\underline{w}} = (1/n)\sum_{i=1}^n X_i =^d (1/\sqrt{n})X_1$. Consequently, by positive homogeneity of the VaR given by property a3 in Section 3, $VaR_q(Z_{\underline{w}}) = (1/\sqrt{n})VaR_q(X_1) = (1/\sqrt{n})VaR_q(Z_{\overline{w}}) < VaR_q(Z_{\overline{w}})$. That is, the value at risk of the most diversified portfolio with equal weights $\underline{w}$ is less than that of the least diversified portfolio with weights $\overline{w}$ consisting of only one risk Z_1.

Theorem 4.1 shows that similar results also hold for all moderately heavy-tailed risks X_i with arbitrary weights $w = (w_1, ..., w_n) \in \mathbf{R}^n_+$.[15] In all these settings, diversification of a portfolio of

[13]Heterogeneity in distributions of risks may require altering of formalizations of portfolio diversification using majorization. Appendix A3 contains some suggestions on these extensions motivated by VaR analysis for skewed and non-identically distributed risks.

[14]The functions ϕ_α have the same form as measures of diversification considered in Bouchaud & Potters (2004), Ch. 12, p. 205.

[15]In particular, the results Theorems 4.1 and 4.2 and their analogues under dependence provided by Theorems 5.1 and 5.2 substantially generalize the riskiness analysis for uniform (equal weights) portfolios of independent stable risks considered, among others, in Fama (1965*a*), Ross (1976), Samuelson (1967): these theorems demonstrate that the formalization of portfolio diversification on the basis of majorization pre-ordering allows one to obtain comparisons of riskiness for portfolios of heavy-tailed and possibly dependent risks with arbitrary, rather than equal, weights.

$X_i's$ leads to a decrease in the value at risk of its return $Z_w = \sum_{i=1}^n w_i X_i$.

Theorem 4.1 *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) $VaR_q(Z_v) < VaR_q(Z_w)$ if $v \prec w$ and v is not a permutation of w (in other words, the function $\psi(w, q) = VaR_q(Z_w)$ is strictly Schur-convex in $w \in \mathbf{R}^n_+$).

(ii) In particular, $VaR_q(Z_{\underline{w}}) < VaR_q(Z_w) < VaR_q(Z_{\overline{w}})$ for all $q \in (0, 1/2)$ and all weights $w \in \mathcal{I}_n$ such that $w \neq \underline{w}$ and w is not a permutation of $\overline{w}$.

Let us illustrate the settings where diversification is suboptimal in the value at risk framework. Let $q \in (0, 1)$ and let $X_1, ..., X_n$ be i.i.d. risks with a Lévy distribution $S_{1/2}(\sigma, 1, 0)$ with $\alpha = 1/2$, $\beta = 1$ and the density $f(x) = (\sigma/(2\pi))^{1/2} \exp(-\sigma/(2x))x^{-3/2}$. Using (3) with $\alpha = 1/2$ for the portfolio of $X_i's$ with equal weights $w_i = 1/n$, we get $Z_{\underline{w}} = (1/n)\sum_{i=1}^n X_i =^d nX_1$. Consequently, by positive homogeneity of the VaR (see property a3 in Section 3), $VaR_q(Z_{\underline{w}}) = nVaR_q(X_1) = nVaR_q(Z_{\overline{w}}) > VaR_q(Z_{\overline{w}})$. Thus, the VaR of the least diversified portfolios with weights $\overline{w}$ that consists of only one risk is less than the value at risk of the most diversified portfolio with equal weights $\underline{w}$.

Theorem 4.2 shows that similar conclusions hold for portfolio VaR comparisons with arbitrary weights $w = (w_1, ..., w_n) \in \mathbf{R}^n_+$ under the general assumption that the distributions of the risks $X_1, ..., X_n$ are extremely heavy-tailed. In such settings, the results in Theorem 4.1 are reversed and diversification of a portfolio of the risks X_i increases the value at risk of its return.

Theorem 4.2 *Let $q \in (0, 1/2)$ and let X_i, $i = 1, ..., n$, be i.i.d. risks such that $X_i \sim \underline{\mathcal{CS}}(1)$, $i = 1, ..., n$. Then*

(i) $VaR_q(Z_v) > VaR_q(Z_w)$ if $v \prec w$ and v is not a permutation of w (in other words, the function $\psi(w, q) = VaR_q(Z_w)$, is strictly Schur-concave in $w \in \mathbf{R}^n_+$).

(ii) In particular, $VaR_q(Z_{\overline{w}}) < VaR_q(Z_w) < VaR_q(Z_{\underline{w}})$ for all $q \in (0, 1/2)$ and all weights $w \in \mathcal{I}_n$ such that $w \neq \underline{w}$ and w is not a permutation of $\overline{w}$.

Let us consider the portfolio value at risk dealt with in Theorems 4.1 and 4.2 in the borderline case $\alpha = 1$ which corresponds to i.i.d. risks $X_1, ..., X_n$ with a symmetric Cauchy distribution $S_1(\sigma, 0, 0)$. As discussed in Section 2, these distributions are exactly at the dividing boundary between the class $\overline{\mathcal{CSLC}}$ in Theorem 4.1 and the class $\underline{\mathcal{CS}}(1)$ in Theorem 4.2. Using (3) with $\alpha = 1$ we get that, for all $w = (w_1, ..., w_n) \in \mathcal{I}_n$, $Z_w = \sum_{i=1}^n w_i X_i =^d X_1$. Consequently, for all $q \in (0, 1)$, the value at risk $VaR_q(Z_w) = VaR_q(X_1)$ is independent of w and is the same for all portfolios of risks X_i with weights $w \in \mathcal{I}_n$, $i = 1, ..., n$. Thus, in such a case, diversification of a portfolio has no effect on riskiness of its return. Similarly, for general weights $w = (w_1, ..., w_n) \in \mathbf{R}^n_+$, property (3) with $\alpha = 1$ implies $Z_w = \sum_{i=1}^n w_i X_i =^d (\sum_{i=1}^n w_i) X_1$. Thus, the value at risk $VaR_q(Z_w) = (\sum_{i=1}^n w_i) VaR_q(X_1)$ is independent of w so long as $\sum_{i=1}^n w_i$ is fixed. Consequently, $VaR_q(Z_w)$ is both Schur-convex (as in Theorem 4.1) and Schur-concave (as in Theorem 4.2) in $w \in \mathbf{R}^n_+$ for i.i.d. risks $X_i \sim S_\alpha(\sigma, 0, 0)$ that have symmetric Cauchy distributions with $\alpha = 1$ (see Proschan (1965) and Marshall & Olkin (1979), p. 374, for similar properties of tail probabilities of Cauchy distributions).[16]

Remark 4.1 Theorems A3.1 and A3.2 in Appendix A3 provide extensions of Theorems 4.1 and 4.2 in the case of skewed and heterogeneous risks. Theorems A3.3-A3.5 in the same appendix contain the results on sharp VaR bounds for portfolios of possibly skewed and heterogenous heavy-tailed risks. In particular, Theorem A3.5 provides the following extensions of Theorems 4.1 and 4.2 to the classes $\overline{\mathcal{CS}}(r)$ and $\underline{\mathcal{CS}}(r)$. Let $q \in (0, 1/2)$, $r \in (0, 2)$ and let $X_1, ..., X_n$ be i.i.d. risks such that $X_i \sim \overline{\mathcal{CS}}(r)$. Then $VaR_q(Z_v) < VaR_q(Z_w)$ if $(v_1^r, ..., v_n^r) \prec (w_1^r, ..., w_n^r)$ and $(v_1^r, ..., v_n^r)$ is not a permutation of $(w_1^r, ..., w_n^r)$ (that is, the function $\psi(w, q) = VaR_q(Z_w)$, $w \in \mathbf{R}^n_+$, is strictly Schur-convex in $(w_1^r, ..., w_n^r)$). In addition, the following sharp bounds hold: $n^{1-1/r}(\sum_{i=1}^n w_i^r)^{1/r} VaR_q(Z_{\underline{w}}) < VaR_q(Z_w) < (\sum_{i=1}^n w_i^r)^{1/r} VaR_q(Z_{\overline{w}})$ for all weights $w \in \mathcal{I}_n$ such that $w \neq \underline{w}$ and w is not a permutation of $\overline{w}$. The above inequalities for portfolio VaR are reversed for i.i.d. risks $X_i \sim \underline{\mathcal{CS}}(r)$ (so that the function $\psi(w, q) = VaR_q(Z_w)$, $w \in \mathbf{R}^n_+$, is strictly Schur-concave in $(w_1^r, ..., w_n^r)$ for such risks). Similar to the discussion following Theorem 4.2, it is also easy to see that these inequalities hold as equalities for i.i.d. risks $X_i \sim S_r(\sigma, 0, 0)$ with symmetric stable distributions with the index of stability $\alpha = r$ that are at the dividing boundary between the classes $\overline{\mathcal{CS}}(r)$ and $\underline{\mathcal{CS}}(r)$ (see the discussion in Section 2).

