

THE LIMITS OF DIVERSIFICATION WHEN LOSSES MAY BE LARGE*

Rustam Ibragimov^{a1}, Johan Walden^{b 2}

^a *Department of Economics, Harvard University, Cambridge, MA 02138, USA*

^b *Haas School of Business, University of California at Berkeley, Berkeley, CA 94720, USA*

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 markets for risky assets where the number of assets is limited compared with the (bounded) distribution 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.

JEL classification: G11

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

* An extended working paper version of this paper is available as Ibragimov and Walden (2006). The authors are grateful to Gary Chamberlain and Oliver Hart for helpful comments and for posing questions about the stylized facts on diversification under boundedness assumptions that are answered in this paper. We also thank Donald Andrews, Donald Brown, Xavier Gabaix, Tilmann Gneiting, Dwight Jaffee, Peter Phillips, Jacob Sagi, Herbert Scarf, Martin Weitzman 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. Finally, we thank a referee for helpful comments and suggestions.

¹Tel.: +1 617 496 4795; fax: +1 617 495 7730. *Email address:* ribragim@fas.harvard.edu

²Tel.: +1 510 643 0547; fax: +1 510 643 1420. *Email address:* walden@haas.berkeley.edu

1 Introduction

1.1 Background

Recent research has shown that, under general conditions, the stylized fact of portfolio diversification always being preferable is reversed for extremely heavy-tailed risks, with infinite first moments and unbounded distribution support. Ibragimov (2004a, b, 2005, 2007) developed a unified approach to analyzing portfolio theory for such heavy-tailed risks and a number of other problems in economics using new majorization theory for linear combinations of thick-tailed random variables (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, providing alternatives to the traditional expected utility framework (see, e.g., the papers in Szegö, 2004, and Fabozzi, Focardi and Kolm, 2006, for a review of properties of value at risk and other measures of risk and Roy's, 1952, safety-first approaches to portfolio selection). 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.³

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

³Several 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 (see Artzner, Delbaen, Eber and Heath, 1999, and Szegö, 2004, for the definition of coherency for measures of risk and an overview of its implications). 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 existence of means of the risks in considerations is also required for finiteness of coherent spectral measures of risk (see Acerbi, 2002, and Cotter and Dowd, 2006) that generalize the expected shortfall.

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 and 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* risks⁴ 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.

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 setting⁵, 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,⁶ then our non-diversification results may continue to hold (Theorem 3).

⁴We will, somewhat contradictory, refer to distributions of such risks as *bounded heavy-tailed distributions* as opposed to the standard (unbounded) heavy-tailed distributions.

⁵As originally shown in a general setting in Samuelson (1967).

⁶Convexity 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.

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. 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 α -symmetric distributions and their analogues with non-identical marginals (Theorem 4).

1.3 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 Fama, 1965, and the papers in Mandelbrot, 1997), 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, then⁷

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

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), Gabaix, Gopikrishnan, Plerou and Stanley (2003), Rachev, Menn and Fabozzi (2005) and references therein. Specifically,

⁷Here and throughout the paper, $f(x) \sim g(x)$ means that $0 < c \leq f(x)/g(x) \leq C < \infty$ for large x , for constants c and C .

Chapter 11 in Rachev, Menn and Fabozzi (2005) discusses and reviews the vast literature that supports heavy-tailedness and the stable Paretian hypothesis (with $1 < \alpha < 2$) for equity and bond return distributions. The following estimates of the tail parameters α for returns on various stocks and stock indices were reported in the literature using different models and statistical techniques: $3 < \alpha < 5$ (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); $\alpha \approx 3$ (Gabaix et al., 2003).

As discussed in, e.g., Lux (1996), Guillaume, Dacorogna, Dave et al. (1997) and Gabaix et al. (2003), tail exponents are similar for financial and economic time series in different countries. Some studies have indicated that the tail exponent is close to one or slightly less than one for such financial time series as Bulgarian lev/US dollar exchange spot rates and increments of the market time process for Deutsche Bank price record (see Rachev and Mittnik, 2000). As discussed by Nešlehova, Embrechts and Chavez-Demoulin (2006), tail indices less than one are observed for empirical loss distributions of a number of operational risks. Furthermore, Scherer, Harhoff and Kukies (2000) and Silverberg and Verspagen (2004) report the tail indices α to be considerably less than one for financial returns from technological innovations.

The fact that a number of economic and financial time series have tail exponents of approximately equal to or 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.

