HIER

Harvard Institute of Economic Research

Discussion Paper Number 2104

The Limits of Diversification When Losses May Be Large

by

Rustam Ibragimov
and
Johan Walden

February 2006

HARVARD UNIVERSITY
Cambridge, Massachusetts

THE LIMITS OF DIVERSIFICATION WHEN LOSSES MAY BE LARGE[1]

Rustam Ibragimov

Department of Economics, Harvard University

Johan Walden

Haas School of Business, University of California at Berkeley

January 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

[1]The authors are grateful to Gary Chamberlain and Oliver Hart for helpful comments and for posing a question about the stylized facts on diversification under boundedness assumptions that is answered in this paper. We also thank Donald Andrews, Donald Brown, Xavier Gabaix, Tilmann Gneiting, Dwight Jaffee, Peter Phillips, Jacob Sagi, Herbert Scarf 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. We are grateful to Jonathan Ingersoll and David Pollard for useful discussions on extremal problems for expectations of functions of random variables with fixed moments and for proposing alternative solutions to the problem of determining the extrema of expectations $Eu(X)$ for concave functions $u : [a, b] \rightarrow \mathbf{R}$ over r.v.'s X on $[a, b]$ with fixed mean $EX = c$ discussed in Remark 10.

ABSTRACT

Recent results in value at risk analysis show that, for extremely heavy-tailed risks with unbounded distribution support, diversification may increase value at risk, and that, generally, it is difficult to construct an appropriate risk measure for such distributions. We further analyze the limitations of diversification for heavy-tailed risks. We provide additional insight in two ways. First, we show that similar nondiversification results are valid for a large class of risks with bounded support, as long as the risks are concentrated on a sufficiently large interval. The required length of the support depends on the number of risks available and on the degree of heavy-tailedness. Second, we relate the value at risk approach to more general risk frameworks. We argue that in financial markets where the number of assets is limited compared with the (bounded) distributional support of the risks, unbounded heavy-tailed risks may provide a reasonable approximation. We suggest that this type of analysis may have a role in explaining various types of market failures in markets for assets with possibly large negative outcomes.

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

JEL Classification: G11

1 Introduction

1.1 Background

Recently, Ibragimov (2004a,b, 2005) developed a unified approach to analyzing portfolios of risks with heavy-tailed distributions, using new majorization theory for linear combinations of thick-tailed random variables (r.v.'s).[2] Those works showed that the stylized fact of portfolio diversification always being preferable is reversed for extremely heavy-tailed risks, with infinite first moments and unbounded distribution support (see Proposition 1 in Subsection 3.1 of this paper). Specifically, for such distributions, the value at risk (VaR) is a strictly increasing function in the degree of diversification.

Value at risk and the closely related safety-first principle are frequently used in models in economics, finance and risk management,[3] providing alternatives to the traditional expected utility framework. For extremely heavy-tailed distributions the expected utility framework is not readily available, since it typically involves assumptions on the existence of moments for the risks in consideration. The safety-first and VaR approaches to portfolio selection have thus, in many regards, been the only ones available in the presence of extreme thick-tailedness.[4]

This has also meant that the relationship between traditional diversification results that are based on expected utility and thin tails, and the non-diversification results that are based on VaR and thick tails have been somewhat unclear. Specifically, one may ask whether non-diversification strictly depends on the asymptotic behavior of the distributions far out in the tails. If this is the case, the theoretical results may have few applications in a world, in which distributions may have bounded support. Furthermore, one may ask whether the VaR non-diversification results are due to imperfections of VaR as a risk measure, and how they relate to expected utility based risk measures, e.g. stochastic dominance. In this paper, we suggest that the non-diversification results may be robust to such objections.

1.2 Main contributions of paper

The main results of this paper are provided in Theorems 1-4 and Table 1. First, we demonstrate that the above VaR results continue to hold for a wide class of *bounded* risks[5] concentrated on a sufficiently large interval (Theorem 1). We also study how the length of distributional support needed for our results to hold depends on the number of risks in the portfolio and the degree of heavy-tailedness of the unbounded distributions.

[2]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 the majorization ordering and functions that preserve it, has found applications in disciplines ranging from statistics, probability theory and economics to mathematical genetics, linear algebra and geometry (see the discussion in Ibragimov, 2004a, b).

[3]See, e.g., Fabozzi, Focardi and Kolm (2006) for a review of Roy's (1952) safety-first approaches to portfolio selection, value at risk and other measures of risk.

[4]One should note here that several recent papers (see among others, Acerbi and Tasche, 2002, and Tasche, 2002) 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, and Cotter and Dowd, 2002) that generalize the expected shortfall.

[5]We will, somewhat contradictory, refer to distributions of such risks as *bounded heavy-tailed distributions* as opposed to the standard (unbounded) heavy-tailed distributions.

Second, we relate our results to the expected utility framework. For risks with unbounded heavy-tailed distributions, we provide a natural generalization of the second order stochastic dominance concept, originally introduced in Rothschild and Stiglitz (1970). We provide a rigorous motivation for that diversification increases risk for such distributions (Theorem 2). Furthermore, we relate our results on bounded risks to the traditional results on diversification. With bounded supports, diversification will always be preferable in an expected utility setting[6], contrary to our value at risk results. We show that the traditional results crucially depend on the tail properties of the expected utility function and that if investors' utility function at any point in the domain of large negative outcomes becomes convex,[7] then our non-diversification results may continue to hold (Theorem 3). This provides additional support for our view that the theory of unbounded heavy-tailed distributions may provide a good approximation for markets with a limited number of bounded heavy-tailed risks.

Third, we provide numerical results that show when not to diversify (Table 1), depending on the types of distributions, the length of distributional support and the number of risks at hand. In the non-diversification region, the implications for asset pricing may be large. It will be difficult to create risk sharing, idiosyncratic risk will matter, and risk premia may be high. We suggest that this could explain puzzling properties of risky assets for which losses may be large, e.g., catastrophe insurance and the under-diversification puzzle. This is a natural future application of the results in this paper.

Fourth, we obtain extensions of the above results for a wide class of dependent risks. We show that Theorem 1 continues to hold for convolutions of dependent risks with joint truncated $\alpha-$symmetric distributions and their analogues with non-identical marginals (Theorem 4).[8,9]

To summarize, the stylized facts on portfolio diversification are not robust to distributional assumptions involving truncated extremely heavy-tailed distributions. In our view, this is important since it demonstrates that the "unpleasant" properties of value at risk as a risk measure under heavy-tailedness does not arise from properties of the extreme tails of the distributions. On the contrary, the conclusion that diversification is always to be preferred depends on properties of the utility function for extreme outcomes, or on the number of assets being very large. In the real world, distributions may be finitely supported, the risk averse expected utility specification of investor behavior may break down outside reasonable domains and the number of assets is finite. Which model approximation is most appropriate must depend on the situation at hand.

[6]As originally shown in a general setting in Samuelson (1967).

[7]Convexity of utility functions in the loss domain being one of the key foundations of Prospect theory (Kahneman and Tversky 1979). It also effectively arises if there is limited liability.

[8]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 $\exp\big[-\lambda\big(\sum_{i=1}^{n} t_i^2\big)^{\beta/2}\big]$, $0 < \beta \leq 2$. Obviously, spherically symmetric stable distributions are particular examples of $\alpha-$symmetric distributions with $\alpha = 2$ (that is, of spherical distributions) and $\phi(x) = \exp(-x^{\beta})$.

[9]The class of $\alpha-$symmetric distributions is very wide and includes, in particular, 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 (see, among others, Praetz, 1972, Blattberg and Gonedes, 1974, and Glasserman, Heidelberger and Shahabuddin, 2002). Moreover, the class of $\alpha-$symmetric distributions includes a wide class of convolutions of models with common shocks affecting all risks (such as macroeconomic or political ones, see Andrews, 2003) which are of great importance in economics and finance.

1.3 Extensions of 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 *truncated* stable distributions and their convolutions. More precisely, the results of the paper for the bounded analogues of extremely heavy-tailed densities are first presented and proved using the framework of convolutions of truncated stable distributions with characteristic exponents α less than one.

As indicated above, in Section 4 we show, however, that the results on nondiversification continue to hold for a wide class of multivariate distributions for which marginals are bounded and dependent (Theorem 4). These distributions can be non-identical and, in addition, can be truncations of distributions with finite variances (unlike bounded analogues of stable distributions and their convolutions with infinite second moments). As indicated before, according to these extensions, all the results in the paper continue to hold for convolutions of truncated $\alpha-$symmetric distributions and their analogues with non-identical one-dimensional marginals. Similar to the framework based on stable distributions, the stylized facts on portfolio diversification are reversed in the case of convolutions of truncated $\alpha-$symmetric distributions with $\alpha < 1$ that have extremely heavy-tailed marginal distributions with infinite means.

Also, the results in the paper are available for the case of skewed distributions (see Remark 6), including truncated skewed stable distributions (such as, for instance, extremely heavy-tailed Lévy distributions with $\alpha = 1/2$ concentrated on the positive semi-axis) and, according to the extensions discussed above, truncated $\alpha-$symmetric distributions with skewed marginals.[10] Therefore, this paper, in fact, succeeds in the unification of the robustness of safety-first portfolio selection theory to all the main distributional properties: boundedness, heavy-tailedness, dependence, skewness and the case of non-identical one-dimensional distributions.[11]

1.4 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 thick-tailed phenomena. This stream of literature goes back to Mandelbrot (1963) (see also the papers in Mandelbrot, 1997, and Fama, 1965), 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 thick-tailed distribution, then[12]

$$P\Big(|X| > x\Big) \sim x^{-\alpha}. \quad (1)$$

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 $p \geq \alpha$.

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 and Phillips, 1994, Meerschaert and Scheffler, 2000, Gabaix, Gopikrishnan, Plerou and Stanley, 2003, and references therein). Motivated by these empirical findings, a number

[10] 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.

[11] According to well-established parlance in the many scientific literatures, robustness is understood to mean insensitivity to deviations from distributional assumptions. In this paper, the use of the term "robustness" accords with this tradition.

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

of studies in financial economics have focused on portfolio and value-at-risk modeling with heavy-tailed returns (see, e.g., the reviews in Duffie and Pan, 1997, Uchaikin and Zolotarev, 1999, Ch. 17, and Glasserman, Heidelberger and Shahabuddin, 2002). Several authors considered problems of statistical inference for data from thick-tailed populations (see Loretan and Phillips, 1994, the papers in Adler, Feldman and Taqqu, 1998, 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 and de Vries, 1991),

$2 < \alpha < 4$ (Loretan and Phillips, 1994),

$1.5 < \alpha < 2$ (McCulloch, 1996, 1997),

$0.9 < \alpha < 2$ (Rachev and Mittnik, 2000).

Recent studies (see Gabaix, Gopikrishnan, Plerou and Stanley, 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 on financial markets obey the power laws (1) with $\alpha \approx 1.5$ 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 demonstrates that the above power laws 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 $\alpha \approx 1$. Power laws (1) with $\alpha \approx 1$ (Zipf laws) have also been found to hold for firm sizes (see Axtell, 2001) and city sizes (see Gabaix, 1999a, b for the discussion and explanations of the Zipf law for cities).

De Vany and Walls (2004) presented evidence that stable distributions with tail indices $1 < \alpha < 2$ provide a good model for distributions of profits in motion pictures. One should also note that some studies also 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 and Mittnik, 2000). Furthermore, Scherer, Harhoff and Kukies (2000) and Silverberg and Verspagen (2004) report the tail indices ξ to be considerably less than one for financial returns from technological innovations.

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 truncations of 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.

1.5 Organization of paper

The paper is organized as follows: Section 2 contains notations and definitions of classes of heavy-tailed distributions used throughout the paper. It also reviews their properties. In Section 3, we present the main results of the paper on the effects on riskiness of diversification of bounded risks. We also relate our results to the expected utility/stochastic dominance framework. We provide tables for when it is optimal not to diversify and finally discuss some puzzles in financial markets for which the limits of diversification may play a role. Section 4 discusses extensions of the results in the paper to the case of dependence, including convolutions of truncated α-symmetric and spherical distributions and models with common shocks. In Section 5, we make some concluding remarks.

2 Notation

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

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

$x \in \mathbf{R}$, 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. 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).[13] 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). 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.

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 $a_i \geq 0$, $i = 1, ..., n$,

[13] The 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.

$$\sum_{i=1}^{n} a_i X_i / \Big(\sum_{i=1}^{n} a_i^{\alpha}\Big)^{1/\alpha} \sim S_\alpha(\sigma, \beta, \mu). \tag{2}$$

For a detailed review of properties of stable distributions the reader is referred to, e.g., the monographs by Zolotarev (1986), Embrechts, Klupperberg and Mikosch (1997), Uchaikin and Zolotarev (1999), Rachev and Mittnik (2000) and Rachev, Menn and Fabozzi (2005).

For $0 \le r < 1$, we denote by $\mathcal{CS}(r)$ the class of distributions which are convolutions of symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with indices of stability $\alpha \in (r, 1)$ and $\sigma > 0$.[14] That is, $\mathcal{CS}(r)$ consists of distributions of r.v.'s X for which, with some $k \ge 1$, $X = Y_1 + ... + Y_k$, where Y_i, $i = 1, ..., k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (r, 1)$, $\sigma_i > 0$, $i = 1, ..., k$.

A linear combination of independent stable r.v.'s, each having 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 $\mathcal{CS}(r)$ of *convolutions* of symmetric stable distributions with *different* indices of stability $\alpha \in (r, 1)$ is wider than the class of *all* symmetric stable distributions $S_\alpha(\sigma, 0, 0)$ with $\alpha \in (r, 1)$ and $\sigma > 0$.

