

HIER

Harvard Institute of Economic Research

Discussion Paper Number 2116

Portfolio Diversification Under Local, Moderate
And Global Deviations From Power Laws

by

Rustam Ibragimov
and
Johan Walden

March 2006

HARVARD UNIVERSITY
Cambridge, Massachusetts

This paper can be downloaded without charge from:
<http://post.economics.harvard.edu/hier/2006papers/2006list.html>

The Social Science Research Network Electronic Paper Collection:
<http://ssrn.com/abstract=890641>

**PORTFOLIO DIVERSIFICATION UNDER LOCAL, MODERATE
AND GLOBAL DEVIATIONS FROM POWER LAWS¹**

Rustam Ibragimov

Department of Economics, Harvard University

Johan Walden

Haas School of Business, University of California at Berkeley

February 2006

Address for manuscript correspondence:

Rustam Ibragimov
Department of Economics
Harvard University
Littauer Center
1875 Cambridge St.
Cambridge, MA 02138

Email: ribragim@fas.harvard.edu
Phone: (617) 496-4795
Fax: (617) 495-7730

Johan Walden
Haas School of Business
University of California at Berkeley
545 Student Services Building, #1900
Berkeley, CA 94720-1900

Email: walden@haas.berkeley.edu
Phone: (510) 643-0547
Fax: (510) 643-1420

¹The authors are grateful to Donald Andrews, Xavier Gabaix, Peter Phillips, Herbert Scarf, Michael Schwarz 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 for helpful discussions.

ABSTRACT

This paper focuses on the analysis of portfolio diversification for a wide class of nonlinear transformations of heavy-tailed risks. We show that diversification of a portfolio of nonlinear transformations of thick-tailed risks increases riskiness if expectations of these functions are infinite. In addition, coherency of the value at risk measure is always violated for such portfolios. On the contrary, for nonlinearly transformed heavy-tailed risks with finite expectations, the stylized fact that diversification is preferable continues to hold. Moreover, in the latter setting, the value of risk is a coherent measure of risk. The framework of transformations of long-tailed random variables includes many models with Pareto-type distributions that exhibit local, moderate and global deviations from power tails in the form of additional slowly varying or exponential factors. This leads to a refined understanding of under what distributional assumptions diversification increases riskiness.

KEYWORDS: heavy-tailed risks; nonlinear transformations; portfolios; diversification; riskiness; value at risk; coherent measures of risk; risk bounds; robustness; Pareto-type distributions; power laws; local, moderate and global deviations

JEL Classification: G11

1 Introduction

1.1 Background

In the recent four decades, we have witnessed rapid expansion of the study of heavy-tailedness and the extreme outliers phenomena in economics and finance. Following Mandelbrot (1963) (see also the papers in Mandelbrot 1997, Fama 1965*b*), numerous studies have documented that time series encountered in many fields in economics and finance are typically thick-tailed and have infinite moments of order $p \geq \alpha$ for certain $\alpha > 0$ (see the discussion in Loretan & Phillips 1994, Meerschaert & Scheffler 2000, Gabaix et al. 2003, Ibragimov 2004*a,b*, 2005, Ibragimov & Walden 2006, and references therein).

In models involving a thick-tailed cdf F with infinite moments of order greater than or equal to α , it is typically assumed that F has Pareto (power) tails:

$$F(x) = \frac{c_1 + o(1)}{|x|^\alpha}, \quad x \rightarrow -\infty, \quad (1)$$

$$1 - F(x) = \frac{c_2 + o(1)}{x^\alpha}, \quad x \rightarrow +\infty, \quad (2)$$

or, more generally, that F is of Pareto-type, so that

$$F(x) = \frac{c_1 + o(1)}{|x|^\alpha} l(|x|), \quad x \rightarrow -\infty, \quad (3)$$

$$1 - F(x) = \frac{c_2 + o(1)}{x^\alpha} l(x), \quad x \rightarrow +\infty, \quad (4)$$

where c_1, c_2 are some positive constants and $l(x)$ is a slowly varying function at infinity:

$$\frac{l(\lambda x)}{l(x)} \rightarrow 1$$

as $x \rightarrow +\infty$ for all $\lambda > 0$. If a cdf F satisfies the Pareto-type law (3)-(4) with $\alpha \in (0, 2)$, then it belongs to the domain of attraction of a stable distribution with the characteristic exponent α (stable distributions are those that are closed under portfolio formation; see Section 2 for the definition and the review of the main properties of stable distributions). This means that, for a sequence of i.i.d. r.v.'s X_t , $t \geq 1$, with cdf's F there exist sequences of numbers $a_t > 0$, $t \geq 1$, and $b_t \in \mathbf{R}$, $t \geq 1$, such that the distributions of the sums

$$Z_n = \frac{1}{a_n} \sum_{t=1}^n X_t - b_n \quad (5)$$

weakly converge to a stable cdf G whose tails satisfy (1)-(2). If the tails of a cdf F satisfy the power law (3)-(4) with $\alpha > 2$, then F belongs to the domain of attraction of a normal distribution, that is, sums (5) converge weakly to a Gaussian r.v. for some sequences $a_t > 0$, $t \geq 1$, and $b_t \in \mathbf{R}$, $t \geq 1$.

Mandelbrot (1963) presented evidence that historical daily changes of cotton prices have the tail index $\alpha \approx 1.7$, and thus have infinite variances. As discussed in, e.g., Ibragimov (2004*a,b*, 2005), subsequent research

reported the following estimates of the tail parameter 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).

Recent studies (see Gabaix et al. 2003, and references therein) have found that the returns on many stocks and stock indices have the tail exponent $\alpha \approx 3$, while the distributions of trading volume and the number of trades in financial markets have the tail indices $\alpha \approx 1.5$ and $\alpha \approx 3.4$, respectively. As discussed in Gabaix et al. (2003), these estimates of the tail indices α are robust to different types and sizes of financial markets, market trends and are similar for different countries. Motivated by these empirical findings, Gabaix et al. (2003) proposed a model that demonstrated that the above values of the tail indices for stock returns, trading volume and the number of trades are explained by trading of large market participants, namely, the largest mutual funds whose sizes have the tail exponent $\alpha \approx 1$. Tail indices of $\alpha \approx 1$ (Zipf laws) have also been found for firm sizes (Axtell 2001) and city sizes (see Gabaix 1999*a,b*, for the discussion and explanations of the Zipf law for cities). Moreover, according to the results obtained by Schwarz (2000), distributions exhibiting Pareto-type behavior (3)-(4) with $\alpha = 1$ arise naturally as posteriors in several problems of decision making under uncertainty (see Theorem 5.3 in Schwarz 2000).

De Vany & Walls (2004) showed that stable distributions with tail indices $1 < \alpha < 2$ that obey (1)-(2) provide a good model for distributions of profits in motion pictures. Moreover, some studies have indicated that the tail exponent is close to one or is 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). Furthermore, Scherer et al. (2000) and Silverberg & Verspagen (2004) reported the tail indices α to be considerably less than one for profit outcomes from technological innovations.

Recently, Ibragimov (2004*a,b*, 2005) developed a unified approach to the analysis of value at risk (VaR) theory and other economic models under heavy-tailedness using new majorization theory for linear combinations of thick-tailed random variables (r.v.'s). Generalizing the results on portfolio choice in the stable framework and riskiness analysis for uniform portfolios of stable risks (see Fama 1965*a*, Samuelson 1967, Ross 1976), Ibragimov (2004*a,b*, 2005) showed, in particular, that the stylized fact of portfolio diversification being preferable is reversed for riskiness comparisons of portfolios with arbitrary weights for a wide class of extremely thick-tailed risks whose cdf's F satisfy power law (1)-(2) with $\alpha < 1$ and thus have infinite first moments. Specifically, for such distributions, the VaR is strictly increasing in the degree of diversification (see Proposition 4 in the present paper). According to the results in Ibragimov (2004*a,b*, 2005), the stylized facts on portfolio diversification are robust to heavy-tailedness of risks or returns, as long as their cdf's are not extremely thick-tailed. They continue to hold for distributions that satisfy relations (1)-(2) with $\alpha > 1$ (and thus have finite means). Ibragimov & Walden (2006) demonstrated, among other results, that the above VaR results hold as well for a wide class of bounded risks concentrated on a sufficiently large interval.²

²More specifically, according to the results in Ibragimov (2004*a,b*, 2005), the stylized facts on diversification being preferable (as well as the properties of many economic models) are reversed for convolutions of α -symmetric distributions with $\alpha < 1$. These stylized facts, however, continue to hold for convolutions of α -symmetric distributions with $\alpha > 1$. The results in Ibragimov & Walden (2006) cover, among others, truncated versions of convolutions of α -symmetric densities. Convolutions of α -symmetric distributions provide a natural framework for modeling distributions that exhibit both heavy-tailedness in marginals and dependence among them. An n -dimensional distribution is called α -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 $\alpha > 0$. The class of α -symmetric distributions is very wide and includes, in particular, i.i.d.

