Available online at www.sciencedirect.com
ScienceDirect
Journal of Empirical Finance 14 (2007) 383–400
Journal of EMPIRICAL FINANCE
www.elsevier.com/locate/jempfin

Portfolio selection with heavy tails

Namwon Hyung [a,b,*], Casper G. de Vries [b,c,d]

[a] *Faculty of Economics, The University of Seoul, Seoul, 130-743, South Korea*
[b] *Tinbergen Institute, The Netherlands*
[c] *Department of Accounting and Finance H8-33 Erasmus University Rotterdam, PO Box 1738, 3000 DR, Rotterdam, The Netherlands*
[d] *EURANDOM, The Netherlands*

Accepted 29 June 2006
Available online 31 January 2007

Abstract

Consider the portfolio problem of choosing the mix between stocks and bonds under a downside risk constraint. Typically stock returns exhibit fatter tails than bonds corresponding to their greater downside risk. Downside risk criteria like the safety first criterion therefore often select corner solutions in the sense of a bonds only portfolio. This is due to a focus on the asymptotically dominating first order Pareto term of the portfolio return distribution. We show that if second order terms are taken into account, a balanced solution emerges. The theory is applied to empirical examples from the literature.

JEL classification: G11
Keywords: Safety first; Heavy tails; Portfolio diversification

1. Introduction

Consider the portfolio problem of choosing the mix between a stock index and a government bond index. The mean variance criterion selects non-zero proportions of each as long as stocks have higher expected returns and higher variance. Investors nevertheless in addition often worry about the downside risk features of their portfolio, witness the popularity of policies with put protection that lock in gains, portfolio insurance, capital buffers at pension funds, Value at Risk (VaR) exercises at banks, etc. It is a fact that asset return distributions exhibit fat tails, i.e. are asymptotic to a Pareto distribution. Typically stocks exhibit fatter tails than bonds, i.e. have

* Corresponding author. Faculty of Economics, The University of Seoul, Seoul, 130-743, South Korea.
E-mail addresses: nhyung@uos.ac.kr (N. Hyung), cdevries@few.eur.nl (C.G. de Vries).

0927-5398/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.jempfin.2006.06.004

smaller hyperbolic Pareto coefficient, corresponding to the greater downside risk of stocks. Downside risk criteria like the safety first criterion therefore often select corner solutions in the sense of a bonds only portfolio. This is due to a focus on the tail of the asset return distributions whereby only the asymptotically dominating first order Pareto term is taken into account. In this note we show that if the second order terms are considered as well, a more balanced solution emerges. The theory is applied to examples from the literature.

Portfolio risk and its upside potential are in an important way driven by the 'abnormal' returns emanating from heavy-tailed distributed asset returns. Therefore the financial industry often employs so called downside risk measures to characterize the asset and portfolio risk, since it is widely recognized that large losses are more frequent than a normal distribution based statistic like the standard deviation suggests. A formal portfolio selection criterion which incorporates the concern for downside risk is the safety first criterion, see Roy (1952) and Arzac and Bawa (1977). The paper by Gourieroux, Laurent and Scaillet (2000) analyzes the sensitivity of VaR with respect to portfolio allocation, which is essentially the same problem as portfolio selection with the safety first criterion. Gourieroux et al. (2000) show how to check for the convexity of the estimated VaR efficient portfolio set. Jansen, Koedijk and de Vries (2000) and Jansen (2000) apply the safety first criterion and exploit the fact that returns are fat-tailed. They propose a semi-parametric method for modeling tail events and use extreme value theory to measure the downside risk. This method was subsequently used by Susmel (2001) in an application involving Latin American stock markets.

If one selects assets on the basis of the tail properties of the return distribution, there is a tendency to end up with a corner solution whereby the asset with the highest tail coefficient (thinnest tail) is selected, see e.g. Straetmans (1998, ch.5), Jansen et al. (2000), Hartmann, Straetmans and de Vries (2004) and Poon, Rockinger and Tawn (2003). This follows from Geluk and de Haan (1987), who show that a convolution of two regularly varying variables produces a random variable which has the same tail properties as the fattest tail of the two convoluting variables, i.e. the fattest tail (lowest tail coefficient) dominates. In case the tails are equally fat, the scales of the two random variables have to be added. In this paper we show how to extend the first order convolution result to a second order asymptotic expansion. Whereas in the first order convolution result only the fattest of the two tails plays a role, in the second order expansion often both tails play a role. We show that with a second order expansion of the downside risk, the portfolio solution yields a balanced solution, i.e. both assets are held in non-zero proportion, whereas the first order expansion selects the corner solution. In an extension we also consider the case of dependent returns in multi asset portfolios. In the empirical application, we follow up on Jansen et al. (2000) and Susmel (2001), who apply the safety first criterion to a number of portfolio problems. In several cases Jansen et al. (2000) end up with a corner solution. We calculate the downside risk using the second order expansion and show how this implies a move towards the interior.