Remark 4.2 Proposition A1.5 in Appendix A1 implies that value at risk comparisons for risks with symmetric unimodal densities are closed under convolutions. In particular, if X_1, X_2 and Y_1, Y_2 are independent risks with symmetric unimodal densities such that, for $i = 1, 2$, and all $q \in (0, 1/2)$, $VaR_q(X_i) < VaR_q(Y_i)$, then $VaR_q(X_1 + X_2) < VaR_q(Y_1 + Y_2)$ for all $q \in (0, 1/2)$. The densities of the returns Z_w on the portfolios of risks in the classes $\overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}}(r)$, $r \in (0, 2]$, and $\underline{\mathcal{CS}}(r)$, $r \in (0, 2]$, are symmetric and unimodal by the results summarized in Appendix A1. The above VaR and unimodality properties provide one of the key arguments in the proof of Theorems 4.1 and 4.2.

[16] From the proof of Theorems 4.1 and 4.2 and this property it follows that the theorems continue to hold for convolutions of distributions from the classes $\overline{\mathcal{CSLC}}$ and $\underline{\mathcal{CS}}(1)$ with Cauchy distributions $S_1(\sigma, 0, 0)$.

Remark 4.3 From Theorem 4.1 it follows that, if X_1 and X_2 are i.i.d. risks such that $X_i \sim \overline{\mathcal{CSLC}}$, $i = 1, 2$, then $VaR_q(X_1 + X_2) < VaR_q(X_1) + VaR_q(X_2)$ and $VaR_q(\lambda X_1 + (1-\lambda)X_2) < \lambda VaR_q(X_1) + (1-\lambda)VaR_q(X_2)$ for all $q \in (0, 1/2)$ and any $\lambda \in (0, 1)$. That is, the value at risk exhibits subadditivity and convexity is thus a coherent measure of risk for the class $\overline{\mathcal{CSLC}}$ (see Section 3 for the definition of coherent risk measures and coherency axioms in the case of the VaR). On the other hand, Theorem 4.2 implies that $VaR_q(X_1) + VaR_q(X_2) < VaR_q(X_1 + X_2)$ and $\lambda VaR_q(X_1) + (1-\lambda)VaR_q(X_2) < VaR_q(\lambda X_1 + (1-\lambda)X_2)$ for all $q \in (0, 1/2)$, $\lambda \in (0, 1)$ and i.i.d. risks $X_1, X_2 \sim \underline{\mathcal{CS}}(1)$. Consequently, subadditivity and convexity are always violated for risks with extremely heavy-tailed distributions. In such a case, the value at risk is not a coherent risk measure even in the case of independence which is "the worst case scenario" for diversification failure.

Remark 4.4 From the counterexamples constructed in Artzner, Delbaen, Eber & Heath (1999) and Embrechts, McNeil & Straumann (2002) it follows that the value at risk, in general, fails to satisfy the subadditivity and convexity properties. From the analysis similar to Examples 6 and 7 in Embrechts, McNeil & Straumann (2002) and Chapter 12 in Bouchaud & Potters (2004) it follows that subadditivity of the VaR holds for distributions with power-law tails (1) and sufficiently small values of the loss probability q if $\alpha > 1$. Subadditivity is violated for power-law distributions (1) and sufficiently small values of the loss probability q if $\alpha < 1$. More generally, let, similar to Example 7 in Embrechts, McNeil & Straumann (2002) and Section 12.1.2 in Bouchaud & Potters (2004), X_1 and X_2 be two i.i.d. risks with regularly varying heavy tails: $P(X_1 > x) = L(x)/x^\alpha$, where $\alpha > 0$ and $L(x)$ is a slowly varying at infinity function, that is $L(\lambda x)/L(x) \to 1$, as $x \to +\infty$, for all $\lambda > 0$ (see Embrechts, Klüppelberg & Mikosch (1997) and Zolotarev (1986), p. 8). Using the property that $\lim_{x\to+\infty} P(X_1 + X_2 > x)/P(X_1 > x/2^{1/\alpha}) = 1$ (see Lemma 1.3.1 in Embrechts, Klüppelberg & Mikosch (1997) and Section 12.1.2 in Bouchaud & Potters (2004)), one gets that $\lim_{q\to 0} VaR_q(X_1 + X_2)/(VaR_q(X_1) + VaR_q(X_2)) = 2^{1/\alpha - 1}$. Consequently, the subadditivity property holds for the value at risk *asymptotically* as $q \to 0$ if $\alpha > 1$ and is violated as $q \to 0$ if $\alpha < 1$. The implications of Theorems 4.1 and 4.2 for the value at risk coherency in Remark 4.3 are qualitatively different from the counterexamples available in the literature and the above asymptotic considerations. This is because the VaR comparisons in Remark 4.3 hold *regardless* of the value of q and are valid for the *whole* wide classes of heavy-tailed risks. From the results in Section 5 it follows that similar value at risk comparisons and conclusions also hold under dependence.

Remark 4.5 It is well-known that if r.v.'s X and Y are such that $P(X > x) \leq P(Y > x)$ for all $x \in \mathbf{R}$, then $EU(X) \leq EU(Y)$ for all increasing functions $U : \mathbf{R} \to \mathbf{R}$ for which the expectations exist (see Shaked & Shanthikumar (2007), pp. 3-4). In addition, for risks

Z_v and Z_w with symmetric distributions, $VaR_q(Z_v) \le VaR_q(Z_w)$ for all $q \in (0, 1/2)$ if and only if $P(|Z_v| > x) \le P(|Z_w| > x)$ for all $x > 0$.[17] These properties, together with Theorems 4.1 and 4.2 and Remark 4.1 imply majorization comparisons for expectations of (risk measure) functions of linear combinations of heavy-tailed r.v.'s. For instance, we get that if $U : \mathbf{R}_+ \to \mathbf{R}$ is an increasing function, then, assuming existence of the expectations, the function $\varphi(w) = EU(|\sum_{i=1}^n w_i X_i|)$, $w \in \mathbf{R}_+^n$, is Schur-convex in $(w_1^r, ..., w_n^r)$ for i.i.d. risks $X_i \sim \overline{\mathcal{CS}}(r)$ and is Schur-concave in $(w_1^r, ..., w_n^r)$ for i.i.d. risks $X_i \sim \underline{\mathcal{CS}}(r)$. In particular, $EU\big(\big|n^{1-1/r}\big(\sum_{i=1}^n w_i^r\big)^{1/r} Z_{\underline{w}}\big|\big) \le EU(|Z_w|) \le EU\big(\big|\big(\sum_{i=1}^n w_i^r\big)^{1/r} Z_{\overline{w}}\big|\big)$ for i.i.d. risks $X_i \sim \overline{\mathcal{CS}}(r)$ and $EU\big(\big|\big(\sum_{i=1}^n w_i^r\big)^{1/r} Z_{\overline{w}}\big|\big) \le EU(|Z_w|) \le EU\big(\big|n^{1-1/r}\big(\sum_{i=1}^n w_i^r\big)^{1/r} Z_{\underline{w}}\big|\big)$ for i.i.d. risks $X_i \sim \underline{\mathcal{CS}}(r)$. By Remark 4.1 and the results in Appendix A3 we also get that the function $\varphi(w)$, $w \in \mathbf{R}_+^n$ is Schur-concave in $(w_1^2, ..., w_n^2)$ if $X_1, ..., X_n$ are i.i.d. risks such that $X_i \sim \underline{\mathcal{CS}}(2)$ or $X_i \sim S_\alpha(\sigma, \beta, 0)$ for some $\sigma > 0$, $\beta \in [-1, 1]$ and $\alpha \in (0, 2)$. The above results extend and complement those obtained by Efron (1969) and Eaton (1970) (see also Marshall & Olkin (1979), pp. 361-365) who studied the classes of functions $U : \mathbf{R} \to \mathbf{R}$ and r.v.'s $X_1, ..., X_n$ for which Schur-concavity of $\varphi(w)$, $w \in \mathbf{R}_+^n$, in $(w_1^2, ..., w_n^2)$ holds. Further, we obtain that $\varphi(w)$ is Schur-convex in $w \in \mathbf{R}_+^n$ under the assumptions of Theorem 4.1 and is Schur-concave in $w \in \mathbf{R}_+^n$ under the assumptions of Theorem 4.2. It is important to note here that in the case of increasing *convex* functions $U : \mathbf{R}_+ \to \mathbf{R}$ and r.v.'s $X_1, ..., X_n$ satisfying the assumptions of Theorem 4.2, the expectations $EU(|\sum_{i=1}^n w_i X_i|)$ are infinite for all $w \in \mathbf{R}_+^n$ (since the function $(f(x) - f(0))/x$ is increasing in $x > 0$ by Marshall & Olkin (1979), p. 453). Therefore, the last result does not contradict the well-known fact that (see Marshall & Olkin (1979), p. 361) the function $Ef(\sum_{i=1}^n w_i X_i)$ is Schur-convex in $(w_1, ..., w_n) \in \mathbf{R}$ for all i.i.d. r.v.'s $X_1, ..., X_n$ and convex functions $f : \mathbf{R} \to \mathbf{R}$ as it might seem on the first sight.