Value at risk and the closely related safety first principle provide natural alternatives to the traditional expected utility approach in the presence of thick-tailedness. For heavy-tailed distributions the expected utility framework is not readily available since it typically involves assumptions about the existence of moments for the risks in consideration. The VaR and safety-first approaches to portfolio selection are thus, in many regards, the only ones available in the presence of heavy-tailedness. Moreover, as noted in Ibragimov & Walden (2006), as the results hold for *arbitrary* VaR levels, they also hold for natural generalizations of the concept of “riskiness” to heavy-tailed distributions. The failure to reduce risk by diversifying, henceforth denoted the *failure of diversification*, is thus genuine in this case. It is not a failure of VaR as a risk measure.

The fact that a number of economic and financial time series have tail exponents approximately equal to (or even slightly less than) one is important in the context of the results obtained in Ibragimov (2004*a,b*, 2005): As those works demonstrate, the value of the tail index $\alpha = 1$ (that is, the existence of the first moment) is exactly the critical boundary between robustness of implications of many economic models to heavy-tailedness assumptions and their reversals. More precisely, the implications of most of the models are robust to thick-tailedness assumptions with tail indices $\alpha > 1$ and finite first moments, but the implications of the models are reversed for extremely heavy-tailed distributions with $\alpha < 1$ for which the first moments are infinite.

1.2 Objectives and key results

In this paper, we analyze portfolio diversification for nonlinear transformations of heavy-tailed risks, using the framework of value at risk as a measure of portfolio riskiness. The main objective of the analysis is to further understand under what distributional assumptions failure of diversification occurs. We refine the previous analyses (of Ibragimov 2004*a,b*, 2005) in several ways. The concept of diversification of risk is fundamental to finance. We therefore believe that the detailed analysis of when it fails provided in this paper is well-motivated.

Our first, and main, result is that diversification of a portfolio of nonlinear functions of thick-tailed risks increases its riskiness if expectations of these functions are infinite (Theorem 2). In addition, coherency of the value of risk is always violated in the world of such portfolios (Corollary 2). However, the stylized fact that diversification is preferable continues to hold for nonlinearly transformed heavy-tailed risks with finite expectations (Theorem 1). Furthermore, the value of risk is a coherent measure of risk in the latter setting (Corollary 1).

The main results on nonlinear transformations of heavy-tailed r.v.’s provides a natural framework for modeling distributions exhibiting deviations from power laws. We use this to obtain a second set of results. We analyze deviations of the form (3)-(4) and also more general deviations (see Section 2). First, let us define recursively the iterations of a logarithm by $\ln_0(x) = x$, $\ln_k(x) = \ln[\ln_{k-1}(x)]$, $k \geq 1$. Let $m \geq 0$ and let $\gamma_0 > 0, \gamma_1, \dots, \gamma_m \in \mathbf{R}$ be some constants. The stochastic framework considered in the paper covers risks Y

stable distributions and spherical distributions corresponding to $\alpha = 2$. Important examples of spherical distributions, in turn, are given by Kotz type, multinormal and logistic distributions and multivariate stable laws. In addition, they include a subclass of mixtures of normal distributions as well as multivariate t -distributions that were used in a number of papers to model heavy-tailedness phenomena with dependence and finite moments up to a certain order. Moreover, the class of α -symmetric distributions includes a wide class of convolutions of models with common shocks affecting all heavy-tailed risks (such as macroeconomic or political ones, see Andrews 2003).

whose tails behave as

$$P(|Y| > x) \asymp x^{-\alpha/\gamma_0} \prod_{i=1}^m [\ln_i(x)]^{\alpha\gamma_i/\gamma_0} = x^{-\alpha/\gamma_0} (\ln x)^{\alpha\gamma_1/\gamma_0} (\ln \ln x)^{\alpha\gamma_2/\gamma_0} \dots (\ln \ln \dots \ln x)^{\alpha\gamma_m/\gamma_0} \text{ as } x \rightarrow \infty \quad (6)$$

(here and throughout the paper, $g(x) \asymp h(x)$ as $x \rightarrow \infty$ denotes that there are constants, c and C such that $0 < c \leq g(x)/h(x) \leq C < \infty$ for large $x > 0$). In particular, the choice $\gamma_0 = 1$, $\gamma_k = -1/\alpha$, $\gamma_s = 0$, $1 \leq s \leq m$, $s \neq k$ produces deviations from power law (1)-(2) of the form

$$P(|Y| > x) \asymp \frac{1}{x^\alpha \ln_k(x)}. \quad (7)$$

The choice $\gamma_0 = 1$, $\gamma_k = 1/\alpha$, $\gamma_s = 0$, $1 \leq s \leq m$, $s \neq k$ corresponds to the deviations from power law (1)-(2) of the form

$$P(|Y| > x) \asymp \frac{\ln_k(x)}{x^\alpha}. \quad (8)$$

If $\alpha = 1$, relations (7) and (8) correspond to deviations from the Zipf law (1)-(2) with $\alpha = 1$ of the form

$$P(|Y| > x) \asymp \frac{1}{x \ln_k(x)}; \quad (9)$$

$$P(|Y| > x) \asymp \frac{\ln_k(x)}{x}. \quad (10)$$

Reminiscent of the terminology in the time series unit root literature (see Phillips 1988, Ibragimov & Phillips 2004, Phillips & Magdalinos 2004), it is natural to refer to distributions whose tails satisfy one of relations (6)-(8) as exhibiting “local” or “moderate” deviations from power laws. Similarly, it is natural to call the departures from power tails with the index $\alpha = 1$ in form (9) or (10) as “local” or “moderate” deviations from the Zipf law. Next, let $\delta_0, \delta_1, \delta_2 > 0$. Nonlinear transformations of heavy-tailed risks considered in the paper model the classes of r.v.’s Y whose tails behave as

$$P(|Y| > x) \asymp x^{-\alpha/\delta_0} \exp\left(-(\delta_1\alpha/\delta_0)x^{\delta_2/\delta_0}\right). \quad (11)$$

It is natural to refer to distributions with the tail behavior in form (11) as exhibiting “global” deviations from power laws.

The main results on nonlinear transformations of heavy-tailed risks imply Corollaries 3 and 4. These corollaries concern portfolio diversification and robustness of economic models under “local”, “moderate” and “global” deviations from power laws in forms (6)-(11): We sharpen the “robustness versus reversals” results in Ibragimov (2004*a,b*, 2005), Ibragimov & Walden (2006) discussed in Subsection 1.1. We show that even local or moderate deviations from the Zipf law or a Cauchy-type distribution (1)-(2) with the index $\alpha = 1$ can lead to reversals of properties of models in finance and economics. For instance, as we demonstrate, if such models involve distributions whose tails are slightly thinner than those of Cauchy distributions and behave as (9), then the properties of the models are similar to those under the standard assumptions of Gaussianity. In particular, for risks with such distributions, the stylized facts on diversification being preferable continue to

hold. More generally, the stylized facts on portfolio diversification continue to hold for portfolios of nonlinear transformations of risks with the tail index $\alpha \geq 1$ whose distributions have tails satisfying (6) with $\gamma_0 < 1$ or with $\gamma_0 = 1, \gamma_1 = 0, \dots, \gamma_{k-1} = 0, \gamma_k > 0$ for some $k \in \{1, 2, \dots, m\}$ (see Corollary 3). Analogous conclusions hold for distributions exhibiting “global” deviations from power laws (11).

If, however, the r.v.’s entering the models’ assumptions have tails (10) even slightly heavier than those of a Cauchy distribution, with the iteration of a logarithm in the numerator, rather than in the denominator in (9), then the properties of the models become completely the opposite to those under the “normal” assumptions. In particular, our results show that, similar to the case of extremely heavy-tailed distributions in Ibragimov (2004a,b, 2005), the stylized fact that portfolio diversification is always preferable is reversed for risks with tails (10). Namely, in the case of risks with such “local to Cauchy” or “local to Zipf” distributions, diversification of a portfolio always leads to an increase in riskiness of the portfolio’s return. The same diversification failure also holds for portfolios of nonlinear functions of risks with $\alpha \leq 1$ whose distributions have tails satisfying (6) with $\gamma_0 > 1$ or with $\gamma_0 = 1, \gamma_1 = 0, \dots, \gamma_{k-1} = 0, \gamma_k < 0$ for some $k \in \{1, 2, \dots, m\}$ (see Corollary 4). These results are thus a refinement of the boundary case, $\alpha = 1$, where diversification reverses from being preferable to being inferior, as shown shown in Figure 1.