2. Extreme value theory

The fat tail property is one of the salient features of asset returns. This can be modeled by letting the tail of the distribution be governed by a power law, instead of an exponential rate. Technically speaking, suppose that the returns are i.i.d. and have tails which vary regularly at infinity. This entails that to a first order

$$P\{X>s\} = As^{-\alpha} + o(s^{-\alpha})$$

as $s \to \infty$, where $\alpha > 0$, $A > 0$. A more detailed parametric form for the tail probability can be obtained by taking a second order expansion at infinity. There are only two non-trivial expansions (de Haan and Stadtmüller, 1996). The first expansion has a second order term which also declines hyperbolically

$$P\{X > s\} = As^{-\alpha}[1 + Bs^{-\beta} + o(s^{-\beta})]$$

as $s \to \infty$, where $\alpha > 0$, $A > 0$, $\beta > 0$ and B is a real number. This expansion applies to the non-normal sum-stable, Student-t, Fréchet, and other fat tailed distributions. The other non-trivial expansion is

$$P\{X > s\} = As^{-\alpha}[1 + B \log s + o(\log s)]$$

which is not considered in this paper.[1]

We assume that the tails of two assets are different but symmetric, and vary regularly at infinity. Consider the following second order expansions,

$$P\{X_1 > s\} = P\{X_1 < -s\} = A_1 s^{-\alpha_1}[1 + B_1 s^{-\beta_1} + o(s^{-\beta_1})] \tag{1}$$

$$P\{X_2 > s\} = P\{X_2 < -s\} = A_2 s^{-\alpha_2}[1 + B_2 s^{-\beta_2} + o(s^{-\beta_2})] \tag{2}$$

as $s \to \infty$. We assume $2 < \alpha_1 \leq \alpha_2$. The assumption of $2 < \alpha_1$ implies that at least the mean and variance exist, which seems to be the relevant case for financial data. Portfolios are essentially (weighted) sums of different random variables. We therefore investigate the tail probability of the convolution $X_1 + X_2$. The case of equal tail indices $\alpha_1 = \alpha_2$ is known from Feller (1971, ch. VIII). In this case $P\{X_1 + X_2 > s\} = (A_1 + A_2)s^{\alpha_1} + o(s^{-\alpha_1})$ as $s \to \infty$. When the tail indices are unequal we have the following results.

Theorem 1. *Suppose that the tails of the distributions of X_1 and X_2 satisfy* Eqs. (1) and (2). *Moreover, assume $2 < \alpha_1 < \alpha_2$ so that $E[X]$ and $E[X^2]$ are bounded. When X_1 and X_2 are independent, the asymptotic 2-convolution up to the second order terms is*

(I) if $\alpha_2 - \alpha_1 < \min(\beta_1, 1)$, then $P\{X_1 + X_2 > s\} = A_1 s^{-\alpha_1} + A_2 s^{-\alpha_2} + o(s^{-\alpha_2})$
(II) if $1 < \alpha_2 - \alpha_1$ and $1 < \beta_1$, then $P\{X_1 + X_2 > s\} = A_1 s^{-\alpha_1} + A_1\alpha_1 E[X_2]\, s^{-\alpha_1 - 1} + o(s^{-\alpha_2})$
(III) if $\beta_1 < \alpha_2 - \alpha_1$ and $\beta_1 < 1$, then $P\{X_1 + X_2 > s\} = A_1 s^{-\alpha_1} + A_1 B_1 s^{-\alpha_1 - \beta_1} + o(s^{-\alpha_1 - \beta_1})$
(IV) if $\alpha_2 - \alpha_1 = 1 < \beta_1$, then $P\{X_1 + X_2 > s\} = A_1 s^{-\alpha_1} + \{A_2 + A_1\alpha_1 E[X_2]\} s^{-\alpha_2} + o(s^{-\alpha_2})$
(V) if $\alpha_2 - \alpha_1 = \beta_1 < 1$, then $P\{X_1 + X_2 > s\} = A_1 s^{-\alpha_1} + \{A_2 + A_1 B_1\} s^{-\alpha_1} + o(s^{-\alpha_1})$
(VI) if $\alpha_2 - \alpha_1 = \beta_1 = 1$, then $P\{X_1 + X_2 > s\} = A_1 s^{-\alpha_1} + \{A_2 + A_1\alpha_1 E[X_2] + A_1 B_1\} s^{-\alpha_2} + o(s^{-\alpha_2})$.