Evidently, the classes $\mathcal{CS}(r)$ are closed under convolutions. In addition, clearly, $\mathcal{CS}(r_1) \subset \mathcal{CS}(r_2)$ if $r_1 < r_2$. In what follows, we write $X \sim \mathcal{CS}(r)$ if the distribution of the r.v. X belongs to the class $X \sim \mathcal{CS}(r)$. The properties of stable distributions discussed above imply that the $p-$th absolute moments $E|X|^p$ of a r.v. $X \sim \mathcal{CS}(r)$, $r \in (0, 1)$, are finite if $p < r$. However, all the r.v.'s $X \sim \mathcal{CS}(r)$, $r \in (0, 1)$ have infinite means: $E|X| = \infty$.

Throughout the paper, given two r.v.'s X and Y, we write $X \overset{\mathrm{d}}{=} Y$ if the distributions of X and Y are the same. In what follows, for $z \in \mathbf{R}$, $sign(z)$ denotes the sign of z: $sign(z) = 1$ for $z > 0$, $sign(z) = -1$ for $z < 0$, and $sign(0) = 0$. In addition, $I(\cdot)$ stands for the indicator function.

We define the a-truncated version of a r.v.: $Y(a) = X$ if $|X| \le a$, $Y(a) = -a$ if $X < -a$ and $Y(a) = a$ if $X > a$. In other words, $Y(a) = a \cdot sign(X_i) + XI(|X| \le a)$.[15] We will also use the notation X^a instead of $Y(a)$ for the a-truncated version of X.

3 Main results: The limits of diversification

3.1 Non-diversification for risks with bounded support

Let $0 \le r < 1$. Following the framework of Roy's (1952) safety-first, given a r.v. (risk) Z, we will interested in analyzing the probability $P(Z > z)$ of going above a certain target or a disaster level $z > 0$.[16] Furthermore, given a

[14] $\mathcal{CS}$ is the abbreviation of "convolutions of stable."

[15] This definition of truncation moves probability mass to the edges of the distributions. As follows from the arguments the results in Section 3.1 continue to hold for the more commonly used truncations $XI(|X| \le a)$ which move probability mass to the center. However, this is not true for the results in Section 3.2.

[16] In what follows, we interpret the *positive* values of Z as a risk holder's losses. This interpretation of losses follows that in Embrechts, McNeil and Straumann (2002) and is in contrast to Artzner, Delbaen, Eber and Heath (1999) who interpret *negative* values of risks as losses.

loss probability $q \in (0, 1/2)$ and 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.[17]

Throughout this section, $X_1, X_2, ...$ is a sequence of i.i.d. risks with distributions from the class $\mathcal{CS}(r)$. For $a > 0$, denote by $Y_i(a)$ the a-truncated versions of X_i's. In what follows, $\mathbf{R}_+$ stands for $\mathbf{R}_+ = [0, \infty)$. Let $\mathcal{I}_n = \{w = (w_1, ..., w_n) \in \mathbf{R}_+^n : \sum_{i=1}^n w_i = 1\}$.

For $w \in \mathcal{I}_n$, denote by X_w the return on the portfolio of risks $X_1, ..., X_n$ with weights w: $X_w = \sum_{i=1}^n w_i X_i$. Similarly, in what follows, for $a > 0$ and $w \in \mathcal{I}_n$, $Y_w(a)$ stands for the return on the portfolio of bounded risks $Y_1(a), ..., Y_n(a)$ with weights w: $Y_w(a) = \sum_{i=1}^n w_i Y_i(a)$.

Evidently, the return on the portfolio of risks $X_1, ..., X_n$ with equal weights $\tilde{w}_n = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$ is given by the sample mean of X_i's: $X_{\tilde{w}_n} = \frac{1}{n}\sum_{i=1}^n X_i$. Similarly, $Y_{\tilde{w}_n}(a)$ is the sample mean of the risks $Y_i(a)$: $Y_{\tilde{w}_n}(a) = \frac{1}{n}\sum_{i=1}^n Y_i(a)$.

The problems faced by a holder of risks $X_1, ..., X_n$ or $Y_1(a), ..., Y_n(a)$ consist in minimizing, respectively, the disaster probabilities $P\Big(\sum_{i=1}^n w_i X_i > z\Big)$ or $P\Big(\sum_{i=1}^n w_i Y_i(a) > z\Big)$ over the portfolio weights $w \in \mathcal{I}_n$.

Let $w_{[1]} \geq ... \geq w_{[n]}$ denote the components of $w \in \mathcal{I}_n$ in decreasing order. Obviously, $w_{[1]} = 1$ implies that w is a permutation of the vector $(1, 0, ..., 0)$. E.g., according to the following proposition, in such a case, obviously, the portfolio with weights w consists of only one risk, and, thus, $X_{\tilde{w}_n}$ has the same distribution as X_1 and $Y_{\tilde{w}_n}(a)$ is distributed as $Y_1(a)$. In addition, for $w \in \mathcal{I}_n$, let $(w^{(1)}, w^{(2)}) = \big(max[0.5, w_{[1]}], min[0.5, 1 - w_{[1]}]\big)$.

According to the following result obtained in Ibragimov (2004a, b, 2005), the stylized facts that portfolio diversification is always preferable are violated for a wide class of *extremely heavy-tailed* risks with *unbounded* distribution support.[18] In such a setting, diversification of a portfolio of the risks increases the probability of going over a given disaster level.

Proposition 1 *(Ibragimov, 2004a, b, 2005). Let $w \in \mathcal{I}_n$ be a vector of weights with $w_{[1]} \neq 1$. Suppose that X_i, $i = 1, ..., n$, are i.i.d. risks such that $X_i \sim \mathcal{CS}(r)$, for some $r \in (0, 1)$, $i = 1, ..., n$. Then, for all $z > 0$, $P\Big(X_w > z\Big) > P\Big(w^{(1)} X_1 + w^{(2)} X_2 > z\Big) > P\Big(X_1 > z\Big)$.*

Remark 1 *If r.v.'s $X_1, ..., X_n$ have a symmetric Cauchy distribution $S_1(\sigma, 0, 0)$ which is exactly at the "upper" boundary of the classes $\mathcal{CS}(r)$, $r \in (0, 1)$, then the disaster probabilities $P\Big(X_w > z\Big)$ are the same for all $w \in \mathcal{I}_n$. Consequently, in such a case, diversification of a portfolio has no effect on its riskiness.*

Remark 2 *It is not difficult to see that Proposition 1 can be equivalently formulated as follows in the framework of the value at risk analysis for financial portfolios. Let $w \in \mathcal{I}_n$ be a vector of weights with $w_{[1]} \neq 1$. Suppose that*

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

[18]The result given by Proposition 1 is a part of Corollary 5.3 in Ibragimov, 2004a and of Theorem 4.2 in Ibragimov, 2004b since the vector $w = (w_1, w_2, w_3, ..., w_n)$ is majorized by (that is, has more diverse components than) the vector $(v_1, v_2, 0, ..., 0)$ which is, in turn, is majorized by the vector $(1, 0, 0, ..., 0)$.

X_i, $i = 1, ..., n$, are i.i.d. risks such that $X_i \sim \mathcal{CS}(r)$, for some $r \in (0, 1)$, $i = 1, ..., n$. Then, for all *loss probabilities $q \in (0, 1/2)$, the return X_w on the portfolio of risks $X_1, ..., X_n$ with weights w is strictly more risky (in terms of the value at risk) than the return $w^{(1)}X_1 + w^{(2)}X_2$ on the portfolio of two risks X_1 and X_2 with weights $w^{(1)}$ and $w^{(2)}$. In turn, the return $w^{(1)}X_1 + w^{(2)}X_2$ is more risky (in terms of the value at risk) than the return X_1 on the portfolio consisting of one risk. In other words, for* any *value of the loss probability $q \in (0, 1/2)$, the following inequalities hold: $VaR_q[X_w] > VaR_q[w^{(1)}X_1 + w^{(2)}X_2] > VaR_q[X_1]$.*

We now expand the analysis to risks with bounded support. A summary of the results we will provide is given in Figure 1. The traditional situation with i.i.d. risks is according to line A in the figure: diversification is always to be preferred, regardless of the number of risks. The other extreme is D, when diversification never will be preferred, as analyzed in Ibragimov (2004a, b, 2005). The intermediate cases are B and C, when diversification is suboptimal up to a certain number of risks (similar to D), but becomes preferable when enough assets are available and/or investors are VaR tolerant, (similar to A).

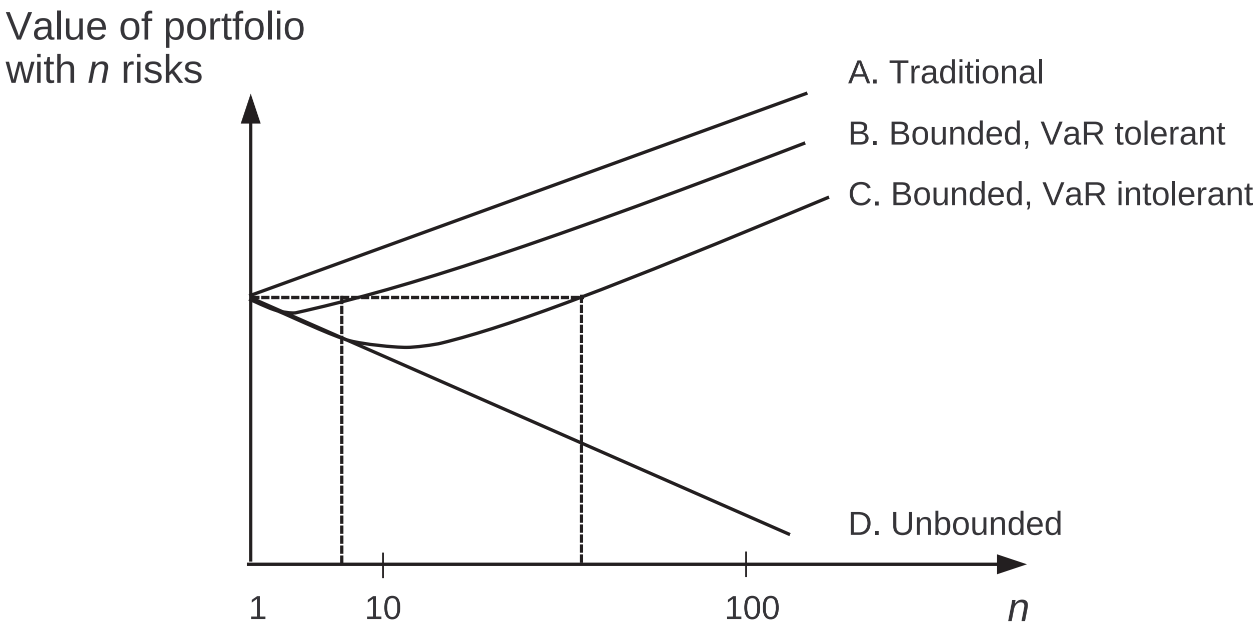

Figure 1: *Illustrative figure of value of diversification. A: Traditional situation ($\alpha > 1$). The value increases monotonically and it is always preferable to add another risk to portfolio. B: New situation ($\alpha < 1$). Bounded heavy-tailed distributions with VaR tolerant investor. For portfolios with few assets, value decreases with diversification. C: New situation ($\alpha < 1$). Bounded heavy-tailed distributions with VaR intolerant investor. For portfolios with few-medium assets, value decreases with diversification. D: Situation in Ibragimov (2004a, b, 2005) ($\alpha < 1$). Unbounded heavy-tailed distributions. Value always decreases with diversification.*

The following theorem is the analogue of Proposition 1 in the case of *bounded* risks. According to the theorem, the diversification continues to be disadvantageous for truncated extremely heavy-tailed distributions. The results demonstrate, in particular, that for any number $n \geq 2$ and any given disaster level $z > 0$, there exist n risks with *finite* support with the property that the return of their portfolio is *more* risky than that of the portfolio consisting *only of one* risk.

In what follows, for $z > 0$ and $w \in \mathcal{I}_n$, we denote by $G(w, z)$ the difference

$$G(w, z) = P\Big(w^{(1)}X_1 + w^{(2)}X_2 > z\Big) - P\Big(X_1 > z\Big), \tag{3}$$

which is positive if $w_{[1]} \neq 1$ since, by Proposition 1 applied to the portfolio of risks X_1, X_2 with weights $(w^{(1)}, w^{(2)})$, $P\Big(w^{(1)}X_1 + w^{(2)}X_2 > z\Big) > P\Big(X_1 > z\Big)$ if $w^{(i)} \neq 1$, $i = 1, 2$.