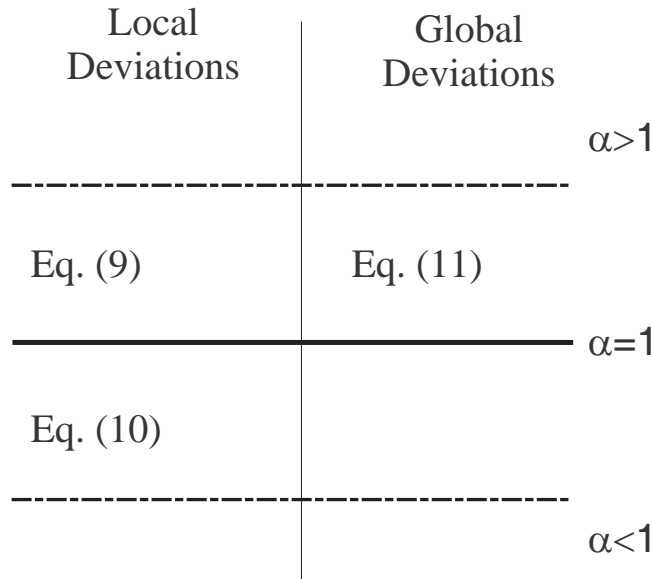


Figure 1: Applications of theory to refining the case $\alpha = 1$. With local deviations of the form (9) and global deviations of the form (11), diversification is to be preferred. With local deviations of the form (10), diversification fails.

Our third contribution is to demonstrate that in the world of risks with distributions (9) and (11), the value at risk satisfies the important condition of coherency. This is similar to the results in Ibragimov (2004a,b, 2005). Corollary 3 shows that coherency of the value at risk measure holds for transformations of risks with $\alpha \geq 1$ that satisfy (6) with $\gamma_0 < 1$ or with $\gamma_0 = 1, \gamma_1 = 0, \dots, \gamma_{k-1} = 0, \gamma_k > 0$ for some $k \in \{1, 2, \dots, m\}$. However, coherency of the value at risk is always violated if distributions of risks have local deviations from Cauchy distributions (8). More generally, it is also violated for transformations of risks with $\alpha \leq 1$ whose distributions

satisfy (6) with $\gamma_0 > 1$ or with $\gamma_0 = 1, \gamma_1 = 0, \dots, \gamma_{k-1} = 0, \gamma_k < 0$ for some $k \in \{1, 2, \dots, m\}$ (Corollary 4).

Our fourth, and final contribution is an analysis of portfolio choice for chi-square distributions. Distributions exhibiting “global” deviations from power laws in form (11) naturally appear in models involving power functions and other transformations of normal r.v.’s. In particular, the choice $\delta_0 = \delta_2 = 2\alpha$ and $\delta_1 = 1$ in (11) produces cdf’s F whose tails have the same rate of decline as those of chi-square distributions:

$$1 - F(x) \asymp \frac{\exp(-x/2)}{\sqrt{x}}. \quad (12)$$

Motivated by this property and using the results on extrema of cdf’s of quadratic forms in Gaussian r.v.’s obtained by Székely & Bakirov (2003), we formulate several characterizations of the optimal safety-first portfolio choice in the world of risks with chi-square distributions (Propositions 5 and 6). We show that the optimal safety-first portfolio of chi-square risks is the one with equal weights for large values of the disaster levels z . The optimal portfolio in such a setting is the one consisting of only one risk in the case of small values of the disaster levels. However, in contrast to nonlinear transformations of heavy-tailed risks with symmetric densities, the optimal portfolio choice consists of exactly two chi-square risks if the disaster level is moderate.

The results obtained in this paper further advance those in Ibragimov (2004*a,b*, 2005), Ibragimov & Walden (2006) and illustrate the dangers in misidentification of the distributional tail behavior in many economic and financial models. They further illustrate the dangers associated with the over- or underestimation of the tail indices under heavy-tailedness. In addition, the results on non-coherency of the value at risk under extreme thick-tailedness with infinite first moments obtained in Ibragimov (2004*a,b*, 2005) and in the present paper emphasize the necessity in the development of coherent risk measures that are finite for wide classes of heavy-tailed distributions.³

1.3 Organization of paper

The paper is organized as follows. Section 2 introduces classes of distributions we are dealing with throughout the paper and discusses their main properties. Section 3 discusses how the stochastic framework proposed in the paper incorporates deviations from power law distributions. Sections 4.1 and 4.2 present our main results about value at risk and (non)coherency of portfolios of nonlinear transformations of heavy-tailed risks. Section 4.3 discusses implications of the results for “local”, “moderate” and “global” deviations from power laws of forms (6)-(10) and (11). This leads to a refined understanding of under what distributional assumptions diversification fails, and when coherency is violated. Section 4.4 discusses the results in the paper from a safety-first portfolio choice perspective. Finally, Section 5 makes some concluding remarks. All proofs are left to the Appendix.

³One should note here that several recent papers (see, among others Acerbi & Tasche 2002, Tasche 2002, and references therein) recommended to use the expected shortfall as a coherent alternative to the value at risk. However, the expected shortfall, which is defined as the average of the worst losses of a portfolio, requires existence of the first moments of risks to be finite. It is not difficult to see that assumptions close to existence of means of the risks in considerations are also required for applications of coherent spectral measures of risk (see Acerbi 2002, Cotter & Dowd 2002) that generalize the expected shortfall.

2 Notations and classes of distributions

In this section, we introduce certain classes of distributions we will be dealing with throughout the paper. The notations for some of these classes are similar to those in Ibragimov (2004a,b, 2005) and Ibragimov & Walden (2006).

We say that a r.v. X with density $f : \mathbf{R} \rightarrow \mathbf{R}$ 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 $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq (f(x_1))^\lambda (f(x_2))^{1-\lambda}. \quad (13)$$

(see An 1998). A distribution is said to be log-concave if its density f satisfies (13). Examples of log-concave distributions include (see, for instance Marshall & Olkin 1979, p. 493) the normal distribution $\mathcal{N}(\mu, \sigma^2)$, the uniform density $\mathcal{U}(\theta_1, \theta_2)$, the exponential density, the logistic distribution, the Gamma distribution $\Gamma(\alpha, \beta)$ with the shape parameter $\alpha \geq 1$, the Beta distribution $\mathcal{B}(a, b)$ with $a \geq 1$ and $b \geq 1$; the Weibull distribution $\mathcal{W}(\gamma, \alpha)$ with the shape parameter $\alpha \geq 1$.

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 $x \rightarrow \infty$ and all the power moments $E|X|^\gamma$, $\gamma > 0$, of the r.v. exist (see Corollary 1 in An 1998). This implies, in particular, that distributions with log-concave densities *cannot* be used to model heavy-tailed phenomena.

As in Ibragimov (2004a,b, 2005) and Ibragimov & Walden (2006), we denote by \mathcal{LC} the class of symmetric log-concave distributions.⁴ 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

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

$x \in \mathbf{R}$, where $i^2 = -1$ and $\text{sign}(x)$ is the sign of x defined by $\text{sign}(x) = 1$ if $x > 0$, $\text{sign}(0) = 0$ and $\text{sign}(x) = -1$ otherwise. In what follows, we write $X \sim S_\alpha(\sigma, \beta, \mu)$, if the r.v. X has the stable distribution $S_\alpha(\sigma, \beta, \mu)$.

As is well-known, 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); $\alpha = 1/2$ and $\beta \pm 1$ (Lévy distributions).⁵ 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 $X \sim S_\alpha(\sigma, \beta, \mu)$, then its distribution satisfies power law (1)-(2). 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 infinite otherwise.

⁴ \mathcal{LC} stands for "log-concave".

⁵The densities of Cauchy distributions are $f(x) = \sigma/(\pi(\sigma^2 + (x - \mu)^2))$; as is indicated before, Lévy distributions 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.

The symmetry index β characterizes the skewness of the distribution. The stable distributions with $\beta = 0$ are symmetric about the location parameter μ . 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 $(-\infty, \mu]$ (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 $\beta \neq 0$ for $\alpha = 1$ are called strictly stable. If $X_i \sim S_\alpha(\sigma, \beta, \mu)$, $\alpha \in (0, 2]$, are i.i.d. strictly stable r.v.'s, then, for all $a_i \geq 0$, $i = 1, \dots, n$,

$$\sum_{i=1}^n a_i X_i / \left(\sum_{i=1}^n a_i^\alpha \right)^{1/\alpha} \sim S_\alpha(\sigma, \beta, \mu). \quad (14)$$

For a detailed review of properties of stable distributions the reader is referred to, e.g., the monographs by Zolotarev (1986), Embrechts et al. (1997), Uchaikin & Zolotarev (1999), Rachev & Mittnik (2000), Rachev et al. (2005).