Proof. We only provide the proof of the upper tail case. The proof for the lower tail case only requires a small modification of this proof. Parts of the proof are similar in spirit as the proof in Dacarogna, Müller, Pictet and de Vries (2001, Lemma 4). It is an extension of Feller's original convolution result for regularly varying distributions. We divide the area over which we have to integrate into five parts A, B, C, D and E; where $P\{A\} = P\{X_1 + X_2 \leq s, X_1 > -\frac{s}{2}, X_2 > \frac{s}{2}\}$, $P\{C\} = P\{X_1 \leq \frac{s}{2}, X_2 \leq \frac{s}{2}\}$, $P\{D\} = P\{X_1 + X_2 \leq s, X_1 \leq -\frac{s}{2}, X_2 > \frac{s}{2}\}$, and where $P\{B\}$ and $P\{E\}$ are the counterparts of $P\{A\}$ and $P\{D\}$ respectively. By integration we find $P\{A\}$, $P\{D\}$, and

[1] The slow decay of the second order term makes this class sufficiently different from the other class. The inclusion of this class would make our paper overly long.

$P\{C\}$. The integrals are provided in Appendix A. Adding up and ignoring the terms which are of smaller order than $s^{-\gamma}$, $\gamma=\alpha_1+\beta_1$, α_1+1, $\alpha_2+\beta_2$, or α_2+1, we find that

$$\begin{aligned} P\{X_1+X_2>s\} &= 1-[P\{C\}+P\{A\}+P\{B\}] \\ &= A_1 s^{-\alpha_1}+A_1B_1 s^{-\alpha_1-\beta_1}+A_2 s^{-\alpha_2}+A_2B_2 s^{-\alpha_2-\beta_2}+A_1\alpha_1 E[X_2]s^{-\alpha_1-1} \\ &\quad +A_2\alpha_2 E[X_1]s^{-\alpha_2-1}+o(s^{-\gamma}) \end{aligned}$$

By considering the different parameter configurations (I)–(VI), we obtain the results of Theorem 1. □

What is the relevance of this theorem for portfolio selection? Suppose that portfolio selection is done on the basis of the concern for the downside risk, safety-first criterion using this convolution result. By mapping negative returns into the positive quadrant, this theorem applies to the left tail with a little modification. Let X_i denote the loss returns on two independent project. Under this criterion the problem is to minimize $P\{\omega X_1+(1-\omega)X_2>s\}$ at some large loss levels s by choosing the asset mix ω. Suppose only the first order terms of tail probability $P\{X_i>s\}=A_i s^{-\alpha_i}$ are taken into account. Then for large loss levels s one choose $\omega=0$, if $\alpha_1<\alpha_2$. This corner solution is driven by evaluation of the safety first criterion in the limit (where only the first order term is relevant). In practice what counts are very high, but finite loss levels. Thus a second order expansion in which the second order term still plays a role has practical relevance. To this end we can use the Theorem 1.

Consider first the case III above. Since asset 1 dominates the first two terms in the loss probability, one is still better of by putting all eggs in one basket. Turn to case I. If one would focus on the first term only, i.e. only taking the limit as $s\to\infty$ into consideration, then again only asset two is selected. At any finite loss level s, this solution is, however, suboptimal. Given that $P\{X_1+X_2>s\}\approx A_1 s^{-\alpha_1}+A_2 s^{-\alpha_1}$ in case I, one should take both assets into account and diversify away from the corner solution. This lowers the loss probability $P\{X_1+X_2>s\}$ at any finite loss level s. This idea is put on a firm footing in the next section by investigating the convexity properties of the solutions.

3. The sensitivity and convexity of VaR

The aim of this section is to analyze the sensitivity of VaR with respect to portfolio allocation. Gourieroux et al. (2000) derive analytical expression for the first and second derivatives of the VaR in a general framework, and state sufficient conditions for the VaR efficient portfolio set to be convex. Gourieroux et al. (2000) also provide explicit expression for the first and second derivatives in case of the normal distribution. Here we provide explicit expressions for the class of fat tailed distributions. Moreover, we show how to ensure an interior solution under which the VaR is convex with respect to the portfolio weight. If a risk measure is a convex function of the portfolio allocation, it induces portfolio diversification. From this we can ensure that an interior solution to the safety first problem exists. While Gourieroux et al. (2000) show the convexity of the VaR-efficient portfolio set in general, they do not give conditions to ensure an interior solution for the optimal allocation.