Theorem 1 *Let $n \geq 2$ and let $w \in \mathcal{I}_n$ be a portfolio of weights with $w_{[1]} \neq 1$. For any $z > 0$ and all*

$$a > \Big(\frac{E|X_1|^r(n-1)}{2G(w,z)}\Big)^{1/r}, \tag{4}$$

the following inequality holds:

$$P\Big(Y_w(a) > z\Big) > P\Big(Y_1(a) > z\Big). \tag{5}$$

Proof. We have

$$P\Big(X_w > z\Big) \leq P\Big(Y_w(a) \geq X_w > z\Big) + P\Big(X_w > z, X_w > Y_w(a)\Big) \leq$$
$$P\Big(Y_w(a) > z\Big) + P\Big(X_w > Y_w(a)\Big) \leq P\Big(Y_w(a) > z\Big) + P\Big(X_i > a \text{ for at least one i} \in \{1, 2, ..., n\}\Big) \leq$$
$$P\Big(Y_w(a) > z\Big) + \sum_{i=1}^{n} P\Big(X_i > a\Big) = P\Big(Y_w(a) > z\Big) + nP\Big(X_1 > a\Big). \tag{6}$$

From Proposition 1 it follows that

$$P\Big(X_w > z\Big) > P\Big(w^{(1)}X_1 + w^{(2)}X_2 > z\Big) =$$
$$P(X_1 > z) + G(w, z) = P\Big(X_1 > a\Big) + P\Big(Y_1(a) > z\Big) + G(w, z). \tag{7}$$

Relations (6) and (7) imply that the following inequalities hold:

$$P\Big(Y_w(a) > z\Big) - P\Big(Y_1(a) > z\Big) > G(w, z) - (n-1)P\Big(X_1 > a\Big). \tag{8}$$

Since, under the assumptions of the theorem, $E|X_1|^r < \infty$, by Chebyshev's inequality we get

$$P\Big(X_1 > a\Big) = \frac{1}{2}P\Big(|X_1| > a\Big) \leq \frac{E|X_1|^r}{2a^r}. \tag{9}$$

Estimates (8) and (9) give

$$P\Big(Y_w(a) > z\Big) - P\Big(Y_1(a) > z\Big) > G(w, z) - \frac{(n-1)E|X_1|^r}{2a^r}. \tag{10}$$

Under the conditions of the theorem, the right-hand side of (10) is positive. Consequently,

$$P\Big(Y_w(a) > z\Big) > P\Big(Y_1(a) > z\Big)$$

and, thus, (5) indeed holds. ■

Remark 3 *We note that in the case of a portfolio with equal weights* $\tilde{w}_n = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$, *one has* $(w^{(1)}, w^{(2)}) = \left(\frac{1}{2}, \frac{1}{2}\right)$ *and, thus, (3) becomes*

$$G(\tilde{w}_n, z) = H(z) = P\Big(\frac{X_1 + X_2}{2} > z\Big) - P\Big(X_1 > z\Big). \tag{11}$$

This means that the length of the distributional support providing the diversification failure results in Theorem 1 can be taken to be same for all the portfolios with equal weights $\tilde{w}_n$. *This holds, obviously, for the whole class of the portfolios* w *such that* $w_{[1]} < 1/2$. *Furthermore, a similar result holds as well for the class of portfolios* w *such that* $w_{[1]} < 1 - \epsilon$ *(and, thus,* $w_i < 1 - \epsilon$ *for all* i*), where* $0 < \epsilon < 1/2$. *As follows from the proof of Theorem 1, for all such portfolios* w*, the theorem holds for* $a > \left(\frac{E|X_1|^r(n-1)}{2\tilde{G}(\epsilon,z)}\right)^{1/r}$, *where* $\tilde{G}(\epsilon, z) = P\Big((1-\epsilon)X_1 + \epsilon X_2 > z\Big) < G(w, z)$. *Similar to Proposition 1, the last inequality follows from Corollary 5.3 in Ibragimov, 2004a and Theorem 4.2 in Ibragimov, 2004b since any vector* w *with* $w_{[1]} < 1 - \epsilon$ *is majorized by (that is, has more diverse components than) the vector* $(1 - \epsilon, \epsilon, 0, ..., 0)$.

Remark 4 *From the proof of Theorem 1 it follows that, in the case of portfolios with equal weights* $\tilde{w}_n = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$, *where* $n > 2$*, the length of the interval of truncation* a *can be reduced to a smaller value. In such a case, the theorem holds under the restriction* $a > \left(\frac{E|X_1|^r(n-1)}{2F_n(z)}\right)^{1/r}$, *where*

$$F_n(z) = P\Big(\frac{\sum_{i=1}^n X_i}{n} > z\Big) - P\Big(X_1 > z\Big).$$

Note that, by Proposition 1, $F_n(z) > H(z) = G(\tilde{w}_n, z)$ *for* $n \geq 3$.

Remark 5 *Theorem 1 does not hold uniformly for portfolios arbitrarily close to an undiversified portfolio. Thus, for any* a *and any number of stocks,* n*, it may be preferable to diversify "slightly." An asymptotic analysis shows that the required support,* a*, to ensure that diversification into* $w = (\epsilon, 1 - \epsilon)$ *is not preferred, grows as* $a \sim \epsilon^{-1/r}$. *Therefore, when* ϵ *approaches zero, the length of the distributional support* a *becomes unbounded.*

Remark 6 *As discussed in Ibragimov (2004a, b, 2005) analogues of Proposition 1 hold for i.i.d. risks* $X_1, ..., X_n$ *that have skewed extremely thick-tailed stable distributions with infinite first moments:* $X_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1)$, $\sigma > 0$, $\beta \in [-1, 1]$, $i = 1, ..., n$. *As follows from the proof of Theorem 1, this implies that complete analogues of the results in the present section for bounded versions of* symmetric *risks from the classes* $\mathcal{CS}(r)$ *continue to hold for truncated extremely heavy-tailed stable distributions* $S_\alpha(\sigma, \beta, 0)$ *with* $\alpha \in (0, 1)$, $\sigma > 0$, *and an* arbitrary skewness *parameter* $\beta \in [-1, 1]$. *In particular, Theorem 1 continues to hold for arbitrary skewed risks* $X_i \sim S_\alpha(\sigma, \beta, 0)$, $\alpha \in (0, 1)$, $\sigma > 0$, $\beta \in [-1, 1]$ *if* $a > \left(\frac{E|X_1|^r(n-1)}{G(w,z)}\right)^{1/r}$. [19]

Remark 7 *As indicated in Ibragimov (2004a, b, 2005), Proposition 1 and its extensions imply corresponding results on peakedness, monotone consistency and efficiency properties of linear estimators* $\sum_{i=1}^n w_i X_i$ *under heavy-tailedness. In particular, from the proposition it follows that, the sample mean of observations from extremely heavy-tailed*

[19] The factor 2 in the denominator of the bound for a in Theorem 1 needs to be replaced by 1 in this case because, if X_1 is not necessarily symmetric, then the tail probability $P\Big(X_1 > a\Big)$ on the right-hand side of (9) is, in general, bounded from above by $\frac{E|X_1|^r}{a^r}$.

populations exhibit decreasing *peakedness about the population center as the sample size increases. Thus, having more data is always disadvantageous for inference if the sample mean is used to estimate the population center under extreme thick-tailedness. For instance, if* $Z_i = \mu + X_i$, *where* X_i, $i \geq 1$, *are symmetric i.i.d. shocks with extremely thick-tailed distributions:* $X_i \sim \mathcal{CS}(r)$, *then, for all* $\epsilon > 0$ *and all* $n > 1$, $P\Big(|\overline{Z}_n - \mu| > \epsilon\Big) = P\Big(|\frac{1}{n}\sum_{i=1}^{n} Z_i - \mu| > \epsilon\Big) > P\Big(|Z_1 - \mu| > \epsilon\Big)$. *Similar to the above, Theorem 1 provides econometric and statistical results that concern efficiency, peakedness and monotone consistency properties of linear estimators for data from* bounded *populations. For instance, Theorem 1 and Remark 3 imply the following result that shows that an increase in the sample size is disadvantageous for inference if the sample mean is used to estimate the population center for data with truncated extremely thick-tailed distributions. Let, again,* $X_1, X_2, ...$ *be a sequence of i.i.d. r.v.'s distributions from the class* $\mathcal{CS}(r)$, $r \in (0,1)$. *Further, let* $\epsilon > 0$ *and let* $a > \left(\frac{E|X_1|^r(n-1)}{2H(\epsilon)}\right)^{1/r}$, *where* $H(\epsilon) = G(\tilde{w}_n, \epsilon)$ *is defined in (11). Suppose that* $Z_i = \mu + Y_i(a)$, $i = 1, ..., n$, *where, as before,* $Y_i(a) = X_i I(|X_i| \leq a)$ *denote the truncated versions of* $X_i's$. *Then, according to the results in Theorem 1 and Remark 3, the likelihood of making an error greater than* ϵ *in estimating* μ *using the sample mean of the whole sample of* n *observations* Z_i *is greater than that using only one datum* Z_1: $P\Big(|\overline{Z}_n - \mu| > \epsilon\Big) = P\Big(\Big|\frac{1}{n}\sum_{i=1}^{n} Z_i - \mu\Big| > \epsilon\Big) > P\Big(|Z_1 - \mu| > \epsilon\Big)$.

Remark 8 *Theorem 1 means that, for a* specific *loss probability* $q \in (0, 1/2)$, *there exists a sufficiently large* a *such that the value at risk* $VaR_q\big[Y_w(a)\big]$ *of the return* $Y_w(a)$ *at level* q *is greater than the value at risk* $VaR_q\big[Y_1(a)\big]$ *of the return* $Y_1(a)$ *at the same level:* $VaR_q\big[Y_w(a)\big] > VaR_q[Y_1(a)]$. *On should emphasize that the last inequality between the returns* $Y_w(a)$ *and* $Y_1(a)$ *holds for the particular* fixed *loss probability* q *and, in the comparisons of the values at risks* $VaR_q\big[Y_w(a)\big]$ *and* $VaR_q[Y_1(a)]$, *the length of the interval needed for the reversals of the stylized facts on the portfolio diversification depends on* q *(similar to the fact that in Theorem 1, the length of the distributional support* a *depends on the value of the disaster level* z*). This is the crucial qualitative difference of the results in Theorem 1 for bounded distributions and their implications for the value at risk from those given by Theorem 1 and Remark 2 for unbounded risks where the inequalities hold for* all $z > 0$ *and* all $q \in (0, 1/2)$.

3.2 Non-diversification and general risk rankings

In this section we relate the VaR approach to the expected utility framework, both for unbounded and bounded heavy-tailed risks. As noted in the introduction, with extremely heavy-tailed distributions, a direct expected utility approach does not work as integrals may not be defined. In what follows, we argue that with natural generalizations of the concepts of expected utility and risk, the same type of risk rankings can be applied for a wide class of symmetric heavy-tailed distributions. It allows us to conclude that for such distributions, diversified portfolios are dominated by undiversified ones from a stochastic dominance perspective. In other words, the monotone decrease of the line D in Figure 1 holds from a general risk perspective.

We work with r.v.'s X and Y, with c.d.f.'s F_X and F_Y respectively and, as before, denote their a-truncated versions by X^a and Y^a.[20] The corresponding c.d.f.'s are denoted by F_X^a and F_Y^a. We will also denote their p.d.f.'s by f_X and f_Y. Following Ingersoll (1970), we define a simple mean preserving spread (MPS) of a c.d.f., F, with

[20] As we relate to the expected utility framework, we use the convention that negative valuers of X and Y are losses in this section.

corresponding p.d.f., f, by adding to f a function $\varphi(x)$ satisfying:

$$\varphi(x) = \begin{cases} \alpha & \text{for } c < x < c + t, \\ -\alpha & \text{for } c' < x < c' + t, \\ -\beta & \text{for } d < x < d + t, \\ \beta & \text{for } d' < x < d' + t, \\ 0 & \text{elsewhere,} \end{cases} \tag{12}$$

where $\alpha(c' - c) = \beta(d' - d)$, $\alpha > 0$, $\beta > 0$, $t > 0$, $c + t < c' < d - t$, and $d + t < d'$. If $f(x) + \varphi(x) \geq 0$ for all x, then the function $G(x) = F(x) + \int_{-\infty}^{x} \varphi(s)ds$ is a simple mean preserving spread of F. For any c.d.f., F, we define M_F, the set of c.d.f.'s obtainable by a finite number of simple MPS's on F.

We equip the space of distributions with the Lévy metric

$$d(F_X, F_Y) = \inf\{\epsilon : F_X(x - \epsilon) - \epsilon \leq F_Y(x) \leq F_Y(x + \epsilon) + \epsilon, \text{ for all } x\}. \tag{13}$$

This makes it a complete metric space with the topology of weak convergence.[21] For the MPS condition, we will use the closure of M_F, $\overline{M}_F$, and say that if $G \in \overline{M}_F$ then G can be obtained by a sequence of MPS's on F, or simply that G is an MPS of F. We note that if $\varphi(x)$ satisfies (12), then so does $\varphi(-x)$ and $c\varphi(x)$ for any $c > 0$. It follows that if X and Y are symmetric and $Y \in M_{F_X}$ then Y can be obtained from X by a finite sequence of pairs of simple MPS's, each pair being symmetric. Specifically, if $f_Y(x) = f_X(x) + \varphi(x)$, where $\varphi = \sum_{i=1}^{N} \varphi_i$, each φ_i satisfying (12), then, clearly, $f_Y(x) = f_X(x) + \sum_{i=1}^{N} \Big(\varphi_i(x)/2 + \varphi_i(-x)/2\Big)$.

Rothschild and Stiglitz (1970) introduce four equivalent definitions of risk dominance, one of which is directly related to expected utility theory. The following four conditions B1-B4 are equivalent for two r.v.'s X and Y with bounded support in $[-a, a]$ and c.d.f.'s F_X and F_Y such that $EX = EY$,[22] defining a partial – second order stochastic dominance – ordering over risks and their c.d.f.'s, $X \succeq Y$ and $F_X \succeq F_Y$.

- B1: F_Y can be obtained by a sequence of mean-preserving spreads (MPS) of F_X.
- B2: For all $t \in (a, b)$, $\int_a^t F_X(x)dx \leq \int_a^t F_Y(x)dx$.
- B3: For all concave utility functions, $u : [-a, a] \to \mathbf{R}$: $Eu(X) \geq Eu(Y)$.
- B4: $Y \stackrel{d}{=} X + U$, where U is a r.v. on $[-a, a]$ such that $E(U|X) = 0$.