Further, we consider the class $\overline{\mathcal{CS}}$ of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in (1, 2]$ and $\sigma > 0$.⁶ That is, $\overline{\mathcal{CS}}$ consists of distributions of r.v.'s X such that, for some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (1, 2]$, $\sigma_i > 0$, $i = 1, \dots, k$.

By $\overline{\mathcal{CSLC}}$, we denote the class of convolutions of distributions from the classes \mathcal{LC} and $\overline{\mathcal{CS}}$. That is, $\overline{\mathcal{CSLC}}$ is the class of convolutions of symmetric distributions which are either log-concave or stable with characteristic exponents greater than one.⁷ 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 $\overline{\mathcal{CS}}$.

Also, $\underline{\mathcal{CS}}$ stands for the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with indices of stability $\alpha \in (0, 1)$ and $\sigma > 0$.⁸ That is, $\underline{\mathcal{CS}}$ consists of distributions of r.v.'s X such that, for some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (0, 1)$, $\sigma_i > 0$, $i = 1, \dots, k$.

We note that (see Ibragimov 2004a,b, 2005) the class $\overline{\mathcal{CS}}$ of *convolutions* of symmetric stable distributions with *different* indices of stability $\alpha \in (1, 2]$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (1, 2]$ and $\sigma > 0$. Similarly, the class $\underline{\mathcal{CS}}$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (0, 1)$ and $\sigma > 0$. Clearly, one has $\mathcal{LC} \subset \overline{\mathcal{CSLC}}$, $\overline{\mathcal{CS}} \subset \overline{\mathcal{CSLC}}$ and $\overline{\mathcal{CSLC}} \subset \overline{\mathcal{CSLC}}$. Note also 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 $\sigma > 0$.

Let $\mathbf{R}_+ = [0, \infty)$. Throughout the paper, $\overline{\mathcal{M}}$ denotes the class of differentiable odd functions $f : \mathbf{R} \rightarrow \mathbf{R}$ such that f is concave and increasing on \mathbf{R}_+ and $\underline{\mathcal{M}}$ denotes the class of odd functions $f : \mathbf{R} \rightarrow \mathbf{R}$ such that

⁶Here and below, \mathcal{CS} stands for "convolutions of stable"; the overline indicates relation to stable distributions with indices of stability *greater* than the threshold value 1.

⁷ $\overline{\mathcal{CSLC}}$ stands for "convolutions of stable and log-concave".

⁸The underline in $\underline{\mathcal{CS}}$ indicates relation to stable distributions with indices of stability *less* than the threshold value 1.

f is convex and increasing on \mathbf{R}_+ . Further, \overline{M}' (resp. \underline{M}') denotes the subclass of \overline{M} (resp. \underline{M}) consisting of functions f which are strictly concave (strictly convex) on \mathbf{R}_+ .

By $\overline{\mathcal{CTSLC}}$, we denote the class of convolutions of log-concave distributions and distributions of transforms $f(Y)$, $f \in \overline{M}$, of symmetric stable r.v.'s $Y \sim S_\alpha(\sigma, 0, 0)$ with characteristic exponents $\alpha \in [1, 2]$ and $\sigma > 0$.⁹ More precisely, $\overline{\mathcal{CTSLC}}$ consists of distributions of r.v.'s X such that, for some $k \geq 1$, and independent r.v.'s $Y_0 \sim \mathcal{LC}$ and $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in [1, 2]$, $\sigma_i > 0$, $i = 1, \dots, k$,

$$X = \theta Y_0 + f_1(Y_1) + \dots + f_k(Y_k), \quad (15)$$

where $\theta \in \{0, 1\}$, $f_i \in \overline{M}$, $i = 1, \dots, k$; $f_i \in \overline{M}'$ if $\alpha_i = 1$.¹⁰

It is not difficult to see that if $X \sim \overline{\mathcal{CTSLC}}$ or if $X = f(Y)$, where $Y \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \geq 1$, and $f \in \overline{M}$; $f \in \overline{M}'$ for $\alpha = 1$, then $E|X| < \infty$. Similarly, if $X = f(Y)$, where $f \in \underline{M}$ and $Y \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \geq 1$, then the first moment of X is infinite: $E|X| = \infty$.

In some sense, symmetric (about 0) Cauchy distributions $S_1(\sigma, 0, 0)$ are at the dividing boundary between the classes $\underline{\mathcal{CS}}$ and $\overline{\mathcal{CSLC}}$. Similarly, Cauchy-type distributions $S_1(\sigma, \beta, 0)$ are at the dividing boundary between the class of nonlinear transformations $X = f(Y)$, $f \in \overline{M}$, of stable r.v.'s $Y \sim S_\alpha(\sigma, \beta, 0)$ with $\alpha \geq 1$ and the family of nonlinear functions $X = f(Y)$, $f \in \underline{M}$, of stable r.v.'s $Y \sim S_\alpha(\sigma, \beta, 0)$ with $\alpha \leq 1$.

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

3 Modeling departures from power tails

As indicated in the introduction, the class of nonlinear transformations of heavy-tailed r.v.'s considered in this paper provides a natural framework for modeling risks with distributions exhibiting “local” to “moderate” departures from power laws in form (3)-(4), including (6)-(10), as well as their “global” analogues such as (11).

Indeed, let, throughout this section, Z be a r.v. whose cdf F obeys power law (1)-(2). It is easy to see that if T is an odd increasing on \mathbf{R} function, then the tails of the distribution of the transformation $Y = T(Z)$ of Z satisfy

$$P(|Y| > x) = P(|T(Z)| > x) \asymp \frac{1}{[T^{-1}(x)]^\alpha}, \quad x \rightarrow \infty.$$

Consequently, if T is such that the function $[T^{-1}(x)/x]^\alpha$ is slowly varying at infinity, then the cdf of the r.v. Y has Pareto-type tails in form (3)-(4). In particular, if the cdf of Z satisfies Zipf law, that is, exhibits power decay (1)-(2) with $\alpha = 1$, then the distribution of the transformation $Y = T(Z)$ satisfies

$$P(|Y| > x) \asymp \frac{1}{T^{-1}(x)}, \quad (16)$$

and, thus, has Pareto-type form (3)-(4) with the tail index $\alpha = 1$ if $T^{-1}(x)/x$ is slowly varying at infinity.

⁹ $\overline{\mathcal{CTSLC}}$ stands for “convolutions of transforms of stable and log-concave”.

¹⁰The latter condition implies, in particular, that the distributions of the r.v.'s in the class $\overline{\mathcal{CTSLC}}$ are different from Cauchy distributions $S_1(\sigma, 0, 0)$ with $\alpha = 1$.

Let, as in the introduction, $\ln_s(x)$, $s \geq 0$, stand for iterations of a logarithm: $\ln_0(x) = x$, $\ln_k(x) = \ln[\ln_{k-1}(x)]$, $k \geq 1$. Further, let, as before, $m \geq 0$ and let $\gamma_0 > 0, \gamma_1, \dots, \gamma_m \in \mathbf{R}$ and $x_0 \geq e^m$ be some constants. Consider the odd increasing on \mathbf{R} function V defined by

$$V(x) = (x + x_0)^{\gamma_0} \prod_{i=1}^m \left[\ln_i(x + x_0) \right]^{\gamma_i} - x_0^{\gamma_0} \prod_{i=1}^m \left[\ln_i(x_0) \right]^{\gamma_i}, \quad V(-x) = -V(x), \quad x > 0.$$

It is not difficult to see that the function $V(x)$ is strictly convex on $[0, \infty)$ for a sufficiently large x_0 if $\gamma_0 > 1$ or if $\gamma_0 = 1, \gamma_1 = 0, \dots, \gamma_{k-1} = 0, \gamma_k > 0$ for some $k \in \{1, 2, \dots, m\}$. Similarly, the function $V(x)$ is strictly concave on $[0, \infty)$ for a sufficiently large x_0 if $\gamma_0 < 1$ or if $\gamma_0 = 1, \gamma_1 = 0, \dots, \gamma_{k-1} = 0, \gamma_k < 0$ for some $k \in \{1, 2, \dots, m\}$. It is not difficult to show that, with the above x_0 ,

$$V^{-1}(x) \asymp x^{1/\gamma_0} \prod_{i=1}^m \left[\ln_i(x) \right]^{-\gamma_i/\gamma_0} \quad \text{as } x \rightarrow \infty.$$

Consequently, according to the above discussion, the tails of the distributions of the transformations $Y = V(Z)$ of the r.v. Z satisfy relation (6) that includes, as special cases, deviations from power laws in forms (7)-(10).