1.4 Organization of paper

The paper is organized as follows: Section 2 contains notation 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 of diversification of bounded risks on portfolio riskiness. We also relate our results to the expected utility/stochastic dominance framework and provide tables for when it is optimal not to diversify. 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

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

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

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 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 ($\alpha = 2$). If $X_i \sim S_\alpha(\sigma, \beta, \mu)$,

$\alpha \in (0, 2]$, are i.i.d. stable r.v.'s such that $\mu = 0$ for $\alpha \neq 1$ and $\beta = 0$ for $\alpha = 1$, then, for all $a_i \geq 0$, $i = 1, \dots, n$,

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

For $0 \leq r < 1$, 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 $\sigma > 0$. That is, $\mathcal{CS}(r)$ consists of distributions of r.v.'s X for which, with some $k \geq 1$, $X = Y_1 + \dots + Y_k$, where Y_i , $i = 1, \dots, k$, are independent r.v.'s such that $Y_i \sim S_{\alpha_i}(\sigma_i, 0, 0)$, $\alpha_i \in (r, 1)$, $\sigma_i > 0$, $i = 1, \dots, k$. 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 $X \stackrel{d}{=} Y$ if the distributions of X and Y are the same. In addition, $I(\cdot)$ stands for the indicator function. We define the a -truncated version of a r.v.: $Y(a) = X$ if $|X| \leq a$, $Y(a) = -a$ if $X < -a$ and $Y(a) = a$ if $X > a$. In other words, $Y(a) = a \cdot \text{sign}(X_i) + XI(|X| \leq a)$.⁸ We will also use the notation X^a instead of $Y(a)$ for the a -truncated version of X .

3 Main results: Limits of diversification

3.1 Non-diversification for risks with bounded support

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

⁸This 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 $XI(|X| \leq a)$ which move probability mass to the center. However, this is not true for the results in Section 3.2.

$z > 0$.⁹ Furthermore, given a 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.¹⁰

Throughout this section, X_1, X_2, \dots 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 $\mathcal{I}_n = \{w = (w_1, \dots, w_n) \in \mathbf{R}_+^n : \sum_{i=1}^n w_i = 1\}$. For $w \in \mathcal{I}_n$, denote by X_w the return on the portfolio of risks X_1, \dots, 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 $Y_1(a), \dots, Y_n(a)$ 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, \dots, X_n with equal weights $\tilde{w}_n = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ 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, \dots, X_n or $Y_1(a), \dots, Y_n(a)$ consist in minimizing, respectively, the disaster probabilities $P\left(\sum_{i=1}^n w_i X_i > z\right)$ or $P\left(\sum_{i=1}^n w_i Y_i(a) > z\right)$ over the portfolio weights $w \in \mathcal{I}_n$. Let $w_{[1]} \geq \dots \geq w_{[n]}$ 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 $(1, 0, \dots, 0)$. 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 $(w^{(1)}, w^{(2)}) = (\max[0.5, w_{[1]}], \min[0.5, 1 - w_{[1]}])$.

As shown 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.¹¹ In such a setting, diversification of a portfolio of the risks increases the probability of

⁹In 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 et al. (1999) who interpret *negative* values of risks as losses.

¹⁰That is, in the case of an absolutely continuous risk Z , $P(Z > VaR_q[Z]) = q$.

¹¹The 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 $w = (w_1, w_2, w_3, \dots, w_n)$ is majorized by (that is, has less diverse or more nearly equal components than) the vector $(w^{(1)}, w^{(2)}, 0, \dots, 0)$ which is, in turn, is majorized by the vector $(1, 0, 0, \dots, 0)$.

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 $X_i, i = 1, \dots, n$, are i.i.d. risks such that $X_i \sim \mathcal{CS}(r)$, for some $r \in (0, 1)$, $i = 1, \dots, n$. Then, for all $z > 0$, $P(X_w > z) > P(w^{(1)}X_1 + w^{(2)}X_2 > z) > P(X_1 > z)$.*

Remark 1 *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 $X_i, i = 1, \dots, n$, are i.i.d. risks such that $X_i \sim \mathcal{CS}(r)$, for some $r \in (0, 1)$, $i = 1, \dots, n$. Then, for all loss probabilities $q \in (0, 1/2)$, the return X_w on the portfolio of risks X_1, \dots, 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). The following theorem is the analogue of Proposition 1 in the case of *bounded* risks. The theorem shows that diversification continues to be disadvantageous for truncated extremely heavy-tailed distributions. The results demonstrate, in particular, that for any number