Remark 9 *There is an interesting link between these definitions and assets with option-like payout. Condition (B2) (and, thus, any other of the above conditions) is equivalent to the following:*

- *$\tilde{B}2$: For all $t \in (a, b)$, $E(X - t)_+ \leq E(Y - t)_+$, where $z_+ = \max(z, 0)$ for $z \in \mathbf{R}$.*

While condition B3 means that every risk averse prefers X to Y, the interpretation of condition $\tilde{B}2$ (with $t > 0$) is that the expected payoff of an European call option with the strike price t written on the risk X is not greater than that of the same contingent claim on the risk Y.[23]

[21] See Lukacs (1975).

[22] Equivalently, $\int_{-a}^{a} x dF_X(x) = \int_{-a}^{a} x dF_Y(x)$).

[23] The reader is referred to de la Peña, Ibragimov and Jordan, 2004, and Ibragimov and Brown, 2005, for the discussion of comparisons for c.d.f.'s of transforms of dependent r.v.'s and their relation to sharp bounds on the expected payoffs and fair prices of European call options and other contingent claims on such risks.

Next, following Birnbaum (1948), we define X to be more peaked about 0 than Y if $P(|X| > x) \leq P(|Y| > x)$ for all $x \geq 0$. The VaR results in Ibragimov (2004a, b, 2005) provided by Proposition 1 can also be cast in peakedness terminology: For any $r < 1$ and all i.i.d. risks $X_i \sim \mathcal{CS}(r)$, $i = 1, ..., n$, the r.v. X_1 is more peaked about the origin than the return X_w on the portfolio of X_i's with weights $w \in \mathcal{I}_n$ such that $w_{[1]} \neq 1$.

For symmetric distributions with finite absolute first moments and, in particular, for bounded symmetric distributions, peakedness implies second order stochastic dominance, as the following Lemma 1 demonstrates. Let X and Y be two symmetric risks with the same distribution support $[-a, a] \subseteq \mathbf{R}$.[24] In the case $a = \infty$, we assume that $E|X| < \infty$ and $E|Y| < \infty$.

Lemma 1 *Under the above assumptions, if X is more peaked about 0 than Y, then F_X and F_Y satisfy condition B2 (and, thus, in the bounded case conditions B1, $\tilde{B}2$, B3 and B4).*

Proof. For $t \in [-a, a]$, denote $G(t) = \int_{-a}^{t} \big(F_X(s) - F_Y(s)\big)ds$. We have that $G(-a) = 0$ and $G(a) = \int_{-a}^{a} \big(F_X(s) - F_Y(s)\big)ds = 0$ by symmetry of X and Y. In addition, peakedness comparisons for X and Y imply that $G'(t) = F_X(t) - F_Y(t) \leq 0$ for $t \in [-a, 0]$ and $G'(t) \geq 0$ for $t \in [0, a]$. The above properties of the function G evidently imply that $G(t) \leq G(-a) = 0$ for $t \in [-a, 0]$ and $G(t) \leq G(a) = 0$ for $t \in [0, a]$. Consequently, B2 indeed holds. ■

Thus, for symmetric r.v.'s with finite first absolute moments and, thus, for symmetric bounded r.v.'s, peakedness provides a ranking of risks that is at least as informative as second order stochastic dominance.

Remark 10 *Lemma 1 is related to several extremal results in probability theory that were demonstrated to be helpful in obtaining sharp bounds on the tail probabilities of self-normalized statistics and developing conservative testing procedures for dependent and/or heavy tailed observations (see de la Peña and Ibragimov, 2003) and can also be used, similar to de la Peña, Ibragimov and Jordan, 2004, and Ibragimov and Brown, 2005, to obtain sharp estimates for expected payoffs and fair prices of European options and other contingent claims. For example, from the results obtained by Hunt (1955), it follows that for all continuous concave functions $u : [-a, a] \rightarrow \mathbf{R}$ and all r.v.'s X such that $EX = 0$ and $|X| \leq a$, the following estimate holds: $Eu(X) \geq Eu(Y)$, where Y is a r.v. with the following distribution: $P(Y = a) = P(Y = -a) = 1/2$. This inequality means that condition B3 (and, thus, each of the conditions B1, B2, $\tilde{B}2$ and B4) with the above r.v. Y holds for all r.v.'s X on $[-a, a]$ under the* only *assumption that $EX = 0$; condition B1, on the other hand, is trivially satisfied for the r.v.'s X and Y in the present instance since, for all $a \geq x > 0$, $P(|X| > x) \leq 1 = P(|Y| > x)$. Similar to Hunt's argument, the inequality $Eu(X) \geq Eu(Y)$ can be proved as follows. Consider a r.v. $\tilde{Y}$ that has the following distribution conditional on X: $P\big(\tilde{X} = -a\big|X\big) = (a - X)/(2a)$ and $P\big(\tilde{X} = a\big|X\big) = (X - a)/(2a)$. It is easy to see that the unconditional distribution of $\tilde{X}$ is the same as that of Y because of the assumption $EX = 0$. In addition, $E(\tilde{X}|X) = X$. Therefore, for any concave functions $u : [-a, a] \rightarrow \mathbf{R}$, by the conditional Jensen's inequality, $Eu(Y) = Eu(\tilde{X}) = E\big[E\big(u(\tilde{X})|X\big)\big] \leq Eu\big[E\big(\tilde{X}|X\big)\big] = Eu(X)$. The above comparisons are easily generalized to the case of r.v.'s X defined on an arbitrary interval $[a, b]$. Namely, using the argument similar to that above, one can show, that for all r.v.'s X such that $a \leq X \leq b$ (a.s.) and all concave functions u on $[a, b]$, $Eu(X) \geq Eu(Y)$, where*

[24] The value of a can be infinity so that $[-a, a]$ can be the whole real line: $[-a, a] = \mathbf{R}$.

Y is a two-valued r.v. with the following distribution: $P(Y=a)=\frac{b-EY}{b-a}$, $P(Y=b)=\frac{EY-a}{b-a}$.

We next turn to unbounded symmetric distributions for which the first absolute moments do not exist. Specifically, we study extremely heavy-tailed symmetric distributions and without loss of generality, we assume that the point of symmetry is the origin. We therefore look at the class of distributions $\mathcal{CS}(r)$, $0<r<1$. Ideally, we would like to generalize the equivalence of B1–B4 to distributions in $\mathcal{CS}(r)$ with $r\in(0,1)$. This would provide an unambiguous risk ranking. However, as is usually the case, the picture becomes more complicated with unbounded heavy-tailed risks.

It is evident that, given two symmetric r.v.'s X and Y on $\mathbf{R}$, X is more peaked than Y if and only if, for any $a>0$, the truncated version X^a of X is more peaked than the truncated version Y^a of Y. From Lemma 1 we get, therefore, that if X is more peaked than Y, then, for any $a>0$, the c.d.f.'s F_{X^a} and F_{Y^a} of the truncated versions X^a and Y^a of the r.v.'s satisfy conditions B1, B2, $\tilde{B}2$, B3 and B4.

Below, for a r.v. W, we denote by $\sigma(W)$ the $\sigma-$algebra spanned by it. In addition, $P(\cdot|W)$ denotes the $\sigma(W)-$conditional probabilities. Given two symmetric r.v.'s X and Y with c.d.f.'s F_X and F_Y, we consider the following conditions.

- B0': F_X is more peaked about the origin than F_Y (Peakedness condition).
- B1': $F_Y\in\overline{M}_{F_X}$ (MPS condition).
- B2': There is an a_0 such that for all $a>a_0$: F_X^a and F_Y^a satisfies B2 (Strong integral condition).
- B2": For all $\epsilon>0$, there exists $a>0$ and a c.d.f. $\tilde{F}^\epsilon:[-a,a]\to[0,1]$, such that
  1. For all $t\in(-a,a)$: $\int_{-a}^t F_X^a(x)dx\le\int_{-a}^t\tilde{F}^\epsilon(x)dx$.
  2. $F_Y^a=\tilde{F}^\epsilon+\xi+s$, where ξ is an antisymmetric[25] function satisfying
  $$F_Y(x-\epsilon)-F_Y(x)\ \le\ \xi(x)\ \le\ F_Y(x+\epsilon)-F_Y(x),\quad\text{for all } x \tag{14}$$
  and s is an antisymmetric function with $\Big|s(x)\Big|\le\epsilon$ for (almost) all x.
  3. If $\epsilon\to0$, then $a\to\infty$.

  (Weak integral condition).[26]
- B3': There is an a_0 such that for all $a>a_0$ for all concave u: $Eu(X^a)\ge Eu(Y^a)$ (Expected utility condition).
- B4': There is an a_0 such that for all $a>a_0$, $Y^a\overset{\mathrm{d}}{=}X^a+Z^a$, where Z^a is a $\sigma(X^a)$-measurable r.v. such that $E(Z^a|X^a)=0$ (a.s.) (Fair game condition).

[25]That is, $f(-x)=-f(x)$ for all x.

[26]Clearly, (14) implies that $|\xi(x)|\le1$ (a.s.) and $\int_{-a}^t\xi(x)dx\le\epsilon$ for all t. Thus, the weak integral condition allows for "approximate" MPS's on bounded sets in the sense that F_Y^a is the sum of an MPS (F^ϵ), a term which is "small" in integration (ξ) and a term which is small in maximum norm (s). Moreover, if F_X and F_Y are absolute continuous, then one can choose $\xi=0$. This is the case as $|F_Y(x+\nu)-F_Y(x)|\le C|\nu|$ for all $|\nu|$ and therefore the condition will be satisfied with $\xi=0$ and $|s(x)|\le\tilde{\epsilon}=(C+1)\epsilon$.

- B4”: $Y \stackrel{d}{=} sign(Y)\big[|X|+Z\big] \stackrel{d}{=} X+sign(X)Z$, where $Z \geq 0$ (a.s.) (Conditional absolute symmetry condition).[27]

We show that the conditions are related as in Figure 2, i.e.,

Theorem 2 *For distributions symmetric about the origin: 1. B0’ is equivalent to B4”, 2. B0’ implies B2’, 3. B2’ is equivalent to B3’ and B4’ 4. B2’ implies B2” and 5. B1’ is equivalent to B2”.*

Proof.

1. B0’ $\iff$ B4”: Let $F_{|X|}(t)$ and $F_{|Y|}(t)$, $t \geq 0$, denote the c.d.f.’s of the r.v.’s $|X|$ and $|Y|$ and let, as usual, $F^{-1}_{|X|}(u)$ and $F^{-1}_{|Y|}(u)$, $u \in [0,1]$ stand for their right continuous inverses: $F^{-1}_{|X|}(u) = \sup\{x : F_{|X|}(x) \leq u\}$ and $F^{-1}_{|Y|}(u) = \sup\{x : F_{|Y|}(x) \leq u\}$. The r.v. $F_{|X|}[|X|]$ has the uniform distribution on $[0,1]$ and, therefore, $|Y| \stackrel{d}{=} F^{-1}_{|Y|}\big[F_{|X|}(|X|)\big]$.

Suppose that B0’ holds. Then, as is not difficult to see, $F^{-1}_{|Y|}(t) \geq F^{-1}_{|X|}(t)$ for all $t \geq 0$. Consequently, the $\sigma(|X|)-$measurable r.v. Z defined by $Z = F^{-1}_{|Y|}\big[F_{|X|}(|X|)\big] - F^{-1}_{|X|}\big[F_{|X|}(|X|)\big] = F^{-1}_{|Y|}\big[F_{|X|}(|X|)\big] - |X|$ is a.s. nonnegative: $P\big(Z \geq 0\big) = 1$. We have that $|Y| \stackrel{d}{=} |X| + Z$ and, thus, $Y \stackrel{d}{=} sign(Y)|X| + sign(Y)Z$. Since, evidently, $sign(Y)$ and $sign(X)$ are symmetric Bernoulli r.v.’s independent of $|X|$, this implies that $Y \stackrel{d}{=} sign(X)|X| + sign(X)Z = X + sign(X)Z$, that is, B4” is satisfied.

Suppose now that B4” holds. Then, evidently, for all $x \geq 0$, $P\big(|Y| > x\big) = P\big(|X| + Z > x\big) \geq P\big(|X| > x\big)$, that is, X is more peaked than Y and B0’ is satisfied.

2. B0’ $\implies$ B2’: Clearly, for symmetric unbounded r.v.’s, if X is more peaked about 0 than Y, then X^a is more peaked about 0 than Y^a for all $a \geq 0$. This, together with Lemma 1 implies that B2’ is satisfied.

3. The equivalence follows from Rothschild and Stiglitz (1970). Clearly, the same a_0 can be used to satisfy both B2’, B3’ and B4’.

4. B2’ $\implies$ B2” is immediate, as for any $\epsilon > 0$ and $a > a_0$, the condition is satisfied with $\xi = s = 0$ and $\tilde{F}^\epsilon = F^a_Y$.

5i. B1’ $\implies$ B2”: If $F_Y \in \overline{M}_{F_X}$, then for any $\epsilon > 0$, we can choose a finite sequence of MPS’s to obtain $F^\epsilon \in M_{F_X}$, which is in an ϵ-neighborhood of F_Y, i.e., $d(F^\epsilon, F_Y) \leq \epsilon$. Without loss of generality, we can assume that F^ϵ is symmetric for all ϵ. This follows from our previous symmetrization discussion, i.e., for any nonsymmetric MPS, $\varphi^\epsilon(x)$, we can define $\tilde{\varphi}^\epsilon(x) = \varphi^\epsilon(x)/2 + \varphi^\epsilon(-x)/2$. As convex summation and reflection (of p.d.f.’s) are continuous operators on distributions in the Lévy metric, this ensures that $\tilde{\varphi}^\epsilon$ converges to φ^ϵ as ϵ approaches zero. Thus, by defining $F^\epsilon(x) = F_X(x) + \tilde{\varphi}^\epsilon(x)$, we get a sequence of symmetric distributions in M_{F_X} converging to F_Y.