In order to obtain transformations $Y = V(Z)$ whose cdf's satisfy relation (6), one can equivalently define V to be the inverse $V(x) = H^{-1}(x)$, $x \in \mathbf{R}$, of the function

$$H(x) = (x + x_0)^{1/\gamma_0} \prod_{i=1}^m \left[\ln_i(x + x_0) \right]^{-\gamma_i/\gamma_0} - x_0^{1/\gamma_0} \prod_{i=1}^m \left[\ln_i(x_0) \right]^{-\gamma_i/\gamma_0}, \quad H(-x) = -H(x), \quad x > 0,$$

where $x_0 > 0$ is such that H is well-defined, increasing and strictly convex on $[0, \infty)$ (clearly, such x_0 always exists).

Similarly, let $\delta_0, \delta_1, \delta_2 > 0$ and let W be the inverse $W(x) = Q^{-1}(x)$ of the function

$$Q(x) = (x + x_0)^{1/\delta_0} \exp\left((\delta_1/\delta_0)(x + x_0)^{\delta_2/\delta_0}\right) - x_0^{1/\delta_0} \exp\left((\delta_1/\delta_0)x_0^{\delta_2/\delta_0}\right), \quad Q(-x) = -Q(x), \quad x > 0,$$

where $x_0 > 0$ is such that W is well-defined, increasing and strictly convex on $[0, \infty)$ (as above, such x_0 always exists). Then W is well-defined, odd and increasing function on \mathbf{R} which is strictly concave on $[0, \infty)$. In addition, by construction, the tails of the distributions of the transformations $Y = W(Z)$ of Z satisfy (11).

4 Main results

4.1 Diversification of nonlinear transformations of heavy-tailed risks

This section presents the main results of the paper on portfolio diversification for nonlinear transformations of heavy-tailed risks. These results are given by Theorems 1 and 2 below. Before formulating them, we briefly review the results on diversification for portfolios of light-tailed and heavy-tailed distributions available in the literature.

Let $0 < q < 1/2$. In what follows, given a r.v. (risk) Z , we denote by $VaR_q[Z]$ the value at risk (VaR) of

Z at level q , that is, its $(1 - q)$ -quantile.^{11,12} For $n \geq 1$, the sample mean \bar{X}_n represents the return on the portfolio of risks X_1, \dots, X_n with equal weights $w_n = (1/n, 1/n, \dots, 1/n)$: $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Proschan (1965) obtained results on majorization properties of tail probabilities of linear combinations of log-concavely distributed r.v.'s and transformations of Cauchy r.v.'s from which it follows that the following propositions hold (see the discussion in Ibragimov 2004a,b, 2005).

Proposition 1 (Proschan 1965) *The following conclusions hold.*

- Let X_1, \dots, X_n , $n \geq 2$, be i.i.d risks such that $X_1 \sim \mathcal{LC}$. Then $VaR_q[\bar{X}_n] < VaR_q[\bar{X}_{n-1}] < \dots < VaR_q[X_1]$.
- Let $k \geq 1$ and let X_1, \dots, X_{2^k} be i.i.d. risks such that $X_1 = f(Y_1)$, where $f \in \overline{M}$ and $Y_1 \sim S_1(\sigma, 0, 0)$ is a symmetric Cauchy r.v. Then $VaR_q[\bar{X}_{2^k}] < VaR_q[\bar{X}_{2^{k-1}}] < \dots < VaR_q[X_1]$.

Proposition 2 (Proschan 1965) *Let X_1 and X_2 be i.i.d risks such that $X_1 = f(Y_1)$, where $f \in \underline{M}$ and $Y_1 \sim S_1(\sigma, 0, 0)$ is a symmetric Cauchy r.v. Then $VaR_q[X_1] < VaR_q[\bar{X}_2]$.*

Ibragimov (2004a,b, 2005) showed that the following analogues of the results in Proschan (1965) for distributions from the classes $\overline{\mathcal{CSLC}}$ and $\underline{\mathcal{CS}}$ hold.

Proposition 3 *Let X_1, \dots, X_n , $n \geq 2$, be i.i.d risks such that $X_1 \sim \overline{\mathcal{CSLC}}$ or $X_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (1, 2]$, $\sigma > 0$, $\beta \in [-1, 1]$. Then $VaR_q[\bar{X}_n] < VaR_q[\bar{X}_{n-1}] < VaR_q[X_1]$.*

Proposition 4 *Let X_1, \dots, X_n , $n \geq 2$, be i.i.d risks such that $X_1 \sim \underline{\mathcal{CS}}$ or $X_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1)$, $\sigma > 0$, $\beta \in [-1, 1]$. Then $VaR_q[\bar{X}_n] > VaR_q[\bar{X}_{n-1}] > VaR_q[X_1]$.*

Ibragimov & Walden (2006) showed that analogues of Propositions 3 and 4 hold as well for bounded risks.

According to Theorem 1 below, the stylized facts on diversification being preferable (as formalized by Propositions 1 and 3) continue to hold for nonlinear transformations of heavy-tailed risks if expectations of these transformations are finite.

Theorem 1 *The following conclusions hold.*

- Let $Y_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in [1, 2]$, $\sigma > 0$, $\beta \in [-1, 1]$. Further, let X_1 and X_2 be i.i.d risks such that $X_1 = f(Y_1)$, where $f \in \overline{M}$; $f \in \overline{M}'$ for $\alpha = 1$. Then $VaR_q[\bar{X}_2] < VaR_q[X_1]$.
- Let $k \geq 1$ and let X_1, \dots, X_{2^k} be i.i.d risks such that $X_1 \sim \overline{\mathcal{C}TSLC}$. Then $VaR_q[\bar{X}_{2^k}] < VaR_q[\bar{X}_{2^{k-1}}] < \dots < VaR_q[X_1]$.

¹¹That is, in the case of an absolutely continuous risk Z , $P(Z > VaR_q[Z]) = q$.

¹²In what follows, we interpret the *positive* values of Z as a risk holder's losses. This interpretation of losses follows that in Embrechts et al. (2002) and is in contrast to Artzner et al. (1999) who interpret *negative* values of risks as losses.

As demonstrated by Theorem 2 below, the stylized fact on diversification being preferable is reversed for nonlinearly transformed thick-tailed risks with infinite expectations.

Theorem 2 *Let $Y_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1]$, $\sigma > 0$, $\beta \in [-1, 1]$. Further, let X_1 and X_2 are i.i.d risks such that $X_1 = f(Y_1)$, where $f \in \underline{M}$; $f \in \underline{M}'$ for $\alpha = 1$. Then $VaR_q[X_1] < VaR_q[\overline{X}_2]$.*

Remark 1 *Suppose that the r.v.'s X_1, \dots, X_n have a Cauchy-type distribution $S_1(\sigma, \beta, 0)$ which is exactly at the boundary between the nonlinear transformations of heavy-tailed risks in Theorems 1 and 2. Similar to Ibragimov (2004a,b, 2005), one can observe that, in such a case, the value at risk $VaR_q[\overline{X}_n]$ is the same for all $n \geq 1$. Consequently, in this setting, diversification of a portfolio has no effect on riskiness of its return. Using this fact, one can show, similar to the proof of Theorem 1 that the theorem continues to hold for convolutions of distributions in the class \overline{CTSLC} with symmetric Cauchy distributions $S_1(\sigma, 0, 0)$.*

Remark 2 *In order to highlight the main ideas and concepts discussed in the paper, Theorems 1 and 2 and the results in subsequent sections are formulated in the framework of independent risks that represent “the worst case scenario” for diversification failure. However, combining the methods of proofs of the results in this paper with those in Ibragimov (2004a,b, 2005), Ibragimov & Walden (2006), one can obtain analogues of the results for transformations of wide classes of dependent risks, including those with α -symmetric distributions.*

4.2 (Non)coherency of VaR for nonlinear transformations of thick-tailed risks

This section of the paper discusses implications of the results in Section 4.1 for (non)coherency properties of the value at risk.

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

- A1. (Monotonicity) $\mathcal{R}(X) \geq \mathcal{R}(Y)$ for all $X, Y \in \mathcal{X}$ such that $Y \leq X$ (a.s.), that is, $P(X \leq Y) = 1$.
- 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 $\lambda \geq 0$.
- A4. (Subadditivity) $\mathcal{R}(X + Y) \leq \mathcal{R}(X) + \mathcal{R}(Y)$ for all $X, Y \in \mathcal{X}$.

The above axioms are natural conditions to be imposed on measures of risk in the setting where positive values of r.v.'s $X \in \mathcal{X}$ represent losses of a risk holder.¹³ For instance, subadditivity property is important, among others, from the regulatory point of view because if a firm were forced to meet the requirement of extra capital which is not subadditive, it might be motivated to break up into several separately incorporated affiliates

¹³This interpretation of losses follows that in Embrechts et al. (2002) and is in contrast to Artzner et al. (1999) who interpret *negative* values of risks in \mathcal{X} as losses.