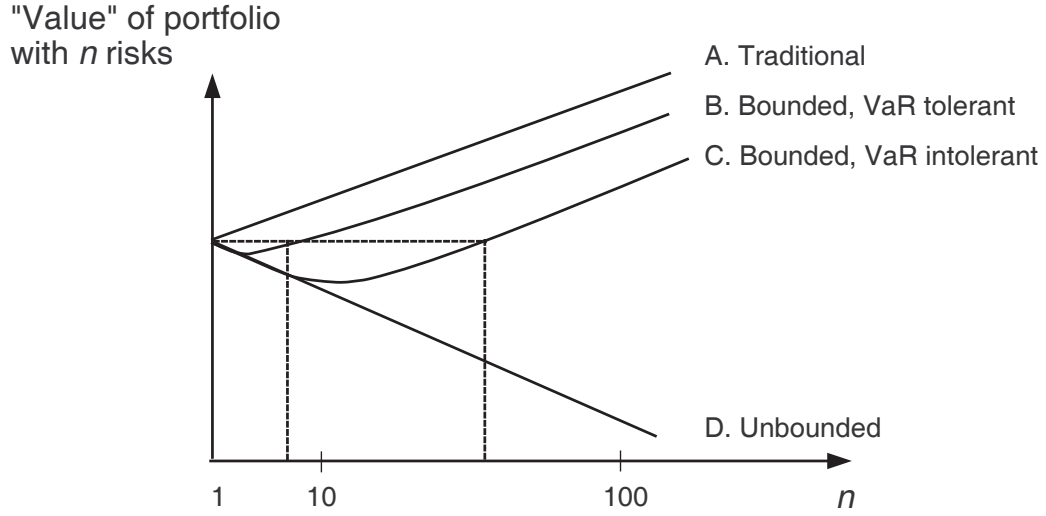


Figure 1: *Illustrative figure of value of diversification. A: Traditional situation ($\alpha > 1$). The value increases monotonically with the number of risks, n , and it is always preferable to add another risk to portfolio. B: New situation ($\alpha < 1$). Bounded heavy-tailed distributions with VaR 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.*

$n \geq 2$ and any given disaster level $z > 0$, there exist n risks with *finite* support with the property that a diversified portfolio is riskier than a portfolio consisting of only one risk.

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

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

which is positive if $w_{[1]} \neq 1$ since, by Proposition 1 applied to the portfolio of risks X_1, X_2 with weights $(w^{(1)}, w^{(2)})$, $P\left(w^{(1)}X_1 + w^{(2)}X_2 > z\right) > P\left(X_1 > z\right)$ 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 $a > \left(\frac{E|X_1|^r(n-1)}{2G(w, z)}\right)^{1/r}$, the following inequality holds: $P\left(Y_w(a) > z\right) > P\left(Y_1(a) > z\right)$.*

Proof. We have $P\left(X_w > z\right) \leq P\left(Y_w(a) \geq X_w > z\right) + P\left(X_w > z, X_w > Y_w(a)\right) \leq P\left(Y_w(a) > z\right) + P\left(X_w > Y_w(a)\right) \leq P\left(Y_w(a) > z\right) + P\left(X_i > a \text{ for at least one } i \in \{1, 2, \dots, n\}\right) \leq$

$P(Y_w(a) > z) + \sum_{i=1}^n P(X_i > a) = P(Y_w(a) > z) + nP(X_1 > a)$. From Proposition 1 it follows that $P(X_w > z) > P(w^{(1)}X_1 + w^{(2)}X_2 > z) = P(X_1 > z) + G(w, z) = P(X_1 > a) + P(Y_1(a) > z) + G(w, z)$. The above relations imply that the following inequalities hold:

$$P(Y_w(a) > z) - P(Y_1(a) > z) > G(w, z) - (n-1)P(X_1 > a). \quad (4)$$

Since, under the assumptions of the theorem, $E|X_1|^r < \infty$, by Chebyshev's inequality we get $P(X_1 > a) = \frac{1}{2}P(|X_1| > a) \leq \frac{E|X_1|^r}{2a^r}$. Consequently, under the conditions of the theorem, the right-hand side of (4) is positive. ■

Remark 2 *From Proposition 1 it follows that value at risk is not subadditive and, thus, its coherency in the sense of Artzner et al. (1999) (see also the papers in Szegö, 2004) is always violated in the class of extremely heavy-tailed risks with infinite first moments. Theorem 1 implies that VaR may also be non-coherent in the world of risks with bounded distributional support.*