First, we derive analytical expression of derivatives of the tail probability at a given quantile in the heavy tail context. This allows us to discuss the convexity properties of VaR. We consider two financial assets whose returns at time t are denoted by X_i, $i=1, 2$. We suppress time indices whenever this is not confusing. The return at t of a portfolio with allocation ω then is $\omega X_1+(1-\omega)X_2$. For a loss probability level p the Value at Risk, VaR(ω, p) is defined by:

$$P\{\omega X_1+(1-\omega)X_2>\mathrm{VaR}(\omega,p)\}=p.$$

In practice, VaR is often computed under the normality assumption for returns. Recently, semi-parametric approaches have been developed, which are based on the extreme value approximation to the tail probability like in the previous section. We derive the first and second derivatives of the probability with respect to portfolio allocation under this approximation. Under the safety first rule an investor specifies a low threshold return s and selects the portfolio of assets which minimizes the probability of a return below this threshold.

3.1. Convexity of the tail probability

Suppose the tails of the distributions of X_1 and X_2 satisfy Eqs. (1) and (2). We obtain the first and second derivatives in the Proof to Proposition 2. We first investigate the case I from the convolution Theorem 1.

Proposition 2. *Under assumptions of* Theorem 1 *and if* $\alpha_2-\alpha_1<\min(\beta_1, 1)$, *there exists a* $\omega^*\in(0, 1)$ *for given large* $s>0$ *such that*

$$P\{\omega^* X_1 + (1-\omega^*)X_2>s\}\leq P\{\omega X_1 + (1-\omega)X_2>s\}$$

for any $0\leq\omega\leq 1$. *The equality holds only when* $\omega=\omega^*$.

Proof. From Theorem 1, the asymptotic 2-convolution up to the second order terms is

$$P\{\omega X_1 + (1-\omega)X_2>s\}\approx\omega^{\alpha_1}A_1 s^{-\alpha_1} + (1-\omega)^{\alpha_2}A_2 s^{-\alpha_2}\equiv p(\omega, s),$$

for given large $s>0$. Note that $p(\omega, s)$ is an approximate asymptotic expansion. We show the function of $p(\omega, s)$ has a minimum for some $\omega\in(0, 1)$. The slope of this function with respect to ω is

$$\frac{\partial p(\omega, s)}{\partial\omega} = \alpha_1\omega^{\alpha_1-1}A_1 s^{-\alpha_1}-\alpha_2(1-\omega)^{\alpha_2-1}A_2 s^{-\alpha_2}$$

for large $s>0$. Thus slopes at the endpoints are

$$\left.\frac{\partial p(\omega, s)}{\partial\omega}\right|_{\omega=0} = -\alpha_2 A_2 s^{-\alpha_2}<0$$

and

$$\left.\frac{\partial p(\omega, s)}{\partial\omega}\right|_{\omega=1} = \alpha_1 A_1 s^{-\alpha_1}>0.$$

for large $s>0$. The slope of this function increases monotonically since the second order derivative of this function is

$$\frac{\partial^2 p(\omega, s)}{\partial\omega^2} = (\alpha_1-1)\alpha_1\omega^{\alpha_1-2}A_1 s^{-\alpha_1} + (\alpha_2-1)\alpha_2(1-\omega)^{\alpha_2-2}A_2 s^{-\alpha_2}$$

which is positive for all $0\leq\omega\leq 1$ provided $\alpha=\min\{\alpha_1, \alpha_2\}>1$. □

In the Proof of the Proposition 2 we show the convexity of $p(\omega, s)\equiv\omega^{\alpha_1} A_1 s^{-\alpha_1}+(1-\omega)^{\alpha_2}A_2 s^{-\alpha_2}$. Note that this expression is only asymptotic to $P\{\omega X_1+(1-\omega) X_2>s\}$ as $s\to\infty$. Therefore $\partial P\{\omega^* X_1+(1-\omega^*)X_2>s\}/\partial\omega$ will typically be close to zero but not be exactly equal to zero. The Proof of the Proposition 2 carries over for the exact expansions under the monotone density condition.

Remark 1. The Proposition 2 implies that if one constructs a portfolio which minimizes the probability of extreme negative returns, one has to assign some weight to the asset with the fatter tail.

Remark 2. Under conditions (II) and (III) from Theorem 1, Proposition 2 has trivial solutions such as $\omega^*=0$ or $\omega^*=1$ depending on the conditions of parameters.