(see the discussion in Artzner et. al., 1999). In addition to that, the properties A1-A4 are important because, as follows from Ch. 10 in Huber (1981) (see also Artzner et al. 1999), in the case of a finite Ω , a risk measure \mathcal{R} is coherent (that is, it satisfies A1-A4) if and only if it is representable as $\mathcal{R}(X) = \sup_{Q \in \mathcal{P}} E_Q(X)$, where \mathcal{P} is some set of probability measures on Ω and, for $Q \in \mathcal{P}$, E_Q denotes the expectation with respect to Q . In other words, the risk measure \mathcal{R} is the worst result of computing the expected loss $E_Q(X)$ over a set \mathcal{P} of “generalized scenarios” (probability measures) Q . A similar representation holds as well in the case of an arbitrary Ω and the space $\mathcal{X} = L^\infty(\Omega, \mathfrak{S}, P)$ of bounded r.v.’s (see Fölmer & Schied (2002); moreover, as discussed in Frittelli & Gianin (2002), by duality theory, the convexity axiom A5 alone implies analogues of such characterizations for an arbitrary Ω and the space $\mathcal{X} = L_p(\Omega, \mathfrak{S}, P)$, $p \geq 1$, of r.v.’s X with a finite p -th moment $E|X|^p < \infty$.

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 et al. (1999) and Embrechts et al. (2002), in general, it fails to satisfy the subadditivity property A4, in particular, for certain Pareto distributions (Examples 6 and 7 in Embrechts et al. 2002).

As discussed in Ibragimov (2004a,b, 2005), from Proposition 3 it follows that the value at risk satisfies subadditivity axiom A4 and is, thus, a coherent measure of risk in the world of not extremely heavy-tailed risks from the class $\overline{\mathcal{C}\mathcal{T}\mathcal{S}\mathcal{L}\mathcal{C}}$. Proposition 4, on the other hand, implies that axiom A2 is *always* violated for risks with extremely heavy-tailed distributions from the class $\underline{\mathcal{C}\mathcal{S}}$. Thus, the value at risk is not a coherent risk measure in the class $\underline{\mathcal{C}\mathcal{S}}$ of extremely long-tailed distributions.

From Theorem 1 it follows that the value at risk is a coherent measure of risks in the world of nonlinear transformations $f \in \overline{\mathcal{M}}$ of heavy-tailed risks with finite expectations.

Corollary 1 *Let X_1 and X_2 be i.i.d risks such that $X_1 \sim \overline{\mathcal{C}\mathcal{T}\mathcal{S}\mathcal{L}\mathcal{C}}$, or $X_1 = f(Y_1)$, where $Y_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in [1, 2]$, and $f \in \overline{\mathcal{M}}$; $f \in \overline{\mathcal{M}'}$ for $\alpha = 1$. For all $q \in (0, 1/2)$, one has $VaR_q(X_1) + VaR_q(X_2) > VaR_q(X_1 + X_2)$. That is, subadditivity axiom A4 holds for VaR and it is a coherent measure of risk for distributions in the class $\overline{\mathcal{C}\mathcal{T}\mathcal{S}\mathcal{L}\mathcal{C}}$ and for transformations $f(Y_i)$ of i.i.d. risks $Y_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in [1, 2]$, where $f \in \overline{\mathcal{M}}$; $f \in \overline{\mathcal{M}'}$ for $\alpha = 1$.*

Theorem 2 implies that coherency of the value at risk is always violated for nonlinear transformations $f \in \underline{\mathcal{M}}$ of heavy-tailed risks for which the first moments are infinite.

Corollary 2 *Let X_1 and X_2 be i.i.d risks such that $X_1 = f(Y_1)$, where $Y_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1]$, and $f \in \underline{\mathcal{M}}$; $f \in \underline{\mathcal{M}'}$ for $\alpha = 1$. For all $q \in (0, 1/2)$ one has $VaR_q(X_1) + VaR_q(X_2) < VaR_q(X_1 + X_2)$. That is, subadditivity axiom A4 is violated for VaR and it is not a coherent measure of risk for transformations $f(Y_i)$ of i.i.d. risks $Y_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1]$, where $f \in \underline{\mathcal{M}}$; $f \in \underline{\mathcal{M}'}$ for $\alpha = 1$.*

Remark 3 *As indicated in the introduction, the reversals vs. robustness results similar to those given by Theorems 1 and 2 and Corollaries 1-4 hold as well for all the economic models considered in Ibragimov (2004a, 2005). These results may be proved similarly to those in Ibragimov (2004a, 2005), with the help of distributional comparisons for transformations of heavy-tailed r.v.’s implied by Theorems 1 and 2.*

4.3 Diversification and value at risk under deviations from power laws

As follows from the discussion in the introduction, the classes of nonlinear transformations of heavy-tailed risks contain many models with distributions exhibiting “local”, “moderate” and “global” departures from power laws (1)-(2) in form (6)-(11). This section of the paper presents implications of the results in Sections 4.1 and 4.2 for such deviations from Pareto distributions. We formulate the results for the functions V and W introduced in Section 3 with the value of x_0 being sufficiently large for V and W to be well-defined, odd and increasing on \mathbf{R} and strictly convex or concave on $[0, \infty)$, as discussed in that section.

According to Corollary 3, portfolio diversification decreases its riskiness in the world of risks whose distributions are even slightly thinner than those of Cauchy r.v.’s, as modelled, by, e.g., distributions with tails in form (9) with an arbitrary iteration of the logarithm in the denominator. This corollary also implies that the stylized fact on diversification being preferable also holds for distributions satisfying (7) with $\alpha > 1$ as well as for distributions exhibiting departures from power laws in form (11) with additional exponential factors.

Corollary 3 *Let V and W be as in Section 3 and let the parameters of the function V be such that $\gamma_0 < 1$ or $\gamma_0 = 1$, $\gamma_1 = 0$, ..., $\gamma_{k-1} = 0$, $\gamma_k < 0$ for some $k \in \{1, 2, \dots, m\}$. Further, let x_0 in the definition of V and W be sufficiently large so that the functions are well-defined, odd and increasing on \mathbf{R} and are strictly concave on $[0, \infty)$. If X_1 and X_2 are i.i.d risks such that $X_1 = V(Y_1)$ or $X_1 = W(Y_1)$, where $Y_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in [1, 2]$, $\sigma > 0$, $\beta \in [-1, 1]$, then the conclusions of Theorem 1 and Corollary 1 hold.*

Corollary 4 shows that riskiness of a portfolio increases with diversification for risks with distributions whose tails are even slightly heavier than those of Cauchy r.v.’s. This is the case, in particular, for distributions with tails in form (10) with an arbitrary iteration of the logarithm in the numerator. According to the corollary, the failure of diversification is also exhibited by distributions satisfying (8) with $\alpha < 1$.

Corollary 4 *Let V be as in Section 3 and let $\gamma_0 > 1$ or $\gamma_0 = 1$, $\gamma_1 = 0$, ..., $\gamma_{k-1} = 0$, $\gamma_k > 0$ for some $k \in \{1, 2, \dots, m\}$. Further, let x_0 in the definition of V be sufficiently large so that the function is well-defined, odd and increasing on \mathbf{R} and is strictly convex on $[0, \infty)$. If X_1 and X_2 are i.i.d risks such that $X_1 = V(Y_1)$, where $Y_1 \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1]$, $\sigma > 0$, $\beta \in [-1, 1]$, then the conclusions of Theorem 2 and Corollary 2 hold.*

4.4 Safety-first approach to portfolio selection under heavy-tailedness

The results on the effects of portfolio diversification on its value at risk discussed in the previous sections can be equivalently casted in the framework of Roy (1952)’s safety-first approach to portfolio selection. Given n risks X_1, \dots, X_n , the safety-first approach to portfolio choice consists in minimizing the probability $P\left(\sum_{i=1}^n w_i X_i > z\right)$ of going above a certain target or a disaster level $z > 0$ over the portfolio weights w_i :

$$\min P\left(\sum_{i=1}^n w_i X_i > z\right) \quad \text{s.t.} \quad \sum_{i=1}^n w_i = 1, \quad w_i \geq 0. \quad (17)$$

For instance, let $n = 2^k$, $k \geq 1$, and let the i.i.d. risks X_i , $i = 1, \dots, n$, be such that $X_1 \sim \overline{CTSLC}$. From Theorem 1 it follows that, for any value of the disaster level $z > 0$, the optimal portfolio of risks X_i in safety-first portfolio choice problem (17) is that with equal weights $(w_1, \dots, w_n) = (\frac{1}{n}, \dots, \frac{1}{n})$. Similarly, let X_1 and X_2 be nonlinear transformations $X_i = f(Y_i)$ of i.i.d. stable r.v.'s $Y_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in [1, 2]$, where $f \in \overline{M}$; $f \in \overline{M}'$ for $\alpha = 1$. In particular, one can take here the function f to be the transformation V defined in Section 3, with the parameters satisfying $\gamma_0 < 1$ or $\gamma_0 = 1$, $\gamma_1 = 0, \dots, \gamma_{k-1} = 0$, $\gamma_k < 0$ for some $k \in \{1, 2, \dots, m\}$. Alternatively, one can take f to be W defined in the same section. From Theorem 1 and Corollary 3 it follows that, for any $z > 0$, the probability of disaster $P(w_1 X_1 + w_2 X_2 > z)$ is minimized for equal weights $(w_1, w_2) = (\frac{1}{2}, \frac{1}{2})$.