Remark 3 *We note that in the case of a portfolio with equal weights $\tilde{w}_n = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$, one has $(w^{(1)}, w^{(2)}) = (\frac{1}{2}, \frac{1}{2})$ and, thus, (3) becomes $G(\tilde{w}_n, z) = H(z) = P(\frac{X_1+X_2}{2} > z) - P(X_1 > z)$. This means that the length of the distributional support 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 $\tilde{G}(\epsilon, z) = P((1-\epsilon)X_1 + \epsilon X_2 > z) < G(w, z)$. 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 less diverse or more nearly equal components than)*

the vector $(1 - \epsilon, \epsilon, 0, \dots, 0)$.

Remark 4 From the proof of Theorem 1 it follows that, in the case of portfolios with equal weights $\tilde{w}_n = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$, $n > 2$, 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 $F_n(z) = P\left(\frac{\sum_{i=1}^n X_i}{n} > z\right) - P(X_1 > z)$. 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 Theorem 1 shows that, for a specific loss probability $q \in (0, 1/2)$, there exists a sufficiently large a such that the value at risk $VaR_q[Y_w(a)]$ of the return $Y_w(a)$ at level q is greater than the value at risk $VaR_q[Y_1(a)]$ of the return $Y_1(a)$ at the same level: $VaR_q[Y_w(a)] > VaR_q[Y_1(a)]$. The last inequality between the returns $Y_w(a)$ and $Y_1(a)$ holds for the particular fixed loss probability q . In the comparisons of the values at risks $VaR_q[Y_w(a)]$ 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 . This is a crucial difference compared with Proposition 1 and Remark 1, where the inequalities hold for all $z > 0$ and all $q \in (0, 1/2)$.

3.2 Non-diversification and 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 .¹² 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 (1987), we define a simple mean preserving spread (MPS) of a c.d.f., F , with corresponding p.d.f., f , by adding to f a function $\varphi(x)$ satisfying $\varphi(x) = \alpha, c < x < c + t; \varphi(x) = -\alpha, c' < x < c' + t; \varphi(x) = -\beta, d < x < d + t; \varphi(x) = \beta, d' < x < d' + t; \text{ and } \varphi(x) = 0$ otherwise, 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 $d(F_X, F_Y) = \inf\{\epsilon : F_X(x - \epsilon) - \epsilon \leq F_Y(x) \leq F_Y(x + \epsilon) + \epsilon \text{ for all } x\}$. This makes it a complete metric space with the topology of weak convergence.¹³ 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 X and Y are symmetric r.v.'s and $Y \in M_{F_X}$ then Y can be obtained from X by a finite sequence of pairs of simple MPS's with symmetric φ 's.

Rothschild and Stiglitz (1970) introduce four equivalent definitions of risk dominance, one of

¹²As we relate to the expected utility framework, we use the convention that negative values of X and Y are losses in this section.

¹³See Lukacs (1975).

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 $EX = EY$,¹⁴ 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, $u : [-a, a] \rightarrow \mathbf{R}$: $Eu(X) \geq Eu(Y)$.
- B4: $Y \stackrel{d}{=} X + U$, where U is a r.v. on $[-a, a]$ such that $E(U|X) = 0$.

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 $X_i \sim \mathcal{CS}(r)$, $i = 1, \dots, n$, 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 $[-a, a] \subseteq \mathbf{R}$.¹⁵ 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, B3 and B4).*

Proof. See Ibragimov and Walden (2006). ■

¹⁴Equivalently, $\int_{-a}^a x dF_X(x) = \int_{-a}^a x dF_Y(x)$.

¹⁵The value of a can be infinity so that $[-a, a]$ can be the whole real line: $[-a, a] = \mathbf{R}$.

Consequently, 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.

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, 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, B3 and B4.

Below, for a r.v. W , we denote by $\sigma(W)$ the σ -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'': (Weak integral condition)¹⁶ : For all $\epsilon > 0$, there exists $a > 0$ and a c.d.f. $\tilde{F}^\epsilon : [-a, a] \rightarrow$

¹⁶Clearly, (5) implies that $|\xi(x)| \leq 1$ (a.s.) and $\int_{-a}^t \xi(x)dx \leq \epsilon$ 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,

[0, 1], such that

1. For all $t \in (-a, a)$: $\int_{-a}^t F_X^a(x)dx \leq \int_{-a}^t \tilde{F}^\epsilon(x)dx$.
2. $F_Y^a = \tilde{F}^\epsilon + \xi + s$, where ξ is an antisymmetric¹⁷ function satisfying

$$F_Y(x - \epsilon) - F_Y(x) \leq \xi(x) \leq F_Y(x + \epsilon) - F_Y(x), \quad \text{for all } x \quad (5)$$

and s is an antisymmetric function with $|s(x)| \leq \epsilon$ for (almost) all x .