As the MPS is in M_{F_X}, the support of $F_X - F^\epsilon$ is bounded and we can define $a = \max\{\max \operatorname{supp} \tilde{\varphi}^\epsilon, 1/\epsilon\} < \infty$,

where $\max \operatorname{supp} \tilde{\varphi}^\epsilon$ is the length of the support of $\tilde{\varphi}^\epsilon$.

[27]For condition B4”, we restrict our attention to absolutely continuous distributions. However, complete analogues of the results below hold as well in the discrete case.

We have

$$\int_{-a}^{t}(F_X^a(x)-F_Y^a(x))dx=\int_{-a}^{t}(F_X^a(x)-F^\epsilon(x))dx+\int_{-a}^{t}(F^\epsilon(x)-F_Y^a(x))dx. \tag{15}$$

From the definition of the Lévy metric, we have $d(F^\epsilon, F_Y^a)\le d(F^\epsilon, F_Y)\le\epsilon$, so

$$\begin{aligned}&F_Y^a(x-\epsilon)-\epsilon\le F^\epsilon(x)\le F_Y^a(x+\epsilon)+\epsilon \Longrightarrow\\ &\quad F^\epsilon(x)-F_Y^a(x)\in\Big[F_Y^a(x-\epsilon)-F_Y^a(x)-\epsilon, F_Y^a(x+\epsilon)-F_Y^a(x)+\epsilon\Big]\Longrightarrow\\ &\qquad F^\epsilon(x)-F_Y^a(x)=F_Y^a(x+\eta(x))-F_Y^a(x)-s(x),\ \ \text{(a.s.)},\end{aligned}$$

where $\eta(x)\in[-\epsilon,\epsilon]$ and $s(x)\in[-\epsilon,\epsilon]$. Furthermore, η and s are antisymmetric from the symmetry of F_Y^a and F^ϵ. Therefore, by defining the antisymmetric function $\xi(x)=F_Y^a(x)-F_Y^a(x+\eta(x))$ we get

$$F_Y^a(x)=F^\epsilon(x)+\xi(x)+s(x). \tag{16}$$

5ii. B2" $\Longrightarrow$ B1': The integral condition implies that $F^\epsilon\in\overline{M}_{F_X^a}$. We define $\tilde{F}^\epsilon=F_X+F^\epsilon-F_X^a$. Clearly, $\tilde{F}^\epsilon\in\overline{M}_{F_X}$. By the triangle inequality, we have

$$d(\tilde{F}^\epsilon,F_Y)\le d(\tilde{F}^\epsilon,F^\epsilon)+d(F^\epsilon,F_Y^a)+d(F_Y^a,F_Y). \tag{17}$$

For the third term on the RHS, we have $\lim_{a\to\infty}d_K(F_Y^a,F_Y)=0$, where $d_K(F,G)$ denotes the Kolmogorov distance, $d_K(F,G)=\sup_x|F(x)-G(x)|$. Similarly, for the first term $d_K(\tilde{F}^\epsilon,F^\epsilon)=d_K(F_X,F_X^a)$, which converges to zero when ϵ approaches zero. As the Kolmogorov topology is stronger than the Lévy topology, this implies that the first and third terms converge to zero.

Finally, for the second term, we have

$$F_Y^a=F^\epsilon+\xi+s, \tag{18}$$

where $F_Y^a(x)-F_Y^a(x+\epsilon)\le\xi(x)\le F_Y^a(x)-F_Y^a(x-\epsilon)$, and $-\epsilon\le s\le\epsilon$. Therefore, $F_Y^a(x-\epsilon)-\epsilon\le F^\epsilon(x)\le F_Y^a(x+\epsilon)+\epsilon$ for all x, i.e., $d(F^\epsilon,F_Y^a)\le\epsilon$.

We have shown that $\lim_{\epsilon\to0}d(\tilde{F}^\epsilon,F_Y)=0$ for a sequence $\tilde{F}^\epsilon\in\overline{M}_{F_X}$ and thus by completeness, $F_Y\in\overline{M}_{F_X}$. ■

Remark 11 *From condition B4' it follows that* $Y\overset{\mathrm{d}}{=}X+H$*, where* H *is a* $\sigma(|X|)-$*conditionally symmetric r.v.:* $P\big(H>t\big||X|\big)=P\big(H<-t\big||X|\big)$ *for all* $t>0$*. One should note that, in representation given by B4' in the above form, the* $\sigma(|X|)-$*conditionally symmetric summand* $H=sign(X)Z$ *depends on* X*. From the results in Birnbaum (1948) (see also Theorem 2.C.3 in Shaked and Shanthikumar, 1994, and the proof of Theorems 4.1-4.4 in Ibragimov, 2004a) it follows that if the* X *and* H *are* independent *symmetric* unimodal *r.v.'s then* X *is less peaked than* $Y=X+H$*. It is important to emphasize that, in general, the fact that*

$$Y\overset{\mathrm{d}}{=}X+H, \tag{19}$$

where X *has a symmetric distribution and* H *is* $\sigma(|X|)-$*conditionally symmetric r.v.,* does not *imply that* Y *is more peaked than* X*, even in the case when* X *and* Z *are independent. The argument for this is similar to Birnbaum's*

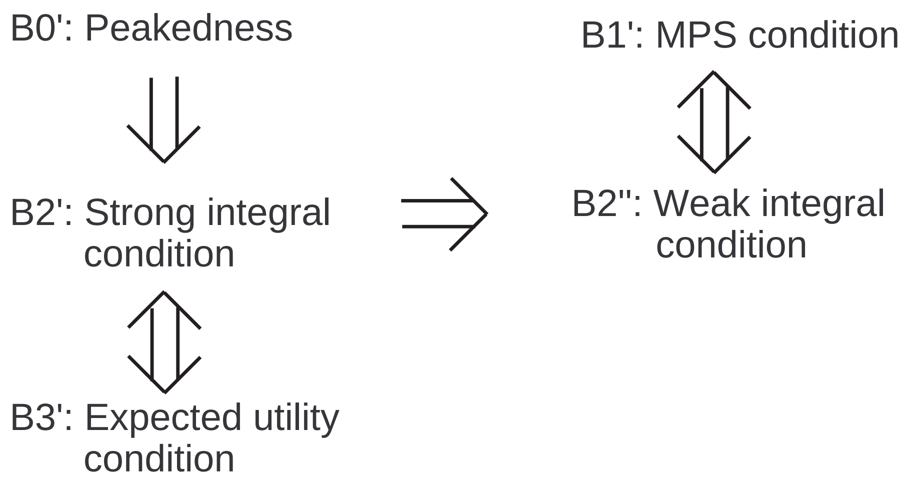

Figure 2: *Relationship between risk rankings for unbounded symmetric distributions.*

(1948) example (see also example 7.1 in Dharmadhikari and Joag-Dev, 1988) that shows that unimodality cannot be completely dispensed for peakedness comparisons to be preserved under convolutions of symmetric distributions. Indeed, suppose that X is a symmetric Bernoulli r.v.: $P(X = 1) = P(X = -1) = 1/2$ and let H be an independent of X r.v. with a uniform distribution on $[-1, 1]$. It is easy to see that the r.v. $X + H$ has a uniform distribution on $[-2, 2]$ and, thus $P(|X + H| > x) > P(|X| > x) = 0$ for all $x > 1$, while $P(|X + H| > x) < P(|X| > x) = 1$ for all $0 \leq x < 1$. That is, the r.v.'s X and $X + H$ are not ordered by peakedness. Since the r.v. X can be approximated by an absolutely continuous r.v. with a $U-$shaped density, representation (19) does not imply $B0'$ in the continuous case as well.

For portfolios of risks in $\mathcal{CS}(r)$, an undiversified portfolio is more peaked about the origin than any diversified portfolio. Thus, an undiversified portfolio also any dominates diversified portfolios in the sense of B1', B2', B2'', B3', B4' and B4''. Therefore the results in Ibragimov (2004a, b, 2005), on diversification always being non preferable, also are true in each of these senses. This concludes our analysis of the limits of diversification for unbounded heavy-tailed risks.

We next compare the VaR results for bounded distributions with the traditional results on diversification. The results in the previous section show that diversification is suboptimal for a large class of distributions with bounded support when value at risk is used as portfolio benchmark measure. This is contrary to the standard view that diversification is always to be preferred. For the case with unbounded risks it can be attributed to the non-existing moments of distributions in $\mathcal{CS}(r)$. However, the distributions in Theorem 1 have bounded (but large) support and finite moments of all orders exist. We therefore analyze what drives the differences compared with the traditional results on diversification.

There are two main motivations for diversification in traditional portfolio theory. The first approach uses the law of large numbers (LLN). The second approach uses expected utility/stochastic dominance.[28] For the first

[28] We view Markowitz' (1952) mean-variance approach as a special case of the latter.

approach, the law of large numbers implies that, for all $\epsilon, \epsilon_1 > 0$, $P\left(\left|\overline{Z}_n - \mu\right| > \epsilon\right) = P\left(\left|\frac{\sum_{i=1}^n Z_i}{n} - \mu\right| > \epsilon\right) < \epsilon_1$ if $n > N(\epsilon_1) > 0$ and the risks $Z_1, Z_2, ...$ are i.i.d. r.v.'s with $EZ_1 = \mu$, and $\overline{Z}_n = n^{-1}\sum_{i=1}^n Z_i$. Thus, as n becomes large, all risk disappears and the diversified portfolio will be preferred. This type of argument has strong asset pricing implications, as shown in the celebrated arbitrage pricing theory (Ross 1976), which analyzes the case when n becomes unbounded. Our approach differs from the LLN approach, in that we asymptotically increase the distributional support, a, as the number of assets, n, increases. This leads to the break-down of the rule. Practically speaking, we assume that the effective distributional support of Z_i is relatively large compared to the number of assets where large is defined by equation (4).

The second motivation for diversification is based on expected utility. Samuelson (1967) showed that any investor with a strictly concave utility function will uniformly diversify among i.i.d. risks with finite second moments, i.e., will choose the portfolio with equal weights and the return $\overline{Z}_n$ among all portfolios. As our previous discussion shows, this breaks down for unbounded extremely heavy-tailed distributions, but it must hold in all situations with bounded support. In light of condition B3, the result in Samuelson (1967) implies that $\overline{Z}_n$ second order stochastically dominates the distributions of all other portfolios.

Why does the expected utility approach favor diversification for any a, even though, as follows from Theorem 1 and Remark 8, for a specific loss probability, $q \in (0, 1/2)$, a can always be chosen large enough so that the diversified portfolio has higher value at risk than the undiversified portfolio: $VaR_q(\overline{Y}_n(a)) > VaR_q(Y_1(a))$? The reason is that regardless of a, there will always be a region further out in the probability tail where the inequality is reversed: for some $\tilde{q} >> q$, $VaR_{\tilde{q}}(\overline{Y}_n(a)) < VaR_{\tilde{q}}(Y_1(a))$. This is contrary to the case when $a = \infty$ in which no such reversal takes place. The argument is illustrated in Figure 3. The concavity of the utility function over the whole real line then implies that diversification is always preferred. Specifically, the concavity of the utility implies that the impact in the tail beyond $\tilde{q}$ will be higher that the impact between q and $\tilde{q}$. Thus, the expected utility argument in favor of diversification with truncated heavy-tailed distributions depends fundamentally on the behavior of the utility function in the domain of extreme negative outcomes. Therefore, under the assumption of strict risk aversion for arbitrary large negative outcomes, the VaR measure is "wrong" regardless of the distributional support, a.

However, there are several situations where assuming concavity over all outcomes may be a stretch. First, experimental results leading to Prospect theory have shown that decision makers' utility functions may be convex in the domain of losses (Kahneman and Tversky, 1979). Second, limited liability introduces an option-like payoff structure, as do several agency problems (see e.g. Stiglitz, 1974, Jensen and Meckling, 1976, Stiglitz and Weiss, 1981, and Gollier, Koehl and Rochet, 2001). This may lead to the expected utility function being effectively convex, with respect to the original distribution. Thus, any of these effects make the assumption on strict concavity of expected utility over the whole real line implausible. In situations where concavity may only be assumed over a bounded domain of outcomes arguments based on asymptotes of the utility function are as dubious as arguments based on asymptotic behavior of the tails of probability distributions.