Let now X_1 and X_2 be nonlinear transformations $X_i = f(Y_i)$, of i.i.d. stable r.v.'s $Y_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1]$, where $f \in \underline{M}$; $f \in \overline{M}'$ for $\alpha = 1$. In particular, one can take here the function f to be transformation V defined in Section 3 with $\gamma_0 > 1$ or with $\gamma_0 = 1$, $\gamma_1 = 0, \dots, \gamma_{k-1} = 0$, $\gamma_k > 0$ for some $k \in \{1, 2, \dots, m\}$. In contrast to the above, from Theorem 2 and Corollary 4 it follows that, for any $z > 0$, the probability of disaster $P(w_1 X_1 + w_2 X_2 > z)$ is minimized for the choice $(w_1, w_2) = (1, 0)$ or $(w_1, w_2) = (0, 1)$, that is, for portfolios consisting of only one risk X_1 or X_2 .

According to the above results, regardless of the disaster level z , the safety-first optimal portfolio for symmetric nonlinear transformations of heavy-tailed risks is the one with equal weights or the one consisting of only one risk. In particular, this is the case for nonlinear transformations of stable risks that exhibit tail behavior (12) similar to that of chi-square distributions. As demonstrated by the following Propositions 5 and 6 that follow from the results obtained by Szekely & Bakirov (2003), the situation is different for chi-square risks concentrated on the right semi-axis. These propositions show that, in the world of risks with chi-square distributions, complete diversification is preferable in the safety-first framework for large values of the disaster level, namely for $z > 2$, and is disadvantageous for small values of z , namely for $z < 1$. That is, in such settings, the optimal safety-first portfolios either have equal weights or consist of only one risk, similar to the framework with nonlinear functions of stable r.v.'s considered in the previous sections. However, in the case of moderate disaster levels between one and two, the optimal portfolio consists of two risks and, furthermore, does not have equal weights at them.

Let $\{\mathcal{N}_t\}_{t=1}^\infty$ be a sequence of i.i.d. standard normal r.v.'s and let, for $d \geq 1$, $\chi^2(d) \sim \sum_{t=1}^d \mathcal{N}_t^2$ be a chi-square r.v. with d degrees of freedom. Proposition 5, which is a consequence of the results in Szekely & Bakirov (2003), provides solutions to the safety-first portfolio selection problem in the case of two chi-square risks.

Proposition 5 *Let $X_t = \mathcal{N}_t^2$, $t = 1, 2$, be two i.i.d. risks with $\chi^2(1)$ distribution. The optimal solution to safety-first portfolio choice problem (17) with $n = 2$ is given by*

$$(w_1, w_2) = \begin{cases} (1, 0) & \text{for } 0 < z \leq 1, \\ (\tilde{w}_1, 1 - \tilde{w}_1) \notin \{(1, 0), (0, 1), (1/2, 1/2)\} & \text{for } 1 < z < 2, \\ (1/2, 1/2) & \text{for } z \geq 2. \end{cases} \quad (18)$$

For $z \in [1, 2]$ denote $s(z) = \inf_{w \in (0, 1/2)} P(w\mathcal{N}_1^2 + (1-w)\mathcal{N}_2^2 > z)$. As follows from Proposition 5, $s(z) > \max\{P(X_1 > z), P(X_2 > z)\} = \max\{P(\chi^2(1) > z), P(\frac{1}{2}\chi^2(2) > z)\}$, $z \in (1, 2)$. According to Proposition 2 in Szekely & Bakirov (2003), for every positive integer $d \geq 1$ there exists a unique intersection point $y(d)$ of $s(z)$

and $P(d^{-1}\chi^2(d) > z) = P\left(\frac{1}{d}\sum_{t=1}^d \mathcal{N}_t^2 > z\right)$. In addition, according to the same proposition, the intersection points $y(d)$ have the following properties: $y(2) = 2 > y(3) > y(4) > \dots > 1 = y(1)$; $y(d) \rightarrow 1$ as $d \rightarrow \infty$. The following proposition that follows from the results in Szekely & Bakirov (2003) provides solutions to safety-first portfolio choice problem (17) in the case of an arbitrary number $d \geq 3$ of chi-square risks X_i .

Proposition 6 *Let $n \geq 3$. The optimal solution to safety-first portfolio selection problem (17) is given by*

$$(w_1, w_2, \dots, w_n) = \begin{cases} (1, 0, \dots, 0) & \text{for } 0 < z \leq 1, \\ (\tilde{w}_1, 1 - \tilde{w}_1) \notin \{(1, 0), (0, 1), (1/2, 1/2)\} & \text{for } 1 < z < y(d), \\ (1/n, 1/n, \dots, 1/n) & \text{for } z \geq y(d). \end{cases} \quad (19)$$

5 Concluding remarks

The value of diversification is a corner stone of financial theory. Understanding when diversification succeeds and when it fails to provide additional value to investors is therefore important. We have provided a refined analysis of when failure of diversification occurs. Our results indicate that the Cauchy distribution is by all means the separating distribution. Distributions with “thinner” tails than Cauchy – no matter how slightly thinner – will always lead to diversification being valuable. On the contrary, distributions with “fatter” tails – no matter how slightly fatter – will always lead to failure of diversification. We believe that the detailed analysis in this paper is well-motivated by the importance of the issue.

Appendix - Proofs

Proof of Theorems 1 and 2. Let $\alpha_1 \in [1, 2]$, $\alpha_2 \in (0, 1]$, and let $f_1 \in \overline{M}$, $f_2 \in \underline{M}$. Suppose that $f_1 \in \overline{M}'$ if $\alpha_1 = 1$ and $f_2 \in \underline{M}'$ if $\alpha_2 = 1$. For $j = 1, 2$, let $Y_1^{(j)}$ and $Y_2^{(j)}$ be i.i.d. r.v.'s such that $Y_i^{(j)} \sim S_{\alpha_j}(\sigma, \beta, 0)$, $\sigma > 0$, $\beta \in [-1, 1]$, $i = 1, 2$, and let $X_i^{(j)} = f_j(Y_i^{(j)})$, $i = 1, 2$, $j = 1, 2$. As in the proof of Lemmas 2.7 and 2.8 in Proschan (1965), by the definition of the classes \overline{M} and \underline{M} we have that

$$|f_1((y_1 + y_2)/2)| \geq |(f_1(y_1) + f_1(y_2))/2|, \quad (20)$$

$$|f_2((y_1 + y_2)/2)| \leq |(f_2(y_1) + f_2(y_2))/2| \quad (21)$$

for all $y_1, y_2 \in \mathbf{R}$. In addition, inequality (20) is strict for $y_1 + y_2 \neq 0$ if $f_1 \in \overline{M}'$. Similarly, inequality (21) is strict for $y_1 + y_2 \neq 0$ if $f_2 \in \underline{M}'$. Since the functions $|f_j(x)|$, $j = 1, 2$, are increasing in $|x|$, we, therefore, get that

$$|f_1((y_1 + y_2)/2^{1/\alpha_1})| \geq |(f_1(y_1) + f_1(y_2))/2|, \quad (22)$$

$$|f_2((y_1 + y_2)/2^{1/\alpha_2})| \leq |(f_2(y_1) + f_2(y_2))/2|. \quad (23)$$

Obviously, inequality in (22) is strict for $y_1 + y_2 \neq 0$ if $\alpha_1 > 1$ and inequality in (23) is strict for $y_1 + y_2 \neq 0$ if $\alpha_2 < 1$. Since $2^{-1/\alpha_j}(Y_1^{(1)} + Y_2^{(1)}) \sim S_{\alpha_j}(\sigma, \beta, 0)$, $j = 1, 2$, and the functions f_j , $j = 1, 2$, are odd, this implies

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

$$\begin{aligned} VaR_q \left[\frac{X_1^{(1)} + X_2^{(1)}}{2} \right] &= VaR_q \left[\frac{f_1(Y_1^{(1)}) + f_1(Y_2^{(1)})}{2} \right] < \\ VaR_q \left[\frac{f_1(Y_1^{(1)} + Y_2^{(1)})}{2^{1/\alpha_1}} \right] &= VaR_q \left[f_1(Y_1^{(1)}) \right] = VaR_q \left[X_1^{(1)} \right] \end{aligned} \quad (24)$$

and

$$\begin{aligned} VaR_q \left[\frac{X_1^{(2)} + X_2^{(2)}}{2} \right] &= VaR_q \left[\frac{f_2(Y_1^{(2)}) + f_2(Y_2^{(2)})}{2} \right] > \\ VaR_q \left[\frac{f_2(Y_1^{(2)} + Y_2^{(2)})}{2^{1/\alpha_2}} \right] &= VaR_q \left[f_2(Y_1^{(2)}) \right] = VaR_q \left[X_1^{(2)} \right]. \end{aligned} \quad (25)$$

According to relation (24), the part of Theorem 1 for transformations of stable r.v.'s holds. Relation (25) shows that Theorem 2 is true.