3. When $\epsilon \rightarrow 0$, $a \rightarrow \infty$.
- B3': There is an a_0 such that for all $a > a_0$ for all concave u : $Eu(X^a) \geq Eu(Y^a)$ (Expected utility condition).
 - B4': There is an a_0 such that for all $a > a_0$, $Y^a \stackrel{d}{=} X^a + Z^a$, where Z^a is a $\sigma(X^a)$ -measurable r.v. such that $E(Z^a|X^a) = 0$ (a.s.) (Fair game condition).
 - B4'': $Y \stackrel{d}{=} \text{sign}(Y)[|X| + Z] \stackrel{d}{=} X + \text{sign}(X)Z$, where $Z \geq 0$ (a.s.) (Conditional absolute symmetry condition).¹⁸

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. See Ibragimov and Walden (2006). ■

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 dominates diversified portfolios in the

then one can choose $\xi = 0$. This is the case as $|F_Y(x + \nu) - F_Y(x)| \leq C|\nu|$ for all $|\nu|$ and therefore the condition will be satisfied with $\xi = 0$ and $|s(x)| \leq \tilde{\epsilon} = (C + 1)\epsilon$.

¹⁷That is, $f(-x) = -f(x)$ for all x .

¹⁸For condition B4'', we restrict our attention to absolutely continuous distributions. However, complete analogues of the results below hold as well in the discrete case.

sense of B1', B2', B2'', B3', B4' and B4''. Therefore, the results in Ibragimov (2004a, b, 2005) on diversification always being nonpreferable are also 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.¹⁹ For the first approach, the law of large numbers implies that, for all $\epsilon, \epsilon_1 > 0$, $P\left(\left|\bar{Z}_n - \mu\right| > \epsilon\right) = P\left(\left|\frac{\sum_{i=1}^n Z_i}{n} - \mu\right| > \epsilon\right) < \epsilon_1$ if $n > N(\epsilon_1) > 0$ and the risks Z_1, Z_2, \dots are i.i.d. r.v.'s with $EZ_1 = \mu$, and $\bar{Z}_n = n^{-1} \sum_{i=1}^n Z_i$. 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 in Theorem 1.

The second motivation for diversification is based on expected utility. Samuelson (1967) showed

¹⁹We view Markowitz' (1952) mean-variance approach as a special case of the latter.

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 \bar{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 \bar{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 6, 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: $VaR_q(\bar{Y}_n(a)) > VaR_q(Y_1(a))$? 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 $\tilde{q} \gg q$, $VaR_{\tilde{q}}(\bar{Y}_n(a)) < VaR_{\tilde{q}}(Y_1(a))$. This is contrary to the case when $a = \infty$ in which no such reversal takes place. 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, 1997). 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 use the following notation. An increasing, strictly concave function: $u : \mathbf{R} \rightarrow \mathbf{R}$ will be called admissible. For any $t > 0$, a continuous function $v : \mathbf{R} \rightarrow \mathbf{R}$ 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 $u(-t) - \lim_{s \rightarrow \infty} v(-s) \leq 1/t$. 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. 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 $Y_i(a)$, $i = 1, \dots, n$ and the diversified portfolio with equal weights $\tilde{w}_n = (\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ and the return $\bar{Y}_n(a) = \frac{1}{n} \sum_{i=1}^n Y_i(a)$.