We formalize this argument by showing that as long as there is a point arbitrarily far out in the domain of negative

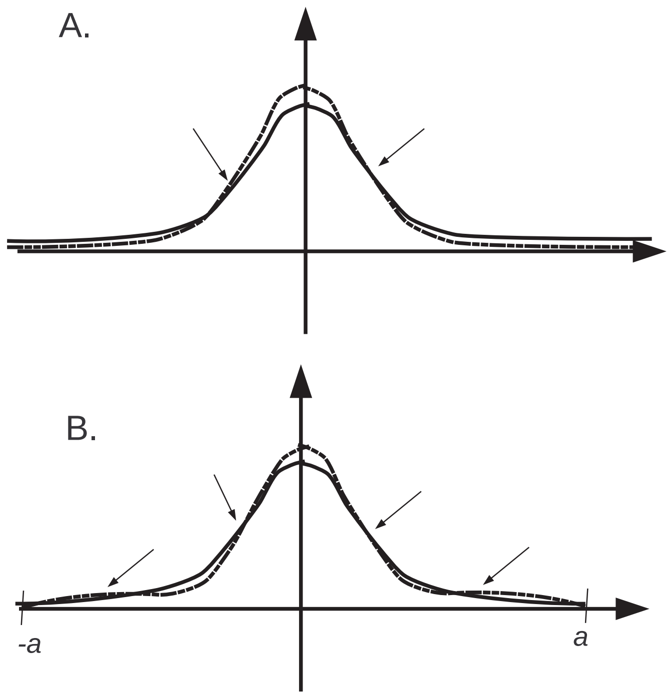

Figure 3: *Differences between untruncated (A) and Truncated (B) distributions. For untruncated, only one crossing on each side of origin. For truncated, (at least) two crossings. Solid lines represent undiversified risk, and dotted line diversified risk.*

outcomes beyond which the utility is convex, increasing and bounded, there is an expected utility specification for which the nondiversification result continues to hold, in line with Theorem 1. We use the following notation. An increasing, strictly concave function: $u : \mathbf{R} \to \mathbf{R}$ will be called admissible. For any $t > 0$, a continuous function $v : \mathbf{R} \to \mathbf{R}$ is called a t-convex regularization of an admissible function u, if $v(x) = u(x)$ for $x \geq -t$, v is increasing and twice continuously differentiable on $(-\infty, -t)$, and $u(-t) - \lim_{s\to\infty} v(-s) \leq 1/t$. For a large t, a t-convex regularization is thus a way of introducing a region of convexity far out in the negative domain of the utility function, while keeping the assumption of strictly positive marginal utility. There are of course many ways to create a t-convex regularization of any admissible function.

As in the previous section, for $r < 1$, and n i.i.d. risks, $X_i \sim \mathcal{CS}(r)$, we consider the truncated r.v.'s $Y_i(a)$, $i = 1, \ldots, n$ and the diversified portfolio with equal weights $\tilde{w}_n = \left(\frac{1}{n}, \frac{1}{n}, ..., \frac{1}{n}\right)$ and the return $\overline{Y}_n(a) = \dfrac{1}{n}\sum_{i=1}^{n} Y_i(a)$.

We have

Theorem 3 *Let $n \geq 2$. Then there exists a t_0, such that for any $t \geq t_0$, there is an admissible utility function u, and $a > 0$, such that any investor with utility function, v, where v is a t-convex regularization of u, will have*

$$Ev\Big(Y_1(a)\Big) > Ev\Big(\overline{Y}_n(a)\Big). \tag{20}$$

Proof. From Theorem 1, we know that for any $t > 0$, we can choose an a such that

$$F_{Y_1(a)}(-t) < F_{\overline{Y}_n(a)}(-t).$$

We have

$$\int_{-\infty}^{-t} -t\, dF_{Y_1(a)}(x) + \int_{-t}^{\infty} x\, dF_{Y_1(a)}(x) = \int_{-\infty}^{-t} -t\, dF_{\overline{Y}_n(a)}(x) + \int_{-t}^{\infty} x\, dF_{\overline{Y}_n(a)}(x) - s, \qquad s > 0.^{29} \tag{21}$$

That $s > 0$ follows from Rothschild and Stiglitz (1970) and that $Y_1(a)$ is a mean preserving spread of $\overline{Y}_n(a)$. Specifically, the integral takes the form $\int QdF$, where $Q(x) = (x+t)_+ - t$ is convex and therefore $-Q$ is concave. For a specific a, we can clearly choose an admissible utility function u, such that

$$\Big| \int_{-t}^{\infty} u(x) dF_{Y_1(a)}(x) - \int_{-t}^{\infty} x dF_{Y_1(a)}(x) \Big| \le \frac{s}{6}, \qquad \Big| \int_{-t}^{\infty} u(x) dF_{\overline{Y}_n(a)}(x) - \int_{-t}^{\infty} x dF_{\overline{Y}_n(a)}(x) \Big| \le \frac{s}{6}, \tag{22}$$

and, for t large enough

$$\Big| \int_{-\infty}^{-t} v(x) dF_{Y_1(a)}(x) - \int_{-\infty}^{-t} -t dF_{Y_1(a)}(x) \Big| \le \frac{s}{6}, \qquad \Big| \int_{-\infty}^{-t} v(x) dF_{\overline{Y}_n(a)}(x) - \int_{-\infty}^{-t} -t dF_{\overline{Y}_n(a)}(x) \Big| \le \frac{s}{6}.$$

We, therefore, have

$$\begin{aligned} \int_{\infty}^{\infty} v(x)[dF_{Y_1(a)}(x) - dF_{\overline{Y}_n(a)}(x)] &= \int_{-\infty}^{\infty} (Q(x) + v(x) - Q(x))[dF_{Y_1(a)}(x) - dF_{\overline{Y}_n(a)}(x)] = \\ s + \int_{-\infty}^{-t} (v(x) - (-t))[dF_{Y_1(a)}(x) - dF_{\overline{Y}_n(a)}(x)] &+ \int_{-t}^{\infty} (u(x) - x)[dF_{Y_1(a)}(x) - dF_{\overline{Y}_n(a)}(x)] \ge \\ \ge s - \Big| \int_{-\infty}^{-t} (v(x) - (-t)) dF_{Y_1(a)}(x) \Big| &- \Big| \int_{-\infty}^{-t} (v(x) - (-t)) dF_{\overline{Y}_n(a)}(x) \Big| - \\ \Big| \int_{-t}^{\infty} (u(x) - x) dF_{Y_1(a)}(x) \Big| &- \Big| \int_{-t}^{\infty} (u(x) - x) dF_{\overline{Y}_n(a)}(x) \Big| \ge \\ & s - 4\frac{s}{6} = \frac{s}{3}. \end{aligned}$$

Altogether, this implies that

$$Ev(Y_1(a)) \ge Ev(\overline{Y}_n(a)) + \frac{s}{3} \tag{23}$$

and as $s > 0$, we are through. ■

In light of this discussion, it is clear that in situations with many assets, or when we can assume that investors' utilities are strictly concave in the whole (efficient) support of distributional outcomes, we expect classical diversification results to hold whenever risks are bounded. However, in situations when the number of risks is not large compared with the number of assets, as defined in Theorem 1 and if utility is non-concave for large negative outcomes, then nondiversification may be optimal even with bounded risks.

3.3 When not to diversify

In this section, we further study the implications of Theorem 1, by analyzing under which conditions it will not be optimal to diversify. To calculate bounds from (4), we need bounds on $E|X|^r$, $G(\omega, z)$, and for uniformly diversified portfolios, on $F_n(z)$.

[29] Note that, due to the boundedness of the risks $Y_i(a)$, the integrals (that is, the means of the r.v.'s $Y_1(a)$ and $\overline{Y}_n(a)$) exist, as opposed to the corresponding integrals involving the risks X_1 and $\overline{X}_n$ (that is, the first moments of the r.v.'s X_1 and $\overline{X}_n$).

We assume i.i.d. risks $X_1, X_2, ..., X_n$ in $S_\alpha(\sigma, \beta, 0)$ with $\alpha \in (r, 1)$, $\beta \in [-1, 1]$ and $\sigma > 0$.[30] From Zolotarev (1986, Property 2.5, p. 63), we have that, for $X \in S_\alpha(\sigma, \beta, 0)$, $r < \alpha < 1$,

$$E|X - med(X)|^r \leq 2^{2+r/\alpha}\sigma^r \Gamma\Big(1 - \frac{r}{\alpha}\Big)\Gamma(r) \sin\Big(\frac{\pi}{2}r\Big), \tag{24}$$

where $med(X)$ denotes the median of X and $\Gamma(x) = \int_0^\infty e^{-t}t^{x-1}dt$ is the Gamma function. Furthermore, according to Theorem 2.4.2 in Zolotarev (1986), if $\alpha \in (0, 1)$, then, using the notation $Q_{\alpha,\beta,\sigma}(x)$ for $P(X > x)$,

$$Q_{\alpha,\beta,\sigma}(x) = \frac{1}{\alpha\pi}\sum_{k=1}^{\infty}(-1)^{k-1}\frac{\Gamma(k\alpha + 1)}{k\Gamma(k + 1)} \sin\Big(\frac{k\pi\alpha(1 + \beta)}{2}\Big)\frac{\sigma^{k\alpha}}{x^{k\alpha}}, \tag{25}$$

$x > 0$.

We use the fact that, by (2),

$$P\Big(w^{(1)}X_1 + w^{(2)}X_2 > z\Big) = P\Big(X_1 > \frac{z}{\big[(w^{(1)})^\alpha + (w^{(2)})^\alpha\big]^{1/\alpha}}\Big),$$

and more generally (for arbitrary nonnegative vectors summing to one, w)

$$P\Big(\sum_{i=1}^{n} w_i X_i > z\Big) = Q_{\alpha,\beta,\sigma}\Big(z/ \;\|\dot{w}\|_\alpha\Big),$$

where $\|\dot{w}\|_\alpha = \Big(\sum_{i=1}^n (w_i)^\alpha\Big)^{1/\alpha}$. Specifically, $1/\;\|\dot{\tilde{w}}_n\|_\alpha = n^{1-1/\alpha}$. Therefore, we have:

$$G(w, z) = Q_{\alpha,\beta,\sigma}\left(\frac{z}{\big[(w^{(1)})^\alpha + (w^{(2)})^\alpha\big]^{1/\alpha}}\right) - Q_{\alpha,\beta,\sigma}(z) \qquad \text{and} \qquad F_n(z) = Q_{\alpha,\beta,\sigma}\big(zn^{1-1/\alpha}\big) - Q_{\alpha,\beta,\sigma}\big(z\big), \tag{26}$$

where $Q_{\alpha,\beta,\sigma}$ is defined in (25).

Remark 12 *If we wish to introduce a time dimension, we can define the T-scaling operator: $\Lambda_T : x \mapsto Tx$. The well-known "$T^{1/2}$" rule for Brownian processes, W, implies that $W \circ \Lambda_T \stackrel{d}{=} T^{1/2} \times W$. For processes in $S_\alpha(\sigma, 0, 0)$, this generalizes to the "$T^{1/\alpha}$" rule (see, e.g. Mandelbrot 1997), i.e., for $X : \mathbf{R}_+ \to \mathbf{R}$, a stable stochastic process with $X(1) \sim S_\alpha(\sigma, 0, 0)$, we have $X \circ \Lambda_T \stackrel{d}{=} T^{1/\alpha} \times X$. Thus, for such processes properties scale-up faster over time than for Brownian processes. With this $T^{1/\alpha}$ scaling in mind, for $X_1, \ldots, X_n$ stable processes $X_i : \mathbf{R}_+ \to \mathbf{R}$ and $X_i(t) \in S_\alpha(t^{1/\alpha}\sigma, 0, 0)$, we can define the truncated processes $X_i^a(T) = X_i(T)$, if $|X_i(T)| \leq aT^{1/\alpha}$, $X_i^a(T) = aT^{1/\alpha}$ if $X_i > aT^{1/\alpha}$ and $X_i^a(T) = -aT^{1/\alpha}$ if $X_i < -aT^{1/\alpha}$. With these definitions, it is clear that σ changes to $(T_2/T_1)^{1/\alpha}\sigma$ in equations (24)-(26) when going from time-scale T_1, to time-scale T_2.*

We first study the symmetric case, i.e., the case when $\beta = 0$. For simplicity, we begin with the case when there are two assets, $n = 2$, and study how a depends on $w^{(1)}$ (and $w^{(2)} = 1 - w^{(1)}$). In this case, the analogue of equation (24) is (see the argument for Property 2.5 in Zolotarev, 1986)

$$E|X|^r \leq 2\sigma^r \Gamma\Big(1 - \frac{r}{\alpha}\Big)\Gamma(r) \sin\Big(\frac{\pi}{2}r\Big). \tag{27}$$

[30]For notational convenience, we restrict ourselves to the case with the location parameter $\mu = 0$. All results are trivially extended to the general case by the translation $X_\mu = X + \mu$, leading to $med(X_\mu) = med(X) + \mu$, $P(X_\mu > z) = P(X > z - \mu)$, etc.

Furthermore, the asymptotic expansion (25) implies the following bounds for the tail of $Q_{\alpha,0,\sigma}$:

$$\frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\left(\frac{\pi\alpha}{2}\right)\frac{\sigma^\alpha}{x^\alpha} - \frac{1}{\alpha\pi}\frac{\Gamma(2\alpha+1)}{4}\sin(\pi\alpha)\frac{\sigma^{2\alpha}}{x^{2\alpha}} < Q_{\alpha,0,\sigma}(x) < \frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\left(\frac{\pi\alpha}{2}\right)\frac{\sigma^\alpha}{x^\alpha}. \tag{28}$$

Using (28) for $G(w,z)$, we get

$$G(w,z) > \frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\left(\frac{\pi\alpha}{2}\right)\frac{\sigma^\alpha}{z^\alpha}\Big((w^{(1)})^\alpha + (w^{(2)})^\alpha - 1\Big) - \frac{1}{\alpha\pi}\frac{\Gamma(2\alpha+1)}{4}\sin(\pi\alpha)\frac{\sigma^{2\alpha}\left[(w^{(1)})^\alpha + (w^{(2)})^\alpha\right]^2}{z^{2\alpha}}. \tag{29}$$

Using bounds (4), (27) and (29) we get that Theorem 1 holds with the following easy to compute estimate for the length of the distribution support:

$$\tilde{a} = \frac{z^{\alpha/r}(\alpha\pi)^{1/r}\sigma^{(r-\alpha)/r}\Big(\Gamma\big(1-\frac{r}{\alpha}\big)\Gamma(r)\sin\big(\frac{\pi}{2}r\big)\Big)^{1/r}(n-1)^{1/r}}{\left[\Gamma(\alpha+1)\sin\left(\frac{\pi\alpha}{2}\right)\Big((w^{(1)})^\alpha + (w^{(2)})^\alpha - 1\Big) - \frac{\Gamma(2\alpha+1)}{4}\sin(\pi\alpha)\frac{\sigma^\alpha\big((w^{(1)})^\alpha+(w^{(2)})^\alpha\big)^2}{z^\alpha}\right]^{1/r}}. \tag{30}$$

Thus, $\tilde{a}$ as a function of $w^{(1)}$ provides a sufficient condition for diversification into (w_1, w_2) not being preferred to holding one asset.