Let now $n = 2^k$, $k \geq 2$, and let X_1, \dots, X_n be i.i.d. r.v.'s such that $X_1 \sim \overline{\mathcal{CTSLC}}$. By definition of the class $\overline{\mathcal{CTSLC}}$, there exist i.i.d. r.v.'s Y_{it} , $t = 0, 1, \dots, k$, $i = 1, \dots, n$, and functions $f_t \in \overline{\mathcal{M}}$, $t = 1, \dots, k$, such that $Y_{i0} \sim \mathcal{LC}$ and $Y_{it} \sim S_{\alpha_t}(\sigma_t, 0, 0)$, $\alpha_t \in (1, 2]$, $\sigma_t > 0$, $t = 1, \dots, k$, and $X_i = \theta Y_{i0} + f_1(Y_{i1}) + \dots + f_k(Y_{ik})$, $\theta \in \{0, 1\}$, $i = 1, \dots, n$. From (24) we have that for all $i = 1, \dots, n/2$ and $t = 1, \dots, k$,

$$VaR_q \left[\frac{f_t(Y_{it}) + f_t(Y_{n/2+i,t})}{2} \right] < VaR_q[f_t(Y_{it})]. \quad (26)$$

In addition, by Proposition 1,

$$VaR_q \left[\frac{Y_{i0} + Y_{n/2+i,0}}{2} \right] < VaR_q[f_t(Y_{i0})]. \quad (27)$$

According to Theorem 2.7.6 in Zolotarev (1986), p. 134, and Theorem 1.10 in Dharmadhikari & Joag-Dev (1988), p. 20, the densities of the r.v.'s Y_{it} , $t = 0, 1, \dots, k$, $i = 1, \dots, n$, are symmetric and unimodal. This implies, as it is not difficult to see, symmetry and unimodality of the densities of the r.v.'s $f_t(Y_{it})$, $f_t \in \overline{\mathcal{M}}$, $t = 1, \dots, k$, $i = 1, \dots, n$. By Theorem 1.6 in Dharmadhikari & Joag-Dev (1988), p. 13, we get, in turn, that the densities of the r.v.'s $(Y_{i0} + Y_{n/2+i,0})/2$ and $(f_t(Y_{it}) + f_t(Y_{n/2+i,t}))/2$, $t = 1, \dots, k$, $i = 1, \dots, n$, are symmetric and unimodal.

From Lemma in Birnbaum (1948) and its proof it follows that if ξ_1, ξ_2 and η_1, η_2 are independent absolutely continuous symmetric unimodal r.v.'s such that, for $j = 1, 2$, and all $q \in (0, 1/2)$, $VaR_q[\xi_j] < VaR_q[\eta_j]$, then $VaR_q[\xi_1 + \xi_2] < VaR_q[\eta_1 + \eta_2]$, $q \in (0, 1/2)$. This, together with (26) and (27), implies by induction (see also Theorem 1 in Birnbaum, 1948, and Theorem 2.C.3 in Dharmadhikari and Joag-Dev, 1988) that

$$\begin{aligned} VaR_q[\overline{X}_n] &= VaR_q \left[\frac{1}{n} \sum_{i=1}^{n/2} \left\{ \theta(Y_{i0} + Y_{n/2+i,0}) + (f_1(Y_{i1}) + f_1(Y_{n/2+i,1})) \dots + (f_k(Y_{ik}) + f_k(Y_{n/2+i,k})) \right\} \right] < \\ VaR_q \left[\frac{2}{n} \sum_{i=1}^{n/2} \left\{ \theta Y_{i0} + f_1(Y_{i1}) + \dots + f_k(Y_{ik}) \right\} \right] &= VaR_q \left[\overline{X}_{n/2} \right]. \end{aligned}$$

This completes the proof of Theorem 1 for distributions from the class $\overline{\mathcal{CTSLC}}$.

Proof of Corollaries 1-4. Corollaries 1 and 2 evidently follow from Theorems 1 and 4 and positive homogeneity of the value at risk (axiom A3 in Section 4.1). Corollary 3 follows from Theorem 1 since, under the assumptions of the corollary, the functions V and W belong to the class $f \in \overline{M}'$. Similarly, Corollary 3 is a consequence of Theorem 1 and the fact that, under its assumptions, the function V belongs to the class $f \in \underline{M}'$.

Proof of Proposition 5-6. The propositions evidently follow from Theorem 2 in Szekely & Bakirov (2003).

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. (2003), 'Cross-section regression with common shocks', *Cowles Foundation Discussion Paper 1428*.
- 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.
- Birnbaum, Z. W. (1948), 'On random variables with comparable peakedness', *Annals of Mathematical Statistics* **19**, 76–81.
- Cotter, D. & Dowd, K. (2002), 'Extreme spectral risk measures: an application to futures clearinghouse margin requirements', *Working paper, Smurfit School of Business, University College Dublin and Nottingham University Business School*.
- 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*, Boston: Academic Press.
- Embrechts, P., Klupperberg, 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.
- Fama, E. (1965a), '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.
- 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.
- Huber, P. (1981), *Robust statistics*, New York: Wiley.
- Ibragimov, R. (2004*a*), ‘On the robustness of economic models to heavy-tailedness assumptions’, *Mimeo, Yale University*. Available at <http://post.economics.harvard.edu/faculty/ibragimov/Papers/HeavyTails.pdf>.
- Ibragimov, R. (2004*b*), ‘Portfolio diversification and value at risk under thick-tailedness’, *Harvard University Research Discussion Paper 2086*. Available at <http://post.economics.harvard.edu/hier/2005papers/HIER2086.pdf>.
- Ibragimov, R. (2005), *New majorization theory in economics and martingale convergence results in econometrics*. Ph.D. dissertation, Yale University.
- Ibragimov, R. & Phillips, P. C. B. (2004), ‘Regression asymptotics using martingale convergence methods’, *Cowles Foundation Discussion Paper 1473*. Available at <http://cowles.econ.yale.edu/P/cd/d14b/d1473.pdf>.
- Ibragimov, R. & Walden, J. (2006), ‘The limits of diversification when losses may be large’, *Harvard University Research Discussion Paper 2104*. Available at <http://post.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.
- Loretan, M. & Phillips, P. C. B. (1994), ‘Testing the covariance stationarity of heavy-tailed time series’, *Journal of Empirical Finance* **1**, 211–248.
- 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.
- Meerschaert, M. M. & Scheffler, H.-P. (2000), Portfolio modelling with heavy tailed random vectors, *in* S. T. Rachev, ed., 'Handbook of Heavy-Tailed Distributions in Finance', Elsevier, North-Holland.
- Phillips, P. C. B. (1988), 'Regression theory for near-integrated time series', *Econometrica* **56**, 1021–1043.
- Phillips, P. C. B. & Magdalinos, T. (2004), 'Limit theory for moderate deviations from a unit root', *Cowles Foundation Discussion Paper 1471*. Available at <http://cowles.econ.yale.edu/P/cd/d14b/d1471.pdf>.
- 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*.
- Roy, A. D. (1952), 'Safety-first and the holding of assets', *Econometrica* **20**, 431–449.
- Samuelson, P. A. (1967), 'Efficient portfolio selection for Pareto-Levy investments', *Journal of Financial and Quantitative Analysis* **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.
- Schwarz, M. (2000), 'Decision making under extreme uncertainty', *Working paper, Harvard University*. Available at <http://rwj.berkeley.edu/schwarz/publications/>.
- Silverberg, G. & Verspagen, B. (2004), 'The size distribution of innovations revisited: an application of extreme value statistics to citation and value measures of patent significance', *Working paper, Maastricht University* **36**, 1703–1706.
- Szekely, G. J. & Bakirov, N. K. (2003), 'Extremal probabilities for Gaussian quadratic forms', *Probability Theory and Related Fields* **126**, 184–202.
- Tasche, D. (2002), 'Expected shortfall and beyond', *Journal of Banking and Finance* **26**, 1519–1533.
- Uchaikin, V. V. & Zolotarev, V. M. (1999), *Chance and stability. Stable distributions and their applications*, Utrecht: VSP.
- Zolotarev, V. M. (1986), *One-dimensional stable distributions*, Providence: American Mathematical Society.