Remark 4.6 Theorems 4.1 and 4.2 and Remark 4.1 imply corresponding results on majorization properties of the tail probabilities $\xi(w, x) = P(\sum_{i=1}^n w_i X_i > x)$, $x > 0$, of linear combinations of heavy-tailed r.v.'s $X_1, ..., X_n$. These implications generalize the results in the seminal work by Proschan (1965) who showed that the tail probabilities $\xi(w, x)$ are Schur-convex in $w = (w_1, ..., w_n) \in \mathbf{R}_+^n$ for all $x > 0$ for i.i.d. r.v.'s $X_i \sim \mathcal{LC}$, $i = 1, ..., n$.[18] Schur-convexity of $\xi(w, x)$ for $X_i \sim \mathcal{LC}$ implies that the value at risk comparisons in Theorem 4.1 hold for i.i.d. log-concavely distributed risks, as stated in Proposition A1.6. The results in Proschan (1965) have been applied to the analysis of many problems in statistics, econometrics, economic theory,

[17] Taking the absolute values here and inside the function U in the comparisons that follow is needed because, for Z_v and Z_w with symmetric distributions, the value at risk comparisons $VaR_q(Z_v) \le VaR_q(Z_w)$, $q \in (0, 1/2)$, imply the opposite inequalities for the tail probabilities with negative $x : P(Z_v > x) \ge P(Z_w > x)$ for $x < 0$.

[18] The main results in Proschan (1965) are reviewed in Section 12.J in Marshall & Olkin (1979). The work by Proschan (1965) is also presented, in a rearranged form, in Section 11 of Chapter 7 in Karlin (1968). Peakedness results in Karlin (1968), Proschan (1965) are formulated for "PF2 densities," which is the same as "log-concave densities."

mathematical evolutionary theory and other fields (see the discussion in Ibragimov (2005)). One should note here that applicability of these majorization results and their analogs for other classes of distributions to portfolio value at risk theory has not been recognized in the previous literature even in the case of i.i.d. log-concavely distributed risks.

A number of papers in probability and statistics have focused on extension of Proschan's results (see, among others, Chan, Park & Proschan (1989), Jensen (1997), Ma (1998) and the review in Tong (1994)). However, in all the studies that dealt with generalizations of the results, the majorization properties of the tail probabilities were of the same type as in Proschan (1965). Namely, the results gave extensions of Proschan's results concerning *Schur-convexity* of the tail probabilities $\xi(a, x)$, $x > 0$, to classes of r.v.'s more general than those considered in Proschan (1965). Analogues of Theorems 4.1 and 4.2 and Remark 4.1 for the tail probabilities $\xi(a, x)$, on the other hand, provide the first general results concerning *Schur-concavity* of $\xi(a, x)$, $x > 0$, for certain wide classes of r.v.'s. According to these results, the class of distributions for which Schur-convexity of the tail probabilities $\xi(a, x)$ is replaced by their Schur-concavity is precisely the class of distributions with extremely heavy-tailed densities.

5 Generalizations to dependence

In this section, we show that the results obtained in the paper continue to hold for dependent risks. More precisely, the results continue to hold for convolutions of risks with joint $\alpha-$symmetric and spherical distributions as well as for several classes of models with common shocks.

According to the definition introduced in Cambanis, Keener & Simons (1983), an $n-$dimensional distribution is called $\alpha-$symmetric if its c.f. can be written as $\phi((\sum_{i=1}^n |t_i|^\alpha)^{1/\alpha})$, where $\phi : \mathbf{R}_+ \to \mathbf{R}$ is a continuous function (with $\phi(0) = 1$) and $\alpha > 0$. That is, a vector $(X_1, ..., X_n)$ has an $\alpha-$symmetric distribution if, for all $t_i \in \mathbf{R}$,

$$E \exp(\sum_{i=1}^n it_iX_i) = \phi((\sum_{i=1}^n |t_i|^\alpha)^{1/\alpha}) \tag{7}$$

with a function ϕ that satisfies the above properties. An important fact is that, similar to strictly stable laws, relation (3) holds if $(X_1, ..., X_n)$ has an $\alpha-$symmetric distribution, (see Fang, Kotz & Ng (1990), Ch. 7). The number α is called the index and the function ϕ is called the c.f. generator of the $\alpha-$symmetric distribution. The class of $\alpha-$symmetric distributions contains, as a subclass, spherical (also referred to as spherically symmetric) 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 (Fang, Kotz & Ng (1990), Ch. 3). Spherically symmetric stable distributions have characteristic functions $\exp\left[-\lambda\left(\sum_{i=1}^n t_i^2\right)^{\gamma/2}\right]$, $0 < \gamma \leq 2$, and are, thus, examples of $\alpha-$symmetric distributions with $\alpha = 2$ and the c.f. generator $\phi(x) = \exp(-x^\gamma)$.

For any $0 < \alpha \leq 2$, the class of $\alpha-$symmetric distributions includes distributions of risks $Q_1, ..., Q_n$ that have the common factor representation

$$(Q_1, ..., Q_n) = (ZY_1, ..., ZY_n), \tag{8}$$

where $Y_i \sim S_\alpha(\sigma, 0, 0)$, $\sigma > 0$, are i.i.d. stable r.v.'s and $Z \geq 0$ is a nonnegative r.v. independent of $Y_i's$ (see Bretagnolle, Dacunha-Castelle & Krivine (1966) and Fang, Kotz & Ng (1990), p. 197). In the case $Z = 1$ (a.s.), model (8) represents vectors with i.i.d. symmetric stable components that have c.f.'s $\exp\left[-\lambda\sum_{i=1}^n |t_i|^\alpha\right]$ which are particular cases of c.f.'s of $\alpha-$symmetric distributions with the generator $\phi(x) = \exp(-\lambda x^\alpha)$.

According to the results in Bretagnolle, Dacunha-Castelle & Krivine (1966) and Kuritsyn & Shestakov (1984), the function $\exp\left(-\left(|t_1|^\alpha + |t_2|^\alpha\right)^{1/\alpha}\right)$ is a c.f. of two $\alpha-$symmetric r.v.'s for all $\alpha \geq 1$ (the generator of the function is $\phi(u) = \exp(-u)$). Zastavnyi (1993) demonstrates that the class of more than two $\alpha-$symmetric r.v.'s with $\alpha > 2$ consists of degenerate variables (so that their c.f. generator $\phi(u) = 1$). For further review of properties and examples of $\alpha-$symmetric distributions the reader is referred to Ch. 7 in Fang, Kotz & Ng (1990) and Gneiting (1998).

The dependence structures considered in this section include vectors $(X_1, ..., X_n)$ given by sums of independent random vectors $(Y_{1j}, ..., Y_{nj})$, $j = 1, ..., k$, where $(Y_{1j}, ..., Y_{nj})$ has an absolutely continuous $\alpha-$symmetric distribution with an index α_j:

$$(X_1, ..., X_n) = \sum_{j=1}^k (Y_{1j}, ..., Y_{nj}). \tag{9}$$

In particular, this framework includes sums of random vectors $(Z_jY_{1j}, ..., Z_jY_{nj})$, $j = 1, ..., k$, in (8) with independent r.v.'s Z_j, Y_{ij}, $j = 1, ..., k$, $i = 1, ..., n$, such that Z_j are positive and absolutely continuous and $Y_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $\alpha_j \in (0, 2]$, $\sigma_j > 0$:

$$(X_1, ..., X_n) = \sum_{j=1}^k (Z_jY_{1j}, ..., Z_jY_{nj}). \tag{10}$$

Although the dependence structure in model (8) alone is restrictive, convolutions (10) of such vectors provide a natural framework for modeling of random environments with different multiple common shocks Z_j, such as macroeconomic or political ones, that affect all risks X_i (see Andrews (2005)).