Remark 3. With conditions (IV), (V) and (VI) from Theorem 1, Proposition 2 has non-trivial solution such that $\omega^* \in (0, 1)$ provided the parameters satisfy additional conditions. We illustrate the case of condition (IV) as an example. Under the condition (IV), $\alpha_2-\alpha_1=1<\beta_1$, then $P\{\omega X_1+(1-\omega)X_2>s\} \approx \omega^{\alpha_1}A_1 s^{-\alpha_1}+(1-\omega)^{\alpha_1}A_2 s^{-\alpha_2}+\omega^{\alpha_2}A_1\alpha_1 E[(1-\omega)X_2]\ s^{-\alpha_1} \equiv q(\omega)$. The slope of this function is

$$\frac{\partial q(\omega)}{\partial \omega} = \omega^{\alpha_1-1}\alpha_1 A_1 s^{-\alpha_1}-(1-\omega)^{\alpha_2-1}\alpha_2 A_2 s^{-\alpha_2} + (\omega^{\alpha_1-1}\alpha_1-\omega^{\alpha_1}(\alpha_1+1))A_1\alpha_1 E[X_2]s^{-\alpha_2}$$

For the corner solution excluding the asset 1 with the heaviest tail

$$\left.\frac{\partial q(\omega)}{\partial \omega}\right|_{\omega=0} = -\alpha_2 A_2 s^{-\alpha_2}<0$$

for large $s>0$. On the other hand, if the following condition is satisfied forlarge $s>0$,

$$\left.\frac{\partial q(\omega)}{\partial \omega}\right|_{\omega=1} = \alpha_1 A_1 s^{-\alpha_1}-\alpha_1 A_1 E[X_2]s^{-\alpha_2}>0$$

then there exists a non-trivial solution under the condition (IV), too. The last condition will be satisfied if $E[X_2]<s$. That is, $E[X_2]$ must not be too large for the given a finite loss level s. This holds certainly as long as the expected return is positive (since the $E[X_2]<0$, recall that a positive X_i reflects a loss).

3.2. Convexity of VaR

We now turn around the question from the previous section, and ask whether the VaR at a given probability level is convex. If the VaR criterion is used as the risk measure for judging the portfolio, and if we can show that the VaR is a convex function of the portfolio allocation, then there is an incentive for portfolio diversification under the VaR objective.

Proposition 3. *Under assumptions of* Theorem 1 *and if* $\alpha_2-\alpha_1<\min(\beta_1, 1)$, *consider the downside risk level*

$$P\{\omega X_1+(1-\omega)X_2>s\} = \omega^{\alpha_1}A_1 s^{-\alpha_1}\left[1+\frac{(1-\omega)^{\alpha_2}A_2}{\omega^{\alpha_1}A_1}s^{-\alpha_2+\alpha_1}+o(s^{-\alpha_2+\alpha_1})\right]$$

and define the VaR implicitly as follows $P\{\omega X_1+(1-\omega)X_2>\mathrm{VaR}(\omega, p)\}=p$. *By De Bruijn's theory on asymptotic inversion*

$$\mathrm{VaR}(\omega,p) = \omega A_1^{\frac{1}{\alpha_1}}p^{-\frac{1}{\alpha_1}}\left[1+\frac{(1-\omega)^{\alpha_2}}{\omega^{\alpha_2}}\frac{A_2}{\alpha_1 A_1^{\alpha_2/\alpha_1}}p^{\frac{\alpha_2-\alpha_1}{\alpha_1}}+o(1)\right]$$

for any $0<\omega<1$.

Proof. Directly follows from de Bruijn's inverse in Theorem 1.5.13 of Bingham, Goldie and Teugels (1987). □

For the given loss probability p, we can find an allocation which minimizes the VaR risk.

Proposition 4. *Under assumptions of* Theorem 1 *and if* $\alpha_2-\alpha_1<min(\beta_1, 1)$, *there exist* $\omega^*\in(0, 1)$ *for given probability level* $\overline{p}$ *such that*

$$VaR(\omega^*,\overline{p})\leq VaR(\omega,\overline{p})$$

for any $0<\omega<1$. *The equality holds only when* $\omega=\omega^*$.

Proof. For a given probability level $\overline{p}$, the first derivative of the VaR is

$$\frac{\partial \mathrm{VaR}(\omega,\overline{p})}{\partial \omega}=A_1^{\frac{1}{\alpha_1}}\overline{p}^{-\frac{1}{\alpha_1}}-\alpha_1^{-1}A_1^{\frac{1-\alpha_2}{\alpha_1}}A_2\overline{p}^{\frac{\alpha_2-\alpha_1-1}{\alpha_1}}\{\alpha_2(1-\omega)^{\alpha_2-1}\omega^{1-\alpha_2}+(\alpha_2-1)(1-\omega)^{\alpha_2}\omega^{-\alpha_2}\}.$$

From this, it follows that

$$\left.\frac{\partial \mathrm{VaR}(\omega,\overline{p})}{\partial \omega}\right|_{\omega=1}=A_1^{\frac{1}{\alpha_1}}\overline{p}^{-\frac{1}{\alpha_1}}>0.$$