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 $Ev(Y_1(a)) > Ev(\bar{Y}_n(a))$.*

Proof. See Ibragimov and Walden (2006). ■

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 in the domain of 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. Consider i.i.d. symmetric stable risks $X_1, X_2, \dots, X_n \sim S_\alpha(\sigma, 0, 0)$ with $\alpha \in (r, 1)$ and $\sigma > 0$.²⁰ We first study the case with two risks, $n = 2$. From Zolotarev (1986, the argument for Property 2.5 on p. 63), we have $E|X|^r \leq 2\sigma^r \Gamma\left(1 - \frac{r}{\alpha}\right) \Gamma(r) \sin\left(\frac{\pi}{2}r\right)$, where $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$ is the Gamma function. In addition, using the asymptotic expansions for stable cdf's given by Theorem 2.4.2 in Zolotarev (1986), we get $G(w, z) > \frac{1}{\alpha\pi} \Gamma(\alpha + 1) \sin\left(\frac{\pi\alpha}{2}\right) \frac{\sigma^\alpha}{z^\alpha} \left((w^{(1)})^\alpha + (w^{(2)})^\alpha - 1 \right) - \frac{1}{\alpha\pi} \frac{\Gamma(2\alpha+1)}{4} \sin(\pi\alpha) \frac{\sigma^{2\alpha} \left[(w^{(1)})^\alpha + (w^{(2)})^\alpha \right]^2}{z^{2\alpha}}$. Using the above inequalities, we obtain that Theorem 1 holds with the following easy to compute estimate for the length of the distribution support:

$$\tilde{a} = \frac{z^{\alpha/r} (\alpha\pi)^{1/r} \sigma^{(r-\alpha)/r} \left(\Gamma\left(1 - \frac{r}{\alpha}\right) \Gamma(r) \sin\left(\frac{\pi}{2}r\right) \right)^{1/r} (n-1)^{1/r}}{\left[\Gamma(\alpha + 1) \sin\left(\frac{\pi\alpha}{2}\right) \left((w^{(1)})^\alpha + (w^{(2)})^\alpha - 1 \right) - \frac{\Gamma(2\alpha+1)}{4} \sin(\pi\alpha) \frac{\sigma^\alpha \left((w^{(1)})^\alpha + (w^{(2)})^\alpha \right)^2}{z^\alpha} \right]^{1/r}}. \quad (6)$$

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 2, 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. 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.

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 value at

²⁰Generalizations of the analysis in the case of skewed stable risks $X_i \sim S_\alpha(\sigma, \beta, 0)$ can be obtained in a similar way and are presented in Ibragimov and Walden (2006).

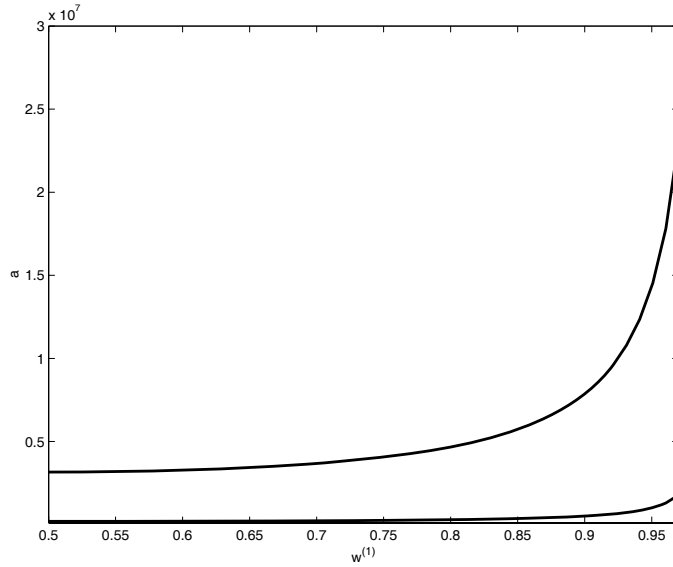


Figure 2: Relationship between distributional support, \tilde{a} , and level of diversification, $w^{(1)}$, for VaR centiles $q = 0.01$ (below) and $q = 0.001$ (above). Parameters: $n = 2$, $\sigma = 1$, $\beta = 0$.

risk, $a(n, z)$ where value at risk is $z = VaR_q(X_1)$, such that Theorem 1 is satisfied for portfolios with equal weights. 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 fraction 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 .²¹ We use the exact formula in (6). For $\alpha = 0.85$, the results are shown in Table 1. A general conclusion

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	9,547	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 . $\alpha = 0.85$, $\beta = 0$.

is that the worst case scenario must be a lot worse than the VaR level chosen, for diversification to be inferior. For example, with a value at risk corresponding to $q = 1\%$, the worst case scenario

²¹Therefore, it also holds for arbitrary time scales, T , according to our previous discussion.

must be almost 2,700 times VaR_q for diversification into two 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, Theorem 1 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 value of 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. As shown in Section 4, diversification also breaks for a wide class of bounded risks that exhibit dependence. The length of distributional support required for diversification failure may therefore be considerably smaller.