Convolutions of $\alpha-$symmetric distributions exhibit both heavy-tailedness in marginals and dependence among them. It is not difficult to show that convolutions of $\alpha-$symmetric distributions with $\alpha < 1$ have extremely heavy-tailed marginals with infinite means.[19] On the other hand, convolutions of $\alpha-$symmetric distributions with $1 < \alpha \leq 2$ can have marginals with power moments finite up to a certain positive order (or finite exponential moments). In particular, finiteness of moments for convolutions of models (8) with $1 < \alpha \leq 2$ depends on heavy-tailedness of the common shock variables Z. For instance, convolutions of models (8) with $1 < \alpha < 2$ and $E|Z| < \infty$ have finite means but infinite variances, however, marginals of such convolutions have infinite means if the r.v.'s Z satisfy $E|Z| = \infty$. Moments $E|ZY_i|^p$, $p > 0$, of marginals in models (8) with $\alpha = 2$ (that correspond to Gaussian r.v.'s Y_i) are finite if and only if $E|Z|^p < \infty$. In particular, all marginal power moments in models (8) with $\alpha = 2$ are finite if $E|Z|^p < \infty$ for all $p > 0$. Similarly, marginals of spherical (that is, 2-symmetric) distributions range from extremely heavy-tailed to extreme lighted-tailed ones. For example, marginal moments of spherically symmetric $\alpha-$stable distributions with c.f.'s $\exp\big[-\lambda\big(\sum_{i=1}^n t_i^2\big)^{\gamma/2}\big]$, $0 < \gamma < 2$, are finite if and only if their order is less than γ. Marginal moments of a multivariate $t-$distribution with k degrees of freedom which is a an example of a spherical distribution are finite if and only the order of the moments is less than k. These distributions were used in a number of works to model heavy-tailedness phenomena with moments up to some order (see, among others, Praetz (1972), Blattberg & Gonedes (1974) and Glasserman, Heidelberger & Shahabuddin (2002)).

Theorems 5.1 and 5.2 show that the results presented in Section 4 for i.i.d. risks continue to hold for convolutions (9) and (10) of $\alpha-$symmetric distributions and models with common shocks.

Let Φ denote the class of c.f. generators ϕ such that $\phi(0) = 1$, $\lim_{t\to\infty}\phi(t) = 0$, and the function $\phi'(t)$ is concave.

Theorem 5.1 *Theorem 4.1 continues to hold if any of the following is satisfied:*

the random vector $(X_1, ..., X_n)$ *is given by (9), where, for* $j = 1, ..., k$, $(Y_{1j}, ..., Y_{nj})$ *has an absolutely continuous* $\alpha-$*symmetric distribution with the c.f. generator* $\phi_j \in \Phi$ *and the index* $\alpha_j \in (1, 2]$. *In particular, Theorem 4.1 holds if* $(X_1, ..., X_n)$ *is given by (9), where* $(Y_{1j}, ..., Y_{nj})$, $j = 1, ..., k$, *have absolutely continuous spherical distributions with c.f. generators* $\phi_j \in \Phi$ *(the case* $\alpha_j = 2$ *for all* j*).*

[19] This is true because if one assumes that r.v.'s $X_1, ..., X_n$, $n \geq 2$, have an $\alpha-$symmetric distribution with $\alpha < 1$ and that $E|X_i| < \infty$, $i = 1, ..., n$, then, by the triangle inequality, $E|X_1+...+X_n| \leq E|X_1|+...+E|X_n| = nE|X_1|$. This inequality, however, cannot hold since, according to (3), $(X_1 + ... + X_n) \sim n^{1/\alpha}X_1$ and thus, under the above assumptions, $E|X_1 + ... + X_n| > nE|X_1|$. Similarly, one can show that $\alpha-$symmetric distributions with $\alpha < r$ have infinite marginal moments of order r.

the random vector $(X_1, ..., X_n)$ is given by (10), where $\alpha_j \in (1, 2]$, $j = 1, ..., k$.

Theorem 5.2 *Theorem 4.2 continues to hold if any of the following is satisfied:*

the random vector $(X_1, ..., X_n)$ is given by (9), where, for $j = 1, ..., k$, $(Y_{1j}, ..., Y_{nj})$ has an absolutely continuous $\alpha-$symmetric distribution with the c.f. generator $\phi_j \in \Phi$ and the index $\alpha_j \in (0, 1)$;

the random vector $(X_1, ..., X_n)$ is given by (10), where $\alpha_j \in (0, 1)$, $j = 1, ..., k$.

Remark 5.1 Similar to the proof of Theorems 5.1 and 5.2, one can also obtain similar extensions of the sharp VaR bounds in Remark 4.1 and Theorem A3.5 for heavy-tailed dependent risks.

6 Concluding remarks

In this paper, we discussed how the majorization pre-ordering and the results in majorization theory can be used in the analysis of portfolio value at risk and diversification. The results in the paper demonstrate that, in the framework of portfolio VaR comparisons, diversification may be inferior under extreme heavy-tailedness in the portfolio components. The paper further shows that portfolio diversification is optimal if the risks in consideration are moderately heavy-tailed.

As shown in Section 5 and Appendix A3, these conclusions hold under dependence and skewness and for not necessarily identically distributed risks. At the same time, dependence and heterogeneity may naturally require altering of formalizations of diversification considered in this work. For instance, the results for non-identically distributed risks presented in Appendix A3 suggest that analogs of Cheng's $p-$majorization (see Ch. 14 in Marshall & Olkin (1979)) may be useful in formalization of portfolio diversification for heterogenous risks.

By definition, the classes $\overline{\mathcal{CS}}(r)$ in Remark 4.1 and Theorem A3.5 include i.i.d. normal distributions which are log-concave, as indicated in Section 2. It is not known, in general, whether the results in that remark and theorem with $r \neq 1$ also hold for a sufficiently wide subclass of i.i.d. log-concave distributions. If this is the case, then the analysis in this paper would allow one to extend the results in Theorem A3.5 to convolutions of distributions in $\overline{\mathcal{CS}}(r)$ and in this subclass. Some results related to this problem with $r = 2$ are provided by the properties of Schur-concavity of the expectations $\varphi(w) = EU(|\sum_{i=1}^n w_i X_i|)$ in $(w_1^2, ..., w_n^2)$ for certain classes of functions U and risks $X_1, ..., X_n$ considered in Remark 4.5.

Analysis of the above issues and extensions of the results in the paper to distributions beyond those considered in Section 5 and and Appendix A3 are left for further research.

Appendix A1: VaR and unimodality properties of log-concave and stable distributions

This appendix summarizes auxiliary VaR and unimodality results for log-concave and stable distributions needed for the analysis in the paper.

Definition A1.1 *A density function f is said to be unimodal if there exists $c \in \mathbf{R}$ such that f is nondecreasing on $(-\infty, c)$ and is nonincreasing on (c, ∞). An absolutely continuous r.v. or distribution with density f is said to be unimodal if f is unimodal.*

Proposition A1.1 (Theorem 2.7.6 in Zolotarev (1986), p. 134) *Each stable r.v. $X \sim S_\alpha(\sigma, \beta, \mu)$ is unimodal.*

Proposition A1.2 is a consequence of Theorem 1.10 in Dharmadhikari & Joag-Dev (1988), p. 20 (see also Dharmadhikari & Joag-Dev (1988), p. 18, and An (1998)).

Proposition A1.2 (Dharmadhikari & Joag-Dev (1988), An (1998)) *Any log-concave density is unimodal.*

Proposition A1.3 below easily follows from a result due to R. Askey, see Theorem 4.1 in Gneiting (1998) and the proof of the proposition in Appendix A2.

Proposition A1.3 *If $(X_1, ..., X_n)$ has an absolutely continuous $\alpha-$symmetric distribution with the c.f. generator $\phi \in \Phi$, then the density of the r.v. $Z_w = \sum_{i=1}^n w_i X_i$ is symmetric and unimodal for all $w_i \in \mathbf{R}$.*

Proposition A1.4 (Theorem 1.6 in Dharmadhikari & Joag-Dev (1988), p. 13) *The convolution of two symmetric unimodal densities is unimodal.*

Proposition A1.5 is an analogue of Lemma and Theorem 1 in Birnbaum (1948) and Theorem 3.D.4 on p. 173 in Shaked & Shanthikumar (2007) in terms of the value at risk inequalities and strict VaR comparisons. Its proof in Appendix A2 follows the same lines as in Birnbaum (1948).

Proposition A1.5 *Let $X_1, ..., X_n$ and $Y_1, ..., Y_n$ be two sets of independent r.v.'s, all having symmetric unimodal densities. Suppose that $VaR_q(X_i) < VaR_q(Y_i)$, $i = 1, ..., n$, for all $q \in (0, 1/2)$. Then $VaR_q(\sum_{i=1}^n X_i) < VaR_q(\sum_{i=1}^n Y_i)$ for all $q \in (0, 1/2)$.*

Proposition A1.6 follows from majorizations results for tail probabilities of linear combinations of log-concavely distributed r.v.'s in Proschan (1965), see the arguments in Appendix A2.

Proposition A1.6 *Theorem 4.1 holds for i.i.d. risks* $X_1, ..., X_n$ *such that* $X_i \sim \mathcal{LC}$, $i = 1, ..., n$.