Moreover, multiplying the derivative by ω^{α_2} and evaluating the resulting expression at $\omega=0$ gives

$$\omega^{\alpha_2}\left.\frac{\partial \mathrm{VaR}(\omega,\overline{p})}{\partial \omega}\right|_{\omega=0}=-\alpha_1^{-1}A_1^{\frac{1-\alpha_2}{\alpha_1}}A_2\overline{p}^{\frac{\alpha_2-\alpha_1-1}{\alpha_1}}(\alpha_2-1)<0.$$

The second-order derivative at $\omega=\omega^*$ with respect to the portfolio allocation is:

$$\frac{\partial^2 \mathrm{VaR}(\omega,\overline{p})}{\partial \omega^2}=\frac{\alpha_2(\alpha_2-1)}{\alpha_1}A_1^{\frac{1-\alpha_2}{\alpha_1}}A_2\overline{p}^{\frac{\alpha_2-\alpha_1-1}{\alpha_1}}\omega^{-3}\left(\frac{1}{\omega}-1\right)^{\alpha_2-2}$$

which is strictly positive for $\omega\in(0, 1)$ under the stated assumptions. Together these derivatives imply there is an interior minimum. □

It follows that the VaR is convex in the portfolio mix if the distribution of returns have tails which vary regularly at infinity. The VaR criterion thus induces diversification, even though it penalizes asset returns which have a higher asymptotic downside risk than others. Under the stated conditions in Proposition 4, the optimal choice includes the riskier asset for the limited downside risk portfolio.

4. Revisit to Jansen et al. (2000)

We now demonstrate the relevance of the above second order expansion by revisiting applications from the literatures. It will be shown how the second order theory modifies the portfolio selected if one only relies on the first order theory. An example is a study of the safety first criterion by Jansen et al. (2000). We first briefly review the safety first criterion and then present our portfolio choices.

4.1. Safety-first portfolio

Portfolio selection is based on a trade-off between expected return and risk. The risk in the safety-first criterion, initially proposed by Roy (1952) and Arzac and Bawa (1977), is evaluated by the probability of failure. A lexicographic form of the safety first principle is:

$$\max_{\omega_i,b}(\pi,\mu) \text{ lexicographically,}$$

subject to

$$\sum_i \omega_i V_{it} + b = W_t,$$

where $\pi=1$ if $p=P\{\Sigma\omega_i V_{it+1}+br\le s\}\le\delta$, and $\pi=1-p$ otherwise. Furthermore let $\mu=E[\Sigma\omega_i V_{it+1}]+br$, V_{it} denotes the initial market values of asset i at time t, W_t is the initial wealth level of the investor, b denotes the amount of lending or borrowing ($b>0$ represents lending), r is the risk-free gross rate of return, ω_i denotes the weight of invested amount in the risky asset i, s is the disaster level of wealth, and δ gives the maximal acceptable probability of this disaster.

Arzac and Bawa (1977) showed that the safety first problem can be separated into two problems: First, the risk averse safety-first investor maximizes the ratio of the risk premium to the return opportunity loss that he is willing to incur with probability δ, that is

$$\max_{\omega_i}\frac{(\overline{R}-r)}{(r-q_\delta(R))}$$

where $R=\Sigma\omega_i V_{it+1}/\Sigma\omega_i V_{it}$ are the gross returns, $\overline{R}=E(R)$, and $q_\delta(R)$ is a quantile (loss level) sucPh that there is δ% probability of returns less than or equal to this value, that is, the VaR. In the second stage the investor determines the scale of the risky portfolio and the amount borrowed from the budget constraint;

$$W_t-b=\frac{s-rW_t}{q_\delta(R)-r}.$$

For further details on this part, we refer to Arzac and Bawa (1977).

4.2. Empirical illustrations

We re-calculate the optimum portfolio weights for the examples in Jansen et al. (2000) which resulted in a corner solution. By using Proposition 2 and the parameter estimates from Jansen et al. (2000) we obtain an interior solution when we apply the second order theory. The problem consists in choosing between investing in a mutual fund of bonds or a mutual fund of stocks over the period 1926.01–1992.12 with 804 monthly observations of a US bond index and a US stock index (from the CRSP database). We also present, separately, an analysis of the two French stocks Thomson-CSF and L'Oreal, covering 546 daily observations, studied both by Jansen et al. (2000) and Gourieroux et al. (2000).

The Table 1 reproduces the summary statistic and tail indices from Jansen et al. (2000). For US assets the tail index is calculated for the lower tails of the distribution of monthly stock and bond returns. For the daily returns of the two French stocks the calculations combined the data from the upper and lower tails upon the assumption of tail symmetry.