In Figure 4, we plot the relationship between $\tilde{a}$ and $w^{(1)}$ for different value at risk and $\sigma = 1$. We see that the bound is fairly constant for $w^{(1)}$, except close to 1 (corresponding to an almost undiversified portfolio) where it rapidly grows. We will therefore compare uniform diversification portfolios going forward. Also, clearly a larger bound is needed for a smaller q (that is, for larger $z = VaR_q(X_1)$). This comes as no surprise, as a smaller q implies that the VaR inequality must hold further out in the tail, which pushes a upward.

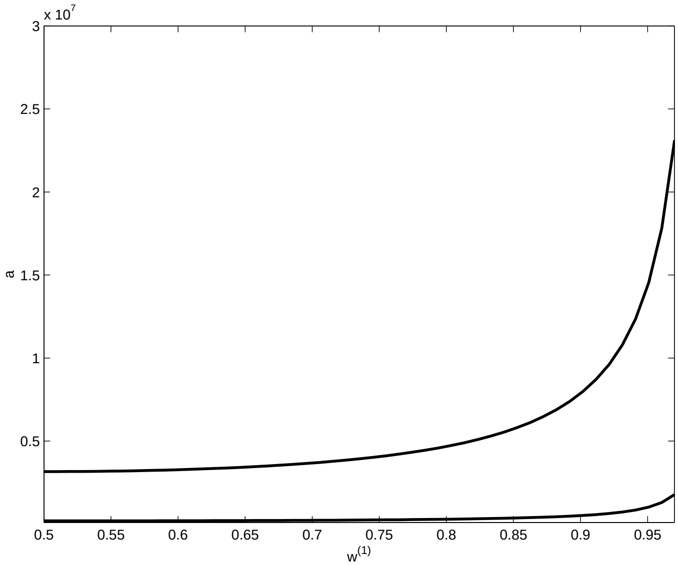

Figure 4: *Relationship between distributional support, $\tilde{a}$, and level of diversification, $w^{(1)}$, for VaR centiles $q = 0.01$ (below) and $q = 0.001$ (above). Parameters: $n = 2$, $\sigma = 1$, $\beta = 0$.*

We next generalize to arbitrary $n \geq 2$, and σ, keeping $\beta = 0$ and fixing $\alpha = 0.85$. We study when holding one risk dominates uniform diversification, i.e., we study a as a function of n, and q (where value at risk is

$z = VaR_q(X_1)$) for (4) to be satisfied is satisfied for $\tilde{w}_n$: $a(n, VaR_q(X_1))$. However, we normalize to $A(n,q) = a(n, VaR_q(X_1))/VaR_q(X_1)$, i.e., for a given percentile, q, the required a as a factor of the value at risk for the untruncated distribution. This normalization is natural as, given the VaR chosen, it is the number of times this level that is the worst possible outcome. The advantage of this normalization is that it is scale free: it holds for arbitrary σ.[31] We use the exact formulae in (24-26). For $\alpha = 0.85$, the results are shown in Table 1. A general

n	2	5	8	11	14	17	20	23	26	29	32	35
q												
0.1	2,288	4,404	6,371	8,279	10,158	12,018	13,868	15,710	17,549	19,385	21,221	23,056
0.05	2,226	4,222	6,056	7,821	09,547	11,248	12,933	14,604	16,267	17,922	19,572	21,218
0.02	2,419	4,560	6,516	8,394	10,225	12,027	13,807	15,571	17,322	19,064	20,799	22,526
0.01	2,672	5,028	7,179	9,240	11,251	13,226	15,178	17,111	19,030	20,938	22,837	24,728
0.005	2,934	5,517	7,874	10,133	12,335	14,499	16,636	18,752	20,853	22,941	25,019	27,088
0.025	3,254	6,118	8,730	11,232	13,671	16,068	18,435	20,779	23,106	25,418	27,719	30,010
0.001	3,691	6,938	9,899	12,736	15,500	18,217	20,899	23,556	26,192	28,813	31,420	34,017
0.0005	4,133	7,768	11,080	14,260	17,355	20,396	23,399	26,373	29,325	32,258	35,177	38,083

Table 1: *Threshold for $A = a/VaR_q(X_1)$, above which diversification is sub-optimal as a function of q and number of risks, n.* $\alpha = 0.85$, $\beta = 0$.

conclusion is that the worst case scenario must be a lot worse than the VaR level chosen, for diversification being inferior. For example, with a value at risk corresponding to $q = 1\%$, the worst case scenario must be almost 2,700 times VaR_q for diversification into 2 assets to be clearly inferior, and the factor increases almost linearly in the number of assets. This might be taken as an indication that the types of limits of diversification discussed in this paper only arises in quite extreme situations, even when distributional support is bounded. We caution against this conclusion for two reasons. First, equation (4) only gives a sufficient condition for diversification to be suboptimal and, in fact, uses rough bounds (Chebyshev's inequality for the marginal distributions). The true A may therefore be considerably smaller. Second, so far, we have for tractability only studied the strongest case for diversification, namely the case with i.i.d. risks. According to the results presented in Section 4, diversification also breaks for a wide range of bounded risks that exhibit dependence modelled by convolutions of $\alpha-$symmetric distributions with $\alpha < 1$, in particular, for certain models with common shocks (see Remark 13). We conjecture, however that, comparing to the independent case, the smallest length of distributional support required for diversification failure is considerably smaller for these as well as for other types of dependence.

Finally, we generalize to the case $\beta \neq 0$. Equation (27) and the right-hand-side inequality in (28) implies the following bound for the median $med(X)$ of a r.v. $X \sim S_\alpha(\sigma, \beta, 0)$:

$$|med(X)| \leq 2^{1/\alpha}\sigma\Big(\frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\Big(\frac{\pi\alpha(1+\beta)}{2}\Big)\Big)^{1/\alpha}.$$

This and (24) imply that

$$E|X|^r \leq 2^{r/\alpha}\sigma^r\Big(\frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\Big(\frac{\pi\alpha(1+\beta)}{2}\Big)\Big)^{r/\alpha} + 2^{2+r/\alpha}\sigma^r\Gamma\Big(1-\frac{r}{\alpha}\Big)\Gamma(r)\sin\Big(\frac{\pi}{2}r\Big). \tag{31}$$

[31] Therefore, it also holds for arbitrary time scales, T, according to our previous discussion.

Similar to (29), we obtain that, in the general case of skewed stable distributions,

$$G(w,z) > \frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\Big(\frac{\pi\alpha(1+\beta)}{2}\Big)\frac{\sigma^\alpha}{z^\alpha}\Big((w^{(1)})^\alpha + (w^{(2)})^\alpha - 1\Big) - \frac{1}{\alpha\pi}\frac{\Gamma(2\alpha+1)}{4}\sin\big(\pi\alpha(1+\beta)\big)\frac{\sigma^{2\alpha}\big[(w^{(1)})^\alpha + (w^{(2)})^\alpha\big]^2}{z^{2\alpha}}. \tag{32}$$

Using bounds (31) and (32), from Remark 6 we obtain that, in the case of general skewed stable risks $X_i \sim S_\alpha(\sigma,\beta,0)$, Theorem 1 holds with the following easy to compute estimate for the length of the distribution support:

$$\tilde{a} > \frac{2^{1/r}z^{\alpha/r}(\alpha\pi)^{1/r}\sigma^{(r-\alpha)/r}\Big(\Big(\frac{1}{\alpha\pi}\Gamma(\alpha+1)\sin\big(\frac{\pi\alpha(1+\beta)}{2}\big)\Big)^{r/\alpha} + 4\Gamma\big(1-\frac{r}{\alpha}\big)\Gamma(r)\sin\big(\frac{\pi(1+\beta)}{2}r\big)\Big)^{1/r}(n-1)^{1/r}}{\Big[\Gamma(\alpha+1)\sin\big(\frac{\pi\alpha(1+\beta)}{2}\big)\Big((w^{(1)})^\alpha + (w^{(2)})^\alpha - 1\Big) - \frac{\Gamma(2\alpha+1)}{4}\sin\big(\pi\alpha(1+\beta)\big)\frac{\sigma^\alpha\big((w^{(1)})^\alpha+(w^{(2)})^\alpha\big)^2}{z^\alpha}\Big]^{1/r}}. \tag{33}$$

The same type of analysis as for the case with $\beta = 0$ could now be carried out for general β's. This concludes our numerical treatment of the limits of diversification for bounded risks.

3.4 Potential implications for financial risks

In this section we briefly discuss two puzzling facts in markets for risky assets, for which the limits of diversification may play a role. An in-depth analysis may be a topic for future research.

First, several facts for markets for catastrophe (re)-insurance may be explained by the limits of diversification: For insurance companies, risk sharing and diversification can be achieved by selling and buying reinsurance in the reinsurance market. Froot (2001) notes the puzzling fact that insurance companies seem to limit their reinsurance and that their degree of reinsurance actually is decreasing in the size of the risk. This is contrary to the implications in Froot, Scharsfstein and Stein (1993) that the largest risks should be reinsured first and most extensively.

An explanation is suggested by lines B-C in Figure 1. If there is a critical size of risks beyond which it is not optimal to diversify, then insurance companies might choose to reinsure smaller risks (corresponding to either a medium-large n on line B, or a large n on line C), whereas, larger risks (corresponding to small-medium n on line C) will not be reinsured.

Second, limits of diversification might have implications for the so-called under-diversification puzzle: Traditional asset pricing theory suggests that investors should invest in the market portfolio, i.e., hold a fraction in each asset (Sharpe 1964, Lintner 1964, Merton 1973). Brennan (1975) noted that the presence of transaction costs leads to less diversified portfolios being optimal. However, the portfolios held by investors seem to be much less diversified than what would motivated by transaction costs (Blume and Friend (1975), Barber and Odean (2000)). This is the under-diversification puzzle. The puzzle is further analyzed in Goetzmann and Kumar (2005).

Several explanations for the under-diversification puzzle have been suggested. Uppal and Wang (2001) and more recently Klibanoff, Marinacci and Mukerji (2005) suggest that the presence of ambiguity averse investors with non-expected utility preferences may lead to low diversification. Bounded rationality in the form of limited information processing ability in combination with learning is suggested as an explanation in Van Nieuwerburgh and Veldkamp (2004). Behavioral explanations for under-diversification have also been suggested. For example, Cao, Hirshleifer and Zhang (2004) suggest that preference for status quo driven by fear of the unfamiliar may lead to under-diversification. Over-confident investors may choose to hold focused portfolios (Odean, 1999), as may investors who suffer from narrow framing (Barberis, Huang and Thaler, 2003).

The analysis in this paper may suggest an additional explanation. From Theorem 1, we know that there are distributions with bounded support for which diversification increases the probabilities of disaster up to a certain number of stocks, n. Even if the number of stocks in the market is larger than n, transaction costs may bring this number down. The number of stocks needed for diversification to be preferred may be so large that transaction costs make such portfolios sub-optimal, even though it is not large enough in absence of transaction costs. The idea is illustrated in Figure 5.

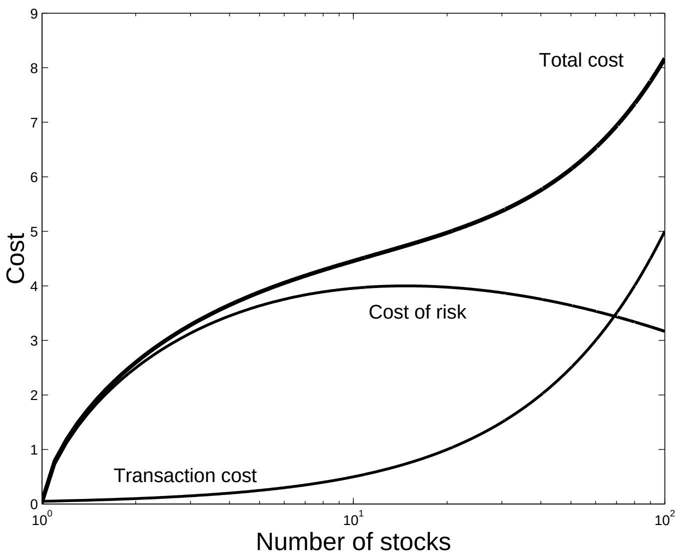

Figure 5: *Illustrative example: Diversification with bounded risks and transaction costs. Total cost increasing function in number of stocks so it is optimal to hold focused portfolio.*

4 Generalizations to dependence and non-identical distributions

As indicated in Subsection 1.3 in the introduction, the results obtained in this paper continue to hold for wide classes of bounded dependent and non-identically distributed risks. More precisely, the results continue to hold for convolutions of r.v.'s with joint truncated $\alpha-$symmetric and spherical distributions and their non-identically distributed versions as well as for a wide class of models with common shocks.