Appendix A2: Proofs

Proof of Theorems 4.1 and 4.2. Let $\alpha \in (0, 2]$, $\sigma > 0$, and let $v = (v_1, ..., v_n) \in \mathbf{R}^n_+$ and $w = (w_1, ..., w_n) \in \mathbf{R}^n_+$ be two vectors of portfolio weights such that $(v_1, ..., v_n) \prec (w_1, ..., w_n)$ and $(v_1, ..., v_n)$ is not a permutation of $(w_1, ..., w_n)$ (clearly, $\sum_{i=1}^n v_i \neq 0$ and $\sum_{i=1}^n w_i \neq 0$). Let $X_1, ..., X_n$ be i.i.d. risks such that $X_i \sim S_\alpha(\sigma, 0, 0)$, $i = 1, ..., n$. From (3) it follows that if $c = (c_1, ..., c_n) \in \mathbf{R}^n_+$, $\sum_{i=1}^n c_i \neq 0$, then $Z_c = \sum_{i=1}^n c_i X_i =^d \left(\sum_{i=1}^n c_i^\alpha\right)^{1/\alpha} X_1$. Using positive homogeneity of the value at risk (property a3 in Section 3), we thus obtain that, for all $q \in (0, 1/2)$,

$$VaR_q(Z_c) = VaR_q(X_1)\Big(\sum_{i=1}^n c_i^\alpha\Big)^{1/\alpha}. \tag{11}$$

Proposition 3.C.1.a in Marshall & Olkin (1979) implies that the function $h(c_1, ..., c_n) = \sum_{i=1}^n c_i^\alpha$ is strictly Schur-convex in $(c_1, ..., c_n) \in \mathbf{R}^n_+$ if $\alpha > 1$ and is strictly Schur-concave in $(c_1, ..., c_n) \in \mathbf{R}^n_+$ if $\alpha < 1$. Therefore, we have $\sum_{i=1}^n v_i^\alpha < \sum_{i=1}^n w_i^\alpha$, if $\alpha > 1$ and $\sum_{i=1}^n w_i^\alpha < \sum_{i=1}^n v_i^\alpha$, if $\alpha < 1$. This, together with (11) implies that, for all $q \in (0, 1/2)$,

$$VaR_q(Z_v) < VaR_q(Z_w) \tag{12}$$

if $\alpha > 1$, and

$$VaR_q(Z_v) > VaR_q(Z_w) \tag{13}$$

if $\alpha < 1$. This completes the proof of parts (i) of Theorem 4.1 and 4.2 in the case of i.i.d. stable risks $X_i \sim S_\alpha(\sigma, 0, 0)$, $i = 1, ..., n$.

Let now $X_1, ..., X_n$ be i.i.d. risks such that $X_i \sim \overline{\mathcal{CSLC}}$, $i = 1, ..., n$. By definition, $X_i = \gamma Y_{i0} + \sum_{j=1}^k Y_{ij}$, $i = 1, ..., n$, where $\gamma \in \{0, 1\}$, $k \geq 0$, $Y_{i0} \sim \mathcal{LC}$, $i = 1, ..., n$, and $(Y_{1j}, ..., Y_{nj})$, $j = 0, 1, ..., k$, are independent vectors with i.i.d. components such that $Y_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $\alpha_j \in (1, 2]$, $\sigma_j > 0$, $i = 1, ..., n$, $j = 1, ..., k$. From (12) and the results in Proschan (1965) for tail probabilities of log-concavely distributed r.v.'s (see Proposition A1.6 and Remark 4.6) it follows that, for all $q \in (0, 1/2)$ and all $j = 0, 1, ..., k$, $VaR_q\big(\sum_{i=1}^n v_i Y_{ij}\big) < VaR_q\big(\sum_{i=1}^n w_i Y_{ij}\big)$. The densities of the r.v.'s Y_{i0}, $i = 1, ..., n$, are symmetric and unimodal by Proposition A1.2. In addition, the densities of the r.v.'s Y_{ij}, $i = 1, ..., n$, $j = 1, ..., k$, are symmetric and unimodal by Proposition A1.1. Using Proposition A1.4, we conclude that the densities of the r.v.'s $\sum_{i=1}^{n} v_i Y_{ij}$ and $\sum_{i=1}^{n} w_i Y_{ij}$, $j = 0, 1, ..., k$, are symmetric and unimodal as well. By Proposition A1.5 we thus obtain

$$VaR_q(Z_v) = VaR_q\big(\sum_{i=1}^{n} v_i X_i\big) = VaR_q\big(\gamma \sum_{i=1}^{n} v_i Y_{i0} + \sum_{j=1}^{k}\sum_{i=1}^{n} v_i Y_{ij}\big) <$$

$$VaR_q\big(\gamma \sum_{i=1}^{n} w_i Y_{i0} + \sum_{j=0}^{k}\sum_{i=1}^{n} w_i Y_{ij}\big) = VaR_q\big(\sum_{i=1}^{n} w_i X_i\big) = VaR_q(Z_w).$$

This completes the proof of part (i) of Theorem 4.1. Part (i) of Theorem 4.2 may be proven in a completely similar way, with the use of relations (13) instead of Proposition A1.6 and inequalities (12). The bounds in parts (ii) of Theorems 4.1 and 4.2 follow from their parts (i) and majorization comparisons (4). Sharpness of the bounds in parts (ii) of the theorems follows from the property that, as follows from the discussion following Theorem 4.2, the bounds become equalities in the limit as $\alpha \to 1$ for i.i.d. risks $X_i \sim S_1(\sigma, 0, 0)$, $i = 1, ..., n$, with symmetric Cauchy distributions. ■

Proof of Theorems 5.1 and 5.2. The proof of the extensions of Theorems 4.1 and 4.2 to the dependent case follows the same lines as the proof of the above theorems since the following properties hold:

Relation (3) holds if $(X_1, ..., X_n)$ has an $\alpha-$symmetric distribution (see Fang, Kotz & Ng (1990), Ch. 7, and Section 5 in this paper);

The densities of the r.v.'s $\sum_{i=1}^{n} v_i Y_{ij}$, $j = 1, ..., k$, and $\sum_{i=1}^{n} w_i Y_{ij}$, $j = 1, ..., k$, are symmetric and unimodal by Proposition A1.3.

The densities of the r.v.'s $Z_j \sum_{i=1}^{n} v_i Y_{ij}$, $j = 1, ..., k$, and $Z_j \sum_{i=1}^{n} w_i Y_{ij}$, $j = 1, ..., k$, are symmetric and unimodal if $Y_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $i = 1, ..., n$, $j = 1, ..., k$, and Z_j are absolutely continuous positive r.v.'s independent of Y_{ij} (this follows by symmetry and unimodality of $\sum_{i=1}^{n} v_i Y_{ij}$ and $\sum_{i=1}^{n} w_i Y_{ij}$ implied by Propositions A1.1 and A1.4, Definition A1.1 and conditioning arguments). ■

Proof of Proposition A1.3. By (7), the c.f. of the r.v. $Z_w = \sum_{i=1}^{n} w_i X_i$ satisfies $E\exp(itZ_w) = f(|t|)$, $t \in \mathbf{R}$, where $f(s) = \phi\big(h(w)s\big)$, $s \in \mathbf{R}_+$, and $h(w) = \sum_{i=1}^{n} |w_i|^{\alpha}$. Since the c.f. is real, the density of Z_w is symmetric. From Theorem 4.1 in Gneiting (1998), $f(|t|)$, $t \in \mathbf{R}$, is a c.f. of a unimodal distribution if the function $f : \mathbf{R}_+ \to \mathbf{R}$ is in Φ : $f \in \Phi$. Evidently, this holds since $\phi \in \Phi$. ■

Proof of Proposition A1.5. It suffices to provide the arguments in the case $n = 2$; the case of general n then follows by induction and Proposition A1.4. Let X_1, X_2 and Y_1, Y_2 be two sets of independent r.v.'s with symmetric unimodal densities f_1, f_2 and g_1, g_2. Suppose that $VaR_q(X_i) < VaR_q(Y_i)$, $i = 1, 2$, for all $q \in (0, 1/2)$ or, equivalently,

$$P(X_i > s) < P(Y_i > s), \quad i = 1, 2, \quad \text{for all} \quad s > 0. \tag{14}$$

As in the proof of Lemma in Birnbaum (1948), for all $x > 0$, we have

$$P(Y_1 + Y_2 > x) - P(X_1 + X_2 > x) = I_1(x) + I_2(x), \tag{15}$$

where

$$I_1(x) = \int_0^{+\infty} [P(X_1 > s) - P(Y_1 > s)][f_2(x + s) - f_2(x - s)]ds,$$
$$I_2(x) = \int_0^{+\infty} [P(X_2 > s) - P(Y_2 > s)][g_1(x + s) - g_1(x - s)]ds.$$

As shown in the proof of Lemma Birnbaum (1948), from symmetry and unimodality of f_2 and g_1 it follows that $f_2(x+s) - f_2(x-s) \leq 0$ and $g_1(x+s) - g_1(x-s) \leq 0$ for all $x \geq 0$ and $s \geq 0$. This, together with (14) and (15), imply that $P(Y_1 + Y_2 > x) - P(X_1 + X_2 > x) \geq 0$ for all $x > 0$. In addition, $P(Y_1 + Y_2 > x) = P(X_1 + X_2 > x)$ for some $x > 0$ if and only $f_2(x+s) = f_2(x-s)$ and $g_1(x+s) = g_1(x-s)$ for all $s \geq 0$, which is impossible. Thus, $P(Y_1+Y_2 > x) - P(X_1+X_2 > x) > 0$ for all $x > 0$, or, equivalently, $VaR_q(X_1 + X_2) < VaR_q(Y_1 + Y_2)$ for all $q \in (0, 1/2)$. ■