Table 1
Summary statistics and estimates of tail indices

	US bonds and stocks		French stocks	
	Corporate bonds	Stocks	Thomson-CSF	L'Oreal
Mean	0.004445	0.007943	0.0000495	0.0005861
S.D.	0.019782	0.055702	0.01261	0.01129
Skewness	0.746	−0.488	−0.239	0.061
Kurtosis	10.027	9.888	4.114	4.311
No. observations	804	804	546	546
m	16	13	21	13
$X_{(m)}$	−0.03843	−0.13150	0.0275	0.0285
α	2.932	2.601	4.370	4.829
$q_{1/2n}$	−0.125	−0.460	−0.063	−0.056

Note: Tables 1 and 2 are from Jansen et al. (2000). US bond index and a US stock index (1926.01–1992.12), Thomson-CSF and L'Oreal, 546 daily observations. $X_{(n-m)}$ denote the m-th lowest observation for US assets, the m-th largest absolute observation for French stocks respectively. q_δ denotes VaR level corresponding to the probability δ.

From Table 1 we see that the first order tail indices differ. In Jansen et al. (2000) for the case of the two French stocks the safety first criterion allocates all wealth to L'Oreal which has the higher tail index. For the US assets, note that with $r=1$ and and a risk level $\delta=0.000625$ all wealth is allocated to the low risk (higher tail index) bond. Our solutions using the second order approach will be different.

We verify whether the conditions for an interior solution from Proposition 2 do apply. Without loss of generality, we set US stock and Thomson-CSF as X_1. We calculate the second order tail index, β_1, by using the estimates from Table 1. One can calibrate the values of the second order coefficient from Table 1 as follows. A consistent estimator for the ratio between the first and second order tail indices is

$$\widehat{\beta/\alpha} = \frac{\ln \hat{m}}{2\ln n - 2\ln \hat{m}},$$

where n is the number of observations, m is the window size for the estimation of the tail index, see Danielsson et al. (2000). By Proposition 1.7 from Geluk and de Haan (1987) on the properties of regularly varying functions we have that $\frac{\ln \hat{m}}{\ln n} \to \frac{2\beta/\alpha}{1+2\beta/\alpha}$ in probability as $n \to \infty$. Then we use the fact that $\hat{m}/m \to 1$ in probability, where $\hat{m}$ is a consistent estimator of m. Thus, for the US assets, $\beta_1=0.809$ and $\alpha_2-\alpha_1=0.311$, in case of the two French stocks, $\beta_1=1.657$ and $\alpha_2-\alpha_1=0.459$. Thus both cases satisfy the conditions of Proposition 2.

To determine the portfolio mix, we follow the same procedure as in Jansen et al. (2000). We first calculate the VaR quantiles for each hypothetical portfolio.[2] These are reported in Table 2.

[2] We can calculate $A_i, i=1, 2$, used in Jansen et al. (2000) by using

$$A_i = \frac{m_i}{n} X_{(m_i)}^{\alpha_i}$$

where $X_{(m)}$ is the m-th largest observation. Then we plug those values in Proposition 3, and solve the following approximation

$$\omega^{\alpha_1} A_1 q_\delta^{-\alpha_1} + (1-\omega)^{\alpha_2} A_2 q_\delta^{-\alpha_2} \approx \delta$$

to get the value q_δ for the given value of ω and δ.

Table 2
Estimated VaR levels corresponding to the stated probabilities

Portfolio of two assets	US bonds and stocks		French stocks
Probabilities	0.0025	0.000625	0.0018
	(2/804)	(0.5/804)	(1/546)
100% asset 2	−0.2695	−0.4593	−0.0487
90% asset 2	−0.2426	−0.4134	−0.0438
80% asset 2	−0.2157	−0.3675	−0.0390
70% asset 2	−0.1888	−0.3217	−0.0344
60% asset 2	−0.1622	−0.2763	−0.0309[a]
50% asset 2	−0.1361	−0.2316	−0.0305*
40% asset 2	−0.1113	−0.1887	−0.0338
30% asset 2	−0.0896	−0.1505	−0.0389
20% asset 2	−0.0752	−0.1236	−0.0443
10% asset 2	−0.0721*	−0.1163*	−0.0499
0% asset 2	−0.0780[a]	−0.1251[a]	−0.0554

Note: The values in parentheses denote the expected number of occurrences. Asset 2 for the US case is US stocks and asset 2 for the French case is the stock of L'Oreal. * indicates the minimum VaR level among available choices on basis of the second order theory, while [a] indicates the portfolio weight with the minimum VaR level from Jansen et al. (2000).