According to the definition introduced by Cambanis, Keener and Simons (1983), an $n-$dimensional distribution is called $\alpha-$symmetric if its characteristic function (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 $\alpha > 0$. An important property of $\alpha-$symmetric distributions is that, similar to strictly stable laws, they satisfy property (2). 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 distributions corresponding to the case $\alpha = 2$ (see Fang, Kotz and 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 et. al., 1990, Ch. 3). Spherically symmetric stable distributions have characteristic functions $\exp\big[-\lambda\big(\sum_{i=1}^n t_i^2\big)^{\gamma/2}\big]$, $0 < \gamma \le 2$, and are, thus, examples of $\alpha-$symmetric distributions with $\alpha = 2$ and the c.f. generator $\phi(x) = \exp(-x^\gamma)$.

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

$$(Q_1, ..., Q_n) = (ZY_1, ..., ZY_n), \tag{34}$$

where $Y_i \sim S_\alpha(\sigma, 0, 0)$ are i.i.d. symmetric stable r.v.'s with $\sigma > 0$ and the index of stability α and $Z \ge 0$ is a nonnegative r.v. independent of $Y_i's$ (see Bretagnolle, Dacuhna-Castelle and Krivine, 1966, and Fang et. al., 1990, p. 197). In the case $Z = 1$ (a.s.), model (34) represents vectors with i.i.d. symmetric stable components that have c.f.'s $\exp\big[-\lambda\sum_{i=1}^n |t_i|^\alpha\big]$ which are particular cases of c.f.'s of $\alpha-$symmetric distributions with the generator $\phi(x) = \exp(-\lambda x^\alpha)$.

The dependence structures considered in this section include, among others, convolutions of models (34). That is, the dependence structures cover vectors $(X_1, ..., X_n)$ which are sums of i.i.d. random vectors $(Z_jV_{1j}, ..., Z_jV_{nj})$, $j = 1, ..., k$, where $V_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $i = 1, ..., n$, $j = 1, ..., k$, and Z_j are positive absolutely continuous r.v.'s independent of V_{ij}:

$$(X_1, ..., X_n) = \sum_{j=1}^k (Z_jV_{1j}, ..., Z_jV_{nj}). \tag{35}$$

Although the dependence structure in model (34) alone is restrictive, convolutions (35) 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, 2003).

According to the results in Bretagnolle et. al. (1966) and Kuritsyn and Shestakov (1984), the function $\exp\big(-\big(|t_1|^\alpha + |t_2|^\alpha\big)^{1/\alpha}\big)$ is a c.f. of two $\alpha-$symmetric r.v.'s for all $\alpha \ge 1$ (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 Fang et. al. (1990, Ch. 7) and Gneiting (1998).

Convolutions of $\alpha-$symmetric distributions are symmetric and unimodal. These convolutions also 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.[32] On the other hand, convolutions of $\alpha-$symmetric distributions with $1 < \alpha \le 2$, and, in particular, convolutions of models (34) with $1 < \alpha \le 2$, can have marginals with power moments finite up to a certain positive order (or finite exponential moments) depending on the choice of the r.v.'s Z. For instance, convolutions of models (34) with $1 < \alpha < 2$ and $E|Z| < \infty$ have finite means but infinite variances, however, marginals of such convolutions have infinite means

[32] This is true because if one assumes that r.v.'s $X_1, ..., X_n$, $n \ge 2$, have an $\alpha-$symmetric distribution with $\alpha < 1$ and that $E|X_i| < \infty$, $i = 1, ..., n$, then, by the triangle inequality, $E|X_1 + ... + X_n| \le E|X_1| + ... + E|X_n| = nE|X_1|$. The latter, however, cannot hold since, according to (2), $(X_1 + ... + X_n) \sim n^{1/\alpha}X_1$ 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.

if the r.v.'s Z satisfy $E|Z| = \infty$. Moments $E|ZY_i|^p$, $p > 0$, of marginals in models (34) 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 (34) with $\alpha = 2$ are finite if $E|Z|^p < \infty$ for all $p > 0$. Similarly, marginals of spherically symmetric (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 provide one of now well-established approaches to modeling heavy-tailedness phenomena with moments up to some order (see Praetz, 1972, Blattberg and Gonedes, 1974, and Glasserman et. al., 2002).

Using the argument similar to that in the proof of Theorem 1 and the results on the value at risk under heavy-tailedness and dependence obtained in Ibragimov (2004a, b, 2005), we obtain the following Theorem 4 that provides precise formulation of the extensions of the results in Subsection 3.1 to the dependent case. According to Theorem 4, the results provided by Theorem 1 for convolutions of truncated i.i.d. stable distributions with indices of stability $\alpha < 1$ continue to hold for convolutions of truncated $\alpha-$symmetric distributions with $\alpha < 1$ that have, as indicated above, extremely thick-tailed marginals with infinite first moments. In particular, Theorem 1 continues to hold for convolutions of truncated analogues of models (34) with common shocks affecting all thick-tailed risks Y_i with tail indices $\alpha < 1$.

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 4 *Theorem 1 continues to hold if any of the following is satisfied:*

The vector of r.v.'s $(X_1, ..., X_n)$ entering its assumptions is a sum of i.i.d. random vectors $(V_{1j}, ..., V_{nj})$, $j = 1, ..., k$, where $(V_{1j}, ..., V_{nj})$ 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 vector of r.v.'s entering the assumptions of the results is a sum of i.i.d. random vectors $(Z_jV_{1j}, ..., Z_jV_{nj})$, $j = 1, ..., k$, where $V_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $i = 1, ..., n$, $j = 1, ..., k$, with $\sigma_j > 0$ and $\alpha_j \in (0, 1)$ and Z_j are positive absolutely continuous r.v.'s independent of V_{ij}.

Proof. Theorem 4 can be derived similar to the proof of Theorem 1 using the fact that, according to the results in Ibragimov (2004a, b, 2005), Proposition 1 continues to hold for vectors of r.v.'s $(X_1, ..., X_n)$ in its assumptions.

Remark 13 *One should emphasize that the values of the interval length, a that lead to the breakdown of the stylized facts on diversification for dependent risks considered in the present section are not greater than those in the case of independence in Section 3. We conjecture, however, that the length of the distributional support required for diversification failure in the case of dependence is, in fact, considerably smaller than that in the independent setting.*

As for generalizations of the results in the paper to the case of non-identical distributions, the following conclusions hold: 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 non-identically distributed stable risks. Using the arguments in this paper together with the fact that, according to the results in Ibragimov (2004a, b, 2005), Proposition 1 holds for risks $X_1, ..., X_n$ if $\sigma_n \geq ... \geq \sigma_1 \geq 0$, we obtain that Theorem 1 is valid under such assumptions.

5 Concluding remarks

We have analyzed the limits of diversification for bounded risks with heavy tails in their support. Our results show that portfolio diversification may not be preferable for a wide class of such bounded risks. Also, value at risk coherency may be violated for this class.

The key parameters for our analysis are the number of risks available, the thickness of the tails and the support of the distributions. If the effective support is large compared with the number of risks, nondiversification may be optimal. This demonstrates that "unpleasant" properties of the value at risk as a risk measure under heavy-tailedness does not arise from the high likelihood of getting very large losses but rather from the fact that there are too few securities available for diversification to work.

The theory can be related to the expected utility model. We show that if there is a point arbitrary far out in the domain of losses beyond which the utility function is not concave, then nondiversification may be optimal also from an expected utility perspective.

Our results suggest that the distributional assumption of unbounded heavy tails may be treated as an appropriate approximation in some situations even though the distributional support may be bounded. In many real world applications, distributions may be bounded, the expected utility specification of investor behavior only makes sense over reasonable domains and the number of assets is finite. Which approximation is most appropriate must then depend on the situation at hand.

Finally, the paper essentially accomplishes the unification of the analysis of robustness of value at risk models to such important distributional phenomena as boundedness, dependence, skewness and the case of non-identical marginals.

The analysis in this paper may help explain observed low diversification in markets where losses may be large. We mention two examples: the low levels of reinsurance in markets for catastrophe reinsurance and the under-diversification puzzle in stock markets. These may be topics for future research. ■

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. and Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance* **26**, 1487-1503.

Adler, R. J., Feldman, R. E. and Taqqu, M. S., eds. (1998). *A practical guide to heavy tails. Statistical techniques and applications.* Boston: Birkhauser.

Andrews, D. W. K. (2003). Cross-section regression with common shocks. *Cowles Foundation Discussion Paper No. 1428*, Yale University.

Artzner, P., Delbaen, F., Eber, J.-M. and 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.

Barber, B. M., and Odean, T. (2000) Trading is hazardous to your wealth: The common stock investor performance of individual investors. *Journal of Finance* **55**, 776-806.

Barberis, N., Huang M., and Thaler R. H. (2003). Individual preferences, monetary gambles and the equity premium. Working paper, University of Chicago, Graduate School of Business.

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

Blattberg, R. C. and Gonedes, R. C. (1974). Comparison of stable and student distributions as statistical models for stock prices. *Journal of Business* **47**, 244-280.

Blume, M. E. and Friend I. (1975). The asset structure of individual portfolios and some implications for utility functions, *Journal of Finance*, 585-603.

Brennan, M. (1975). The optimal number of securities in a risky asset portfolio when there are fixed costs of transacting: Theory and some empirical results. *Journal of Financial and Quantitative Analysis* **10**, 483-496.

Bretagnolle, J., Dacunha-Castelle, D. and Krivine, J. L. (1966). Lois stables et espaces L^p. *Annales de l'Institute H. Poincaré. Section B. Calcul des Probabilités et Statistique.* **2**, 231-259.

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

Cao, H. H., Hirshleifer, D., and Zhang, H. H. (2004). Fear of the unknown: familiarity and economic decisions. Working paper, University of North Carolina, Kenan-Flagler Business school.

Cotter, D. and 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 la Peña, V. H. and Ibragimov, R. (2003). Sharp probability inequalities and conservative testing procedures for Studentized processes and moving averages with applications to econometric models and heavy tails. Mimeo, Yale University. Available at http://post.economics.harvard.edu/faculty/ibragimov/Papers/Selfnorm.pdf

de la Peña, V. H., Ibragimov, R. and Jordan, S. (2004). Option bounds. *Journal of Applied Probability* **41A**, 145-156.

De Vany, A. S. and Walls, W. D. (2004). Motion picture profit, the stable Paretian hypothesis and the curse of the superstar. *Journal of Economic Dynamics and Control* **28**, 1035-1037.

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

Duffie, D. and Pan, J. (1997). An overview of value at risk. *Journal of Derivatives* **4**, 7-49.

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

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

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

Fama, E. (1965). The behavior of stock market prices. *Journal of Business* **38**, 34-105.

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

Froot, K. A. (2001). The market for catastrophe risk: a clinical examination. *Journal of Financial Economics*, **60**, 529-571.

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

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

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

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

Goetzmann W. N. and Kumar A. (2005). Why do individual investors hold under-diversified portfolios, Yale School of Management Working paper 454.

Gollier, C., Koehl, P. F. and Rochet, J. C. (1997) Risk-taking behavior with limited liability and risk aversion. *Journal of Risk and Insurance*, **64**, 347-370.

Ibragimov, R. (2004a). 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 (2004b). Portfolio diversification and value at risk under thick-tailedness. Harvard Institute of Economic 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. and Brown, D. (2005). Sign tests for dependent observations and bounds for path-dependent options. Harvard Institute of Economic Research Discussion Paper #2099. http://post.economics.harvard.edu/hier/2006papers/HIER2099.pdf

Ingersoll, J. E. (1987). *Theory of financial decision making.* Totowa, N.J.: Rowman & Littlefield.

Jansen, D. W. and 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, M. C. and Meckling, W. H. (1976) Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics* **3**, 305-360.

Kahneman, D. and Tversky, A. (1979) Prospect theory: An analysis of decision under risk. *Econometrica* **47**, 263-292.

Klibanoff, P., Marinacci, M., and Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica, forthcoming.*

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

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics* **47**, 13-37

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

Lukacs, E. (1975). *Stochastic Convergence, 2nd edition*, New York: Academic Press.

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.* New York: Springer-Verlag.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance* **7**, 77-91.

McCulloch, J. H. (1996). Financial applications of stable distributions. In: *Handbook of statistics* (G. S. Maddala and C. R. Rao, eds.), Vol. 14 (Amsterdam: Elsevier), 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. and Scheffler, H.-P. (2000). Portfolio modeling with heavy tailed random vectors. In: *Handbook of Heavy-Tailed Distributions in Finance* (S. T. Rachev, ed.), Elsevier North-Holland, New York, 595-640.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica* **41** 373-413.

Odean, T. (1999). Do investors trade too much? *American Economic Review* **89**, 1279-1298.

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

Rachev, S. T., Menn, C. and 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. and Mittnik, S. (2000). *Stable Paretian Models in Finance.* New York: Wiley.

Rothschild, M. and Stiglitz, J.E. (1970). Increasing Risk: I. A definition. *Journal of Economic Theory* **2**, 225-243.

Roy, A. D. (1952). Safety-First and the Holding of Assets. *Econometrica* **20**, 431449.

Samuelson, P. A. (1967). General proof that diversification pays. *The Journal of Financial and Quantitative Analysis* **2** 1-13.

Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic orders and their applications.* Boston: Academic Press.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance* **19**, 425-442.

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

Silverberg, G. and 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.*

Stiglitz, J. E. (1974). Incentives and risk sharing in sharecropping. *Review of Economic Studies* **41**, 219-255.

Stiglitz, J. E., Weiss A. (1981). Credit rationing in markets with imperfect information. *American Economic Review* bf 71, 393-410.

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

Uppal, R. and Wang, T. (2002). Model Misspecification and under-diversification. CEPR Discussion Paper No 3304.

Van Nieuwerburgh, S. and Veldkamp, L. (2004) Information acquisition and portfolio under-diversification. Working paper New York University, Stern School of Business.

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.* Providence: American Mathematical Society.