Proof of Proposition A1.6. Let $X_1, ..., X_n$ be i.i.d. risks such such that $X_i \sim \mathcal{LC}$, $i = 1, ..., n$. From Theorem 2.3 in Proschan (1965) it follows that, for any $x > 0$, the function $\xi(w, x) = P(\sum_{i=1}^n w_i X_i > x)$ is strictly Schur-convex in $w = (w_1, ..., w_n) \in \mathbf{R}_+^n$. This implies that if $v \prec w$ and v is not a permutation of w, then, for all $q \in (0, 1/2)$, $q = P(Z_v > VaR_q(Z_v)) < P(Z_w > VaR_q(Z_v))$. Consequently, $VaR_q(Z_v) < VaR_q(Z_w)$ for all $q \in (0, 1/2)$, and thus part (i) of Theorem 4.1 holds for i.i.d. $X_i \sim \mathcal{LC}$. Part (ii) of the theorem for i.i.d. $X_i \sim \mathcal{LC}$ follows from part (i) for i.i.d. $X_i \sim \mathcal{LC}$ and majorization comparisons (4). ■

References

Acerbi, C. (2002), 'Spectral measures of risk: A coherent representation of subjective risk aversion', *Journal of Banking and Finance* **26**, 1505–1518.

Acerbi, C. & Tasche, D. (2002), 'On the coherence of expected shortfall', *Journal of Banking and Finance* **26**, 1487–1503.

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**, 1551–1585.

Artzner, P., Delbaen, F., Eber, J.-M. & Heath, D. (1999), 'Coherent measures of risk', *Mathematical Finance* **9**, 203–228.

Axtell, R. L. (2001), 'Zipf distribution of U.S. firm sizes', *Science* **293**, 1818–1820.

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

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

Blattberg, R. C. & Gonedes, R. C. (1974), 'A comparison of the stable and student distributions as statistical models for stock prices', *Journal of Business* **47**, 244–280.

Borak, S., Härdle, W. & Weron, R. (2005), 'Stable distributions', *SFB 649 Discussion Paper 2005-08* . Humboldt University, Berlin.

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

Bretagnolle, J., Dacunha-Castelle, D. & Krivine, J. L. (1966), 'Lois stables et espaces L^p', *Annales de l'Institute H. Poincaré. Section B. Calcul des Probabilités et Statistique* **64**, 1278–1302.

Cambanis, S., Keener, R. & Simons, G. (1983), 'On $\alpha-$symmetric distributions', *Journal of Multivariate Analysis* **13**, 213–233.

Chan, W., Park, D. H. & Proschan, F. (1989), Peakedness of weighted averages of jointly distributed random variables, *in* L. J. Gleser, M. D. Perlman, S. J. Press & A. R. Sampson, eds, 'Contributions to Probability and Statistics', pp. 58–62.

Cotter, D. & Dowd, K. (2006), 'Extreme spectral risk measures: An application to futures clearinghouse margin requirements', *Journal of Banking and Finance* **30**, 3469–3485.

De Vany, A. S. & Walls, W. D. (2004), 'Motion picture profit, the stable paretian hypothesis and the curse of the superstar', *Journal of Economic Dynamics and Control* **28**, 1035–1057.

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

Eaton, M. L. (1970), 'A note on symmetric Bernoulli random variables', *Annals of Mathematical Statistics* **41**, 1223–1226.

Efron, B. (1969), 'Student's t-test under symmetry conditions', *Journal of the American Statistical Association* **64**, 1278–1302.

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

Embrechts, P., McNeil, A. & Straumann, D. (2002), Correlation and dependence in risk management: Properties and pitfalls, *in* M. A. H. Dempster, ed., 'Risk Management: Value at Risk and Beyond', Cambridge University Press, Cambridge, pp. 176–223.

Fabozzi, F. J., Focardi, S. M. & Kolm, P. N. (2006), *Financial Modeling of the Equity Market: CAPM to Cointegration*, Wiley, Hoboken, NJ.

Fama, E. (1965*a*), 'Portfolio analysis in a stable paretian market', *Management Science* **11**, 404–419.

Fama, E. F. (1965*b*), 'The behavior of stock market prices', *Journal of Business* **38**, 34–105.

Fang, K.-T., Kotz, S. & Ng, K. W. (1990), *Symmetric multivariate and related distributions*, Chapman and Hall, New York.

Fölmer, H. & Schied, A. (2002), 'Convex measures of risk and trading constraints', *Finance and Stochastics* **6**, 429–447.

Frittelli, M. & Gianin, E. R. (2002), 'Putting order in risk measures', *Journal of Banking & Finance* **26**, 1473–1486.

Gabaix, X. (1999*a*), 'Zipf's law and the growth of cities', *American Economic Review* **89**, 129–132.

Gabaix, X. (1999*b*), 'Zipf's law for cities: An explanation', *Quarterly Journal of Economics* **114**, 739–767.

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

Gabaix, X. & Ibragimov, R. (2006), '$LOG(RANK - 1/2)$: A simple way to improve the OLS estimation of tail exponents', *Harvard Institute of Economic Research Discussion Paper No. 2106* . http://www.economics.harvard.edu/faculty/ibragimov/files/LogLogRevised2.pdf.

Glasserman, P., Heidelberger, P. & Shahabuddin, P. (2002), 'Portfolio value-at-risk with heavy-tailed risk factors', *Mathematical Finance* **12**, 239–269.

Gneiting, T. (1998), 'On $\alpha-$symmetric multivariate characteristic functions', *Journal of Multivariate Analysis* **64**, 131–147.

Guillaume, D., Dacorogna, M., Davé, R., Müller, U. & Olsen, R. (1997), 'From the bird's eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets', *Finance and Stochastics* **1**, 95–129.

Hennessy, D. A. & Lapan, H. E. (2003), 'An algebraic theory of portfolio allocation', *Economic Theory* **22**, 193–210.

Ibragimov, M. & Ibragimov, R. (2007), 'Market demand elasticity and income inequality', *Economic Theory* **32**, 579–587. http://dx.doi.org/10.1007/s00199-006-0125-3.

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

Ibragimov, R., Jaffee, D. & Walden, J. (2008), 'Nondiversification traps in catastrophe insurance markets', *Review of Financial Studies,* doi: 10.1093/rfs/hhn021 .

Ibragimov, R. & Walden, J. (2007), '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 Also available as *Harvard Institute of Economic Research Discussion Paper No. 2104,* http://www.economics.harvard.edu/hier/2006papers/HIER2104.pdf.

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.

Jensen, D. R. (1997), 'Peakedness of linear forms in ensembles and mixtures', *Statistics and Probability Letters* **35**, 277–282.

Kahneman, D. & Tversky, A. (1979), 'Prospect theory: An analysis of decision under risk', *Econometrica* **47**, 263292.

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

Kuritsyn, Y. G. & Shestakov, A. V. (1984), 'On α-symmetric distributions', *Theory of Probability and Its Applications* **29**, 804–806.

Lapan, H. E. & Hennessy, D. A. (2002), 'Symmetry and order in the portfolio allocation problem', *Economic Theory* **19**, 747–772.

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

Lux, T. (1996), 'The stable Paretian hypothesis and the frequency of large returns: an examination of major German stocks.', *Applied Financial Economics* **6**, 463–475.

Ma, C. (1998), 'On peakedness of distributions of convex combinations', *Journal of Statistical Planning and Inference* **70**, 51–56.

Mandelbrot, B. (1963), 'The variation of certain speculative prices', *Journal of Business* **36**, 394–419.

Mandelbrot, B. (1997), *Fractals and scaling in finance. Discontinuity, concentration, risk*, Springer-Verlag, New York.

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

McCulloch, J. H. (1996), Financial applications of stable distributions, *in* G. S. Maddala & C. R. Rao, eds, 'Handbook of statistics, Vol. 14', Elsevier, Amsterdam, pp. 393–425.

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.

Praetz, P. (1972), 'The distribution of share price changes', *Journal of Business* **45**, 49–55.

Proschan, F. (1965), 'Peakedness of distributions of convex combinations', *Annals of Mathematical Statistics* **36**, 1703–1706.

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.

Ross, S. A. (1976), 'A note on a paradox in portfolio theory', *Working paper, University of Pennsylvania* .

Samuelson, P. A. (1967), 'Efficient portfolio selection for pareto-levy investments', *Journal of Financial and Quantitative Analaysis* **2**, 107–122.

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.

Szegö, G., E. (2004), *Risk measures for the 21st century*, Wiley, Chichester.