4 Generalizations to dependence and non-identical distributions

Our results 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 α -symmetric and spherical distributions and their non-identically distributed versions as well as for a wide class of models with common shocks.

Following Cambanis, Keener and Simons (1983), an n -dimensional distribution is called α -symmetric if its characteristic function (c.f.) can be written as $\phi((\sum_{i=1}^n |t_i|^\alpha)^{1/\alpha})$, where $\phi : \mathbf{R}_+ \rightarrow \mathbf{R}$ is a continuous function (with $\phi(0) = 1$) and $\alpha > 0$. The number α is called the index and the function ϕ is called the c.f. generator of the α -symmetric distribution. The class of α -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 α -stable distributions

(Fang, Kotz and Ng, 1990, Ch. 3). Spherically symmetric stable distributions have characteristic functions $\exp[-\lambda(\sum_{i=1}^n t_i^2)^{\gamma/2}]$, $0 < \gamma \leq 2$, and are, thus, examples of α -symmetric distributions with $\alpha = 2$ and the c.f. generator $\phi(x) = \exp(-x^\gamma)$.

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

$$(Q_1, \dots, Q_n) = (ZY_1, \dots, ZY_n), \quad (7)$$

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 $Z \geq 0$ is a nonnegative r.v. independent of Y_i 's (see Bretnagolle, Dacuhna-Castelle and Krivine, 1966, and Fang, Kotz and Ng, 1990, p. 197). In the case $Z = 1$ (a.s.), model (7) represents vectors with i.i.d. symmetric stable components that have c.f.'s $\exp[-\lambda \sum_{i=1}^n |t_i|^\alpha]$ which are particular cases of c.f.'s of α -symmetric distributions with the generator $\phi(x) = \exp(-\lambda x^\alpha)$.

The dependence structures considered in this section include, among others, convolutions of models (7). That is, the dependence structures cover vectors (X_1, \dots, X_n) which are sums of i.i.d. random vectors $(Z_j V_{1j}, \dots, Z_j V_{nj})$, $j = 1, \dots, k$, where $V_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $i = 1, \dots, n$, $j = 1, \dots, k$, and Z_j are positive absolutely continuous r.v.'s independent of V_{ij} :

$$(X_1, \dots, X_n) = \sum_{j=1}^k (Z_j V_{1j}, \dots, Z_j V_{nj}). \quad (8)$$

Although the dependence structure in model (7) alone is restrictive, convolutions (8) of such vectors provide a natural framework for modeling of random environments with different multiple common shocks Z_j , such as macroeconomic or political ones, that affect all risks X_i (see Andrews, 2005). Moreover, convolutions (8) of common factor models (7) can be viewed as generalized multi-factor models, with factors Z_j that have random factor sensitivities, V_{kj} . The results in this section thus have direct implications for multi-factor modeling of financial or insurance variables, including the

returns on stocks and bonds, and general credit risk.

Convolutions of α -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 α -symmetric distributions with $\alpha < 1$ have extremely heavy-tailed marginals with infinite means.²² On the other hand, convolutions of α -symmetric distributions with $1 < \alpha \leq 2$, and, in particular, convolutions of models (7) with $1 < \alpha \leq 2$, 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 (7) with $1 < \alpha < 2$ and $E|Z| < \infty$ have finite means but infinite variances, however, marginals of such convolutions have infinite means if the r.v.'s Z satisfy $E|Z| = \infty$. Moments $E|ZY_i|^p$, $p > 0$, of marginals in models (7) 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 (7) 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 light-tailed ones. For example, marginal moments of spherically symmetric α -stable distributions with c.f.'s $\exp[-\lambda(\sum_{i=1}^n t_i^2)^{\gamma/2}]$, $0 < \gamma < 2$, 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 an example of a spherical distribution are finite if and only if 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, Heidelberger and Shahabuddin, 2002).

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

²²This is true because if one assumes that r.v.'s X_1, \dots, X_n , $n \geq 2$, have an α -symmetric distribution with $\alpha < 1$ and that $E|X_i| < \infty$, $i = 1, \dots, n$, then, by the triangle inequality, $E|X_1 + \dots + X_n| \leq E|X_1| + \dots + E|X_n| = nE|X_1|$. The latter, however, cannot hold since, according to (2), $(X_1 + \dots + X_n) \sim n^{1/\alpha}X_1$ and, thus, under the above assumptions, $E|X_1 + \dots + X_n| > nE|X_1|$. Similarly, one can show that α -symmetric distributions with $\alpha < r$ have infinite marginal moments of order r .