The investor can borrow or lend at the risk-free rate r, and maximizes $(\bar{R}-r)/(r-q_\delta(R))$. The safety-first investor specifies the desired probability of δ level; the calculations are done for two choices of δ, $\delta=0.0025$, and $\delta=0.000625$. Two interest rates are used, $r=1$ and $r=1.00303$ (the latter corresponds to an annual rate of 3.7%, which equals the average returns on the US Treasury bills over 1926–1992). The mean return $\bar{R}$ is taken from Table 2 by weighting the mean returns on the two assets with the indicated portfolio mix. Optimal portfolios in Table 3 are marked with an asterisk. In all four configurations considered, the optimal portfolio contains 20% stocks and 80% bonds. Fig. 1 illustrates the portfolio choice problem, plotting the mean return versus VaR for portfolios of stocks and bonds when $r=1.003$. For the case $r=1$ and $\delta=0.000625$, Jansen et al. (2000) select a corner solution with 100% bonds. In our procedure, however, stocks are still part of the portfolio.

Empirical analyses of the daily data on the two French stocks are presented in Tables 2 and 4. Fig. 2 illustrates that the limited downside risk portfolio selection criterion chooses a portfolio with 30% of Thomson-CSF stocks and 70% of L'Oreal stocks, not the corner solution as in Jansen et al. (2000).

To conclude, if we take into account the second order terms, solutions are often bounded away from the 100% bond portfolio in the example of US assets, while if only the first order terms are taken into account, a corner solution is repeatedly selected. This may make the portfolio overly conservative, giving up quite a bit of upside potential.

We briefly examine another example from the literature. Susmel (2001) investigates the diversification opportunities which the Latin American emerging markets offer to a US safety first investor. From the portfolio choice problem between an equally weighted Latin American Index and US index, the optimal investment in the Latin American Index is 15% in Susmel's (2001) paper. Instead of an equally weighted Latin American Index, we analyze the optimum portfolio weight for each pair of US and Argentina, US and Brazil, US and Chile, US and Mexico respectively. One can verify that the conditions of Proposition 2 are satisfied for all Latin American stocks combined with US from the estimates in Table 4 of Susmel (2001). Using the same procedure as before, we calculate optimal weights for each pair. For the case of $r=1$ and

Table 3
Portfolio selection for monthly US stocks and bonds

Portfolio	$q_\delta(R)$	$(R-r)/(r-q_\delta)$	$(R-r)/(r-q_\delta)$
		$r=1$	$r=1.00303$
Portfolio selection with $\delta=0.0025$			
100% stock	1−0.2695	0.02947	0.01802
90% stock	1−0.2426	0.03130	0.01858
80% stock	1−0.2157	0.03359	0.01927
70% stock	1−0.1888	0.03650	0.02014
60% stock	1−0.1622	0.04034	0.02126
50% stock	1−0.1361	0.04550	0.02274
40% stock	1−0.1113	0.05252	0.02462[a]
30% stock	1−0.0896	0.06133	0.02661
20% stock	1−0.0752	0.06844*	0.02704*
10% stock	1−0.0721	0.06648[a]	0.02348
0% stock	1−0.0780	0.05701	0.01747
Portfolio selection with $\delta=0.000625$			
100% stock	1−0.4593	0.01729	0.01063
90% stock	1−0.4134	0.01838	0.01096
80% stock	1−0.3675	0.01971	0.01137
70% stock	1−0.3217	0.02143	0.01190
60% stock	1−0.2763	0.02369	0.01258
50% stock	1−0.2316	0.02675	0.01349
40% stock	1−0.1887	0.03097	0.01468[a]
30% stock	1−0.1505	0.03653	0.01606
20% stock	1−0.1236	0.04162*	0.01670*
10% stock	1−0.1163	0.04125	0.01480
0% stock	1−0.1251	0.03553[a]	0.01104

Note: * indicates optimal portfolio among available choices on basis of the second order theory, while [a] indicates the optimal choice from Jansen et al. (2000).

$\delta=0.00289$ (1/346), we find only portfolio weights 1%, 2%, 5% and 2%. For the case of $r=1$ and $\delta=0.001445$ (0.5/346), we find only 1%, 1%, 4% and 2% portfolio weights. These low proportions of Latin American stocks are due to the much higher tail risk (low tail indices) compared to the US.[3] Since the estimated tail indices of US and Latin American markets are very different, from 3.2 to 1.8–2.1 the portfolio selection problems have near corner solutions for all cases.

5. Extension to the multi assets with dependence between assets

We extend the theoretical results to the case of multi asset portfolios with dependent returns. This relaxes the two independent asset portfolio case treated above, and gives the theory more scope.

Typically there exist two types of dependence: over time and cross-sectionally. The time series dependency structure is very common in financial time series. Typically asset return series exhibit

[3] Susmel (2001) proceeds along a different line and selects much higher proportions. The reason is that Susmel (2001) estimates different tail indices for each portfolio combination. This approach, however, biases the tail indices upward (causing underestimation of the risk). This is further clarified in Appendix B.