Tasche, D. (2002), 'Expected shortfall and beyond', *Journal of Banking and Finance* **26**, 1519–1533.

Tong, Y. L. (1994), 'Some recent developments on majorization inequalities in probability and statistics', *Linear Algebra and its Applications* **199**, 69–90.

Uchaikin, V. V. & Zolotarev, V. M. (1999), *Chance and stability. Stable distributions and their applications*, VSP, Utrecht.

Weron, R. (2001), 'Levy-stable distributions revisited: Tail index > 2 does not exclude the Levy-stable regime', *International Journal of Modern Physics, C* **12**, 209–223.

Zastavnyi, V. P. (1993), 'Positive definite functions depending on the norm', *Russian Journal of Mathematical Physics* **1**, 511–522.

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

Appendix A3: Extensions to heterogeneity and skewness

Theorems A3.1 and A3.2 in this appendix provide extensions of Theorems 4.1 and 4.2 to the case of skewed and heterogenous risks. We also present, in Theorems A3.3-A3.5, sharp VaR bounds for portfolios of such risks and those in the classes $\overline{\mathcal{CS}}(r)$ and $\underline{\mathcal{CS}}(r)$ that include, as a particular case, the results discussed in Remark 4.1. Theorems A3.3-A3.5 further provide VaR comparisons for portfolios of heavy-tailed risks that are implied by majorizations between vectors of powers of portfolio weights.

Let $\sigma_1, ..., \sigma_n \geq 0$ be some scale parameters and let $X_i \sim S_\alpha(\sigma_i, \beta, 0)$, $\alpha \in (0, 2]$, be independent not necessarily identically distributed stable risks. Further, let $Z_{\tilde{w}} = \sum_{i=1}^n w_{[i]} X_i$ be the return on the portfolio with weights $\tilde{w} = (w_{[1]}, ..., w_{[n]})$, where, as in Section 4, $w_{[1]} \geq ... \geq w_{[n]}$ denote the components of the vector $w = (w_1, ..., w_n) \in \mathbf{R}_+^n$ in decreasing order (a certain ordering in the components of the vector of weights w is necessary for the extensions of the majorization results in this paper to the case of non-identically distributed r.v.'s X_i since property (6) holds for all functions $\phi : \mathbf{R}_+^n \to \mathbf{R}$ that are Schur-convex or Schur-concave).

Observe that, in the case of identically distributed risks X_i, $i = 1, ..., n$, with $\sigma_1 = ... = \sigma_n$, the risk $Z_{\tilde{w}}$ has the same distribution as the return $Z_w = \sum_{i=1}^n w_i X_i$ on the portfolio of X_i's with weights $w = (w_1, ..., w_n)$, and, thus, $VaR_q(Z_{\tilde{w}}) = VaR_q(Z_w)$. That is, the VaR comparisons for $Z_{\tilde{w}}$ with the ordered weights $\tilde{w}$ under the above distributional assumptions cover the case of portfolio returns Z_w with identically distributed possibly skewed risks.

The proof of Theorems A3.1-A3.4 uses Schur-convexity and Schur-concavity properties of the functions

$$\chi(c_1, ..., c_n) = \sum_{i=1}^n \sigma_i^\alpha c_{[i]}^\alpha, \tag{16}$$

$\alpha > 0$, where σ_i, $i = 1, ..., n$, are the scale parameters of the risks in consideration. These properties are similar to the definitions of $p-$majorization for the vectors $(c_1^\alpha, ..., c_n^\alpha)$ with $p = (\sigma_1^\alpha, ..., \sigma_n^\alpha)$ (see Ch. 14 in Marshall & Olkin (1979)). This suggests that extensions of the majorization pre-ordering such as $p-$majorization may be useful in formalizations of the concept of diversification for portfolios of heterogenous risks.

Let $Q \sim S_\alpha(1, \beta, 0)$.

Theorem A3.1 *Let* $0 < q < P(Q > 0)$ *and let* $X_1, ..., X_n$ *be independent risks such that* $X_i \sim S_\alpha(\sigma_i, \beta, 0)$, *where* $\alpha \in (1, 2]$, $\sigma_1 \geq ... \geq \sigma_n > 0$ *and* $\beta \in [-1, 1]$. *Then* *(i)* $VaR_q(Z_{\tilde{v}}) < VaR_q(Z_{\tilde{w}})$ *if* $v \prec w$ *and* v *is not a permutation of* w *(in other words, the function* $\psi(w,q) = VaR_q(Z_{\tilde{w}})$ *is strictly Schur-convex in* $w \in \mathbf{R}_+^n$*).*

(ii) In particular, $VaR_q(Z_{\underline{w}}) < VaR_q(Z_{\tilde{w}}) < VaR_q(Z_{\overline{w}})$ *for all* $q \in (0,1/2)$ *and all weights* $w \in \mathcal{I}_n$ *such that* $w \neq \underline{w}$ *and* w *is not a permutation of* $\overline{w}$.

Theorem A3.2 *Let* $0 < q < P(Q > 0)$ *and let* $X_1, ..., X_n$ *be independent risks such that* $X_i \sim S_\alpha(\sigma_i, \beta, 0)$, *where* $\alpha \in (0,1)$, $\sigma_n \geq ... \geq \sigma_1 > 0$ *and* $\beta \in [-1,1]$. *Then*

(i) $VaR_q(Z_{\tilde{v}}) > VaR_q(Z_{\tilde{w}})$ *if* $v \prec w$ *and* v *is not a permutation of* w *(in other words, the function* $\psi(w,q) = VaR_q(Z_w)$, *is strictly Schur-concave in* $w \in \mathbf{R}_+^n$*).*

(ii) In particular, $VaR_q(Z_{\overline{w}}) < VaR_q(Z_{\tilde{w}}) < VaR_q(Z_{\underline{w}})$ *for all* $q \in (0,1/2)$ *and all weights* $w \in \mathcal{I}_n$ *such that* $w \neq \underline{w}$ *and* w *is not a permutation of* $\overline{w}$.

Theorem A3.3 *Let* $r \in (0,2]$, $0 < q < P(Q > 0)$ *and let* $X_1, ..., X_n$ *be independent risks such that* $X_i \sim S_\alpha(\sigma_i, \beta, 0)$, *where* $\alpha \in (r,2]$, $\sigma_1 \geq ... \geq \sigma_n > 0$, $\beta \in [-1,1]$, $\beta = 0$ *for* $\alpha = 1$. *Then*

(i) $VaR_q(Z_{\tilde{v}}) < VaR_q(Z_{\tilde{w}})$ *if* $(v_1^r, ..., v_n^r) \prec (w_1^r, ..., w_n^r)$ *and* $(v_1^r, ..., v_n^r)$ *is not a permutation of* $(w_1^r, ..., w_n^r)$ *(that is, the function* $\psi(w,q) = VaR_q(Z_{\tilde{w}})$, $w \in \mathbf{R}_+^n$, *is strictly Schur-convex in* $(w_1^r, ..., w_n^r)$*).*

(ii) The following sharp bounds hold:

$$n^{1-1/r}(\sum_{i=1}^n w_i^r)^{1/r} VaR_q(Z_{\underline{w}}) < VaR_q(Z_{\tilde{w}}) < (\sum_{i=1}^n w_i^r)^{1/r} VaR_q(Z_{\overline{w}})$$

for all $q \in (0,1/2)$ *and all weights* $w \in \mathcal{I}_n$ *such that* $w \neq \underline{w}$ *and* w *is not a permutation of* $\overline{w}$.

Theorem A3.4 *Let* $r \in (0,2]$, $0 < q < P(Q > 0)$ *and let* $X_1, ..., X_n$ *be independent risks such that* $X_i \sim S_\alpha(\sigma_i, \beta, 0)$, *where* $\alpha \in (0,r)$, $\sigma_n \geq ... \geq \sigma_1 > 0$, $\beta \in [-1,1]$, $\beta = 0$ *for* $\alpha = 1$. *Then*

(i) $VaR_q(Z_{\tilde{v}}) > VaR_q(Z_{\tilde{w}})$ *if* $(v_1^r, ..., v_n^r) \prec (w_1^r, ..., w_n^r)$ *and* $(v_1^r, ..., v_n^r)$ *is not a permutation of* $(w_1^r, ..., w_n^r)$ *(that is, the function* $\psi(w,q) = VaR_q(Z_{\underline{w}})$, $w \in \mathbf{R}_+^n$ *is strictly Schur-concave in* $(w_1^r, ..., w_n^r)$*).*

(ii) The following sharp bounds hold :

$$(\sum_{i=1}^n w_i^r)^{1/r} VaR_q(Z_{\overline{w}}) < VaR_q(Z_{\tilde{w}}) < n^{1-1/r}(\sum_{i=1}^n w_i^r)^{1/r} VaR_q(Z_{\underline{w}})$$

for all $q \in (0,1/2)$ *and all weights* $w \in \mathcal{I}_n$ *such that* $w \neq \underline{w}$ *and* w *is not a permutation of* $\overline{w}$.