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

- *The vector of r.v.'s (X_1, \dots, X_n) entering its assumptions is a sum of i.i.d. random vectors (V_{1j}, \dots, V_{nj}) , $j = 1, \dots, k$, where (V_{1j}, \dots, V_{nj}) has an absolutely continuous α -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 $(Z_j V_{1j}, \dots, Z_j V_{nj})$, $j = 1, \dots, k$, where $V_{ij} \sim S_{\alpha_j}(\sigma_j, 0, 0)$, $i = 1, \dots, n$, $j = 1, \dots, k$, 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. See Ibragimov and Walden (2006). ■

Theorem 4 provides a precise formulation of the extensions of the results in Subsection 3.1 to the dependent case. In particular, Theorem 1 continues to hold for convolutions of truncated analogues of models (7) with common shocks affecting all thick-tailed risks Y_i with tail indices $\alpha < 1$.

The generalization to non-identical distributions, is straightforward. Let $\sigma_1, \dots, \sigma_n \geq 0$ be 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, \dots, X_n if $\sigma_n \geq \dots \geq \sigma_1 \geq 0$, we obtain that Theorem 1 also holds under these assumptions.

5 Concluding remarks

We have analyzed the limits of diversification for bounded risks with heavy tails in their support. 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. 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.

Our analysis may have implications for several under-diversification puzzles in financial markets and insurance markets (see e.g., Blume and Friend, 1975, Barber and Odean, 2000, and Froot, 2001, for examples of such puzzles). For example, our analysis can be used to explain low levels of reinsurance among insurance providers in markets for catastrophe reinsurance, as shown in Ibragimov, Jaffee and Walden (2006).

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.

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

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

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.

- 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.
- Bretagnolle, J., Dacunha-Castelle, D. and Krivine, J. L. (1966). Lois stables et espaces L^p . *Annales de l'Institut H. Poincaré. Section B. Calcul des Probabilités et Statistique*. **2**, 231-259.
- Cambanis, S., Keener, R. and Simons, G. (1983). On α -symmetric distributions. *Journal of Multivariate Analysis* **13**, 213-233.
- Cotter, D. and Dowd, K. (2006). Extreme spectral risk measures: an application to futures clearinghouse margin requirements. Working paper, University College Dublin and Nottingham University. Forthcoming in *Journal of Banking and Finance*.
- 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., 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.
- 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.
- Guillaume, D, Dacorogna M., Davé, R., Müller, U., Olsen. R. and Pictet, O. (1997). From the bird's eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics* **1**, 95-129.
- 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, R. (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. (2007). Efficiency of linear estimators under heavy-tailedness: Convolutions of α -symmetric distributions. *Econometric Theory*. **23**(3).
- Ibragimov, R. and Walden, J. (2006). The limits of diversification when losses may be large. Harvard Institute of Economic Research Discussion Paper No. 2104. Available at <http://www.economics.harvard.edu/hier/2006papers/HIER2104.pdf>
- Ibragimov, R., Jaffee, D. and Walden, J. (2006). Non-diversification traps in markets for catastrophic risk. Working paper, Harvard University and University of California at Berkeley.
- 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.
- 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.
- Lux, T. (1996). The stable Paretian hypothesis and the frequency of large returns: an examination of major German stocks. *Applied Financial Economics* **6**, 463-475.
- 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.
- Nešlehova, J., Embrechts, P. and Chavez-Demoulin, V. (2006). Infinite mean models and the LDA for operational risk. *Journal of Operational Risk* **1**, 3-25.
- Praetz, P. (1972). The distribution of share price changes. *Journal of Business* **45**, 49-55.
- 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.
- Ross, S. A. (1976). Arbitrage theory of capital asset pricing. *Journal of Economic Theory* **13**, 341-360.
- 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**, 431-449.
- Samuelson, P. A. (1967). General proof that diversification pays. *The Journal of Financial and Quantitative Analysis* **2**, 1-13.
- 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. Forthcoming in *Journal of Econometrics*.
- 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* **71**, 393-410.
- Szegö, G., Ed. (2004). *Risk measures for the 21st century*. Wiley, Chichester.
- Tasche, D. (2002). Expected shortfall and beyond. *Journal of Banking and Finance* **26**, 1519-1533.
- Zolotarev, V. M. (1986). *One-dimensional stable distributions*. Providence: American Mathematical Society.