Theorem A3.5 shows that, in the case of i.i.d. symmetric risks, the results in Theorems A3.3 and A3.4 hold for the convolution classes $\overline{\mathcal{CS}}(r)$ and $\underline{\mathcal{CS}}(r)$. As follows from the discussion at the beginning of this appendix, under the i.i.d. assumption in Theorem A3.5, one has $VaR_q(Z_{\tilde{v}}) = VaR_q(Z_v)$ and $VaR_q(Z_{\tilde{w}}) = VaR_q(Z_w)$.

Theorem A3.5 *Parts (i) and (ii) of Theorem A3.3 hold for all $q \in (0, 1/2)$ and i.i.d. risks $X_i \sim \overline{\mathcal{CS}}(r)$, $i = 1, ..., n$. Parts (i) and (ii) of Theorem A3.4 hold for all $q \in (0, 1/2)$ and i.i.d. risks $X_i \sim \underline{\mathcal{CS}}(r)$, $i = 1, ..., n$.*

Remark A3.1 Using conditioning arguments, one gets that the extensions provided by Theorems Theorems A3.1-A3.4 also hold in the case of random scale parameters σ_i. Similar to the proof of Theorems 5.1 and 5.2 and Theorems A3.1-A3.5, one can also show that analogues of these theorems hold for dependent heterogenous skewed risks, including convolutions (10) of common shock models (8) with skewed non-identically distributed risks Y_{ij}.

Proof of Theorems A3.1-A3.4.

Let $r, \alpha \in (0, 2]$, $\sigma_1, ..., \sigma_n > 0$, and let $v = (v_1, ..., v_n) \in \mathbf{R}^n_+$ and $w = (w_1, ..., w_n) \in \mathbf{R}^n_+$ be two vectors of portfolio weights such that $(v_1^r, ..., v_n^r) \prec (w_1^r, ..., w_n^r)$ and $(v_1^r, ..., v_n^r)$ is not a permutation of $(w_1^r, ..., w_n^r)$ (similar to the proof of Theorems 4.1 and 4.2, these assumptions evidently imply $\sum_{i=1}^n v_i \neq 0$ and $\sum_{i=1}^n w_i \neq 0$). Let $X_1, ..., X_n$ be independent risks such that $X_i \sim S_\alpha(\sigma_i, 0, 0)$, $i = 1, ..., n$. Similar to the proof of Theorems 4.1 and 4.2, we note that from (3) it follows that if $c = (c_1, ..., c_n) \in \mathbf{R}^n_+$, $\sum_{i=1}^n c_i \neq 0$, then $\sum_{i=1}^n c_{[i]} X_i / \left(\sum_{i=1}^n c_{[i]}^\alpha \sigma_i\right)^{1/\alpha} \sim S_\alpha(1, \beta, 0)$. Using positive homogeneity of the value at risk (see property a3 in Section 3), we thus obtain that, for all $0 < q < P(Q > 0)$,

$$VaR_q(\sum_{i=1}^n c_{[i]} X_i) = VaR_q(Q) \Big(\sum_{i=1}^n c_{[i]}^\alpha \sigma_i\Big)^{1/\alpha}. \tag{17}$$

By Theorem 3.A.4 in Marshall & Olkin (1979), the function $\chi(c_1, ..., c_n)$ defined in (16) is strictly Schur-convex in $(c_1, ..., c_n) \in \mathbf{R}^n_+$ if $\alpha > 1$ and $\sigma_1 \geq ... \geq \sigma_n \geq 0$ and is strictly Schur-concave in $(c_1, ..., c_n) \in \mathbf{R}^n_+$ if $\alpha < 1$ and $\sigma_n \geq ... \geq \sigma_1 \geq 0$ (see also Propositions 3.H.2.b and 4.B.7 in Marshall & Olkin (1979)). Therefore, we have $\sum_{i=1}^n v_{[i]}^\alpha \sigma_i^\alpha = \sum_{i=1}^n (v_{[i]}^r)^{\alpha/r} \sigma_i^\alpha < \sum_{i=1}^n (w_{[i]}^r)^{\alpha/r} \sigma_i^\alpha = \sum_{i=1}^n w_{[i]}^\alpha \sigma_i^\alpha$, if $\alpha/r > 1$ and $\sigma_1 \geq ... \geq \sigma_n > 0$. Similarly, $\sum_{i=1}^n w_{[i]}^\alpha \sigma_i^\alpha = \sum_{i=1}^n (w_{[i]}^r)^{\alpha/r} \sigma_i^\alpha < \sum_{i=1}^n (v_{[i]}^r)^{\alpha/r} \sigma_i^\alpha = \sum_{i=1}^n v_{[i]}^\alpha \sigma_i^\alpha$, if $\alpha/r < 1$ and $\sigma_n \geq ... \geq \sigma_1 > 0$. This, together with (17), implies that, for all $q \in (0, 1/2)$, $VaR_q(Z_{\tilde{v}}) < VaR_q(Z_{\tilde{w}})$ if $\alpha > r$ and $\sigma_1 \geq ... \geq \sigma_n \geq 0$ and $VaR_q(Z_{\tilde{v}}) > VaR_q(Z_{\tilde{w}})$ if $\alpha < r$ and $\sigma_n \geq ... \geq \sigma_1 \geq 0$ This completes the proof of parts (i) of Theorems A3.3 and A3.4.

The bounds in parts (ii) of Theorems A3.3 and A3.4 follow from their parts (i) and the property that, by majorization comparisons (4), $\big(\sum_{i=1}^n w_i^r/n, ..., \sum_{i=1}^n w_i^r/n\big) \prec (w_1^r, ..., w_n^r) \prec \big(\sum_{i=1}^n w_i^r, 0, ..., 0\big)$ for all portfolio weights $w \in \mathbf{R}_+^n$ and all $r \in (0, 2]$. Sharpness of the bounds in the theorems follows from the property that, as indicated in Remark 4.1, the bounds become equalities in the limit as $\alpha \to r$ for i.i.d. stable risks $X_i \sim S_r(\sigma, 0, 0)$.

Theorems A3.1 and A3.2 are consequences of Theorems A3.3 and A3.4 with $r = 1$.

■

Proof of Theorem A3.5.

Suppose that $X_1, ..., X_n$ are i.i.d. risks such that $X_i \sim \overline{\mathcal{CS}}(r)$, $i = 1, ..., n$. By definition of the class $\overline{\mathcal{CS}}(r)$, there exist independent r.v.'s Y_{ij}, $i = 1, ..., n$, $j = 1, ..., k$, such that $Y_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $\alpha_j \in (0, r)$, $\sigma_j > 0$, $i = 1, ..., n$, $j = 1, ..., k$, and $X_i = \sum_{j=1}^k Y_{ij}$, $i = 1, ..., n$. Using Theorem A3.3 for the r.v.'s $Y_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $\alpha_j \in (r, 2]$, $\sigma_j > 0$, $i = 1, ..., n$, with $\beta = 0$ and the implied equality $P(Q > 0) = 1/2$, we conclude that, for all $q \in (0, 1/2)$ and all $j = 1, ..., k$,

$$VaR_q\Big(\sum_{i=1}^n v_i Y_{ij}\Big) < VaR_q\Big(\sum_{i=1}^n w_i Y_{ij}\Big). \tag{18}$$

The densities of the r.v.'s Y_{ij}, $i = 1, ..., n$, $j = 1, ..., k$, are symmetric and unimodal by Proposition A1.1. Therefore, Proposition A1.4 implies that the densities of the r.v.'s $\sum_{i=1}^n v_i Y_{ij}$, $j = 1, ..., k$, and $\sum_{i=1}^n w_i Y_{ij}$, $j = 1, ..., k$, are symmetric and unimodal as well. By Proposition A1.5, this, together with relations (18), implies that $VaR_q(Z_v) = VaR_q\big(\sum_{j=1}^k \sum_{i=1}^n v_i Y_{ij}\big) < VaR_q\big(\sum_{j=1}^k \sum_{i=1}^n w_i Y_{ij}\big) = VaR_q(Z_w)$ for all $q \in (0, 1/2)$. Therefore, part (i) of Theorem A3.3 continues to hold for all $q \in (0, 1/2)$ and i.i.d. risks $X_i \sim \overline{\mathcal{CS}}(r)$, $i = 1, ..., n$. Using in the above arguments Theorem A3.4 instead of Theorem A3.3, in complete similarity one obtains that part (i) of Theorem A3.4 continues to hold for all $q \in (0, 1/2)$ and i.i.d. risks $X_i \sim \underline{\mathcal{CS}}(r)$, $i = 1, ..., n$. Parts (ii) of Theorems A3.4 and A3.3 for the classes $\overline{\mathcal{CS}}(r)$ and $\underline{\mathcal{CS}}(r)$ follow in the same way as in the proof of these theorems in the case of independent stable risks $X_i \sim S_\alpha(\sigma_i, \beta, 0)$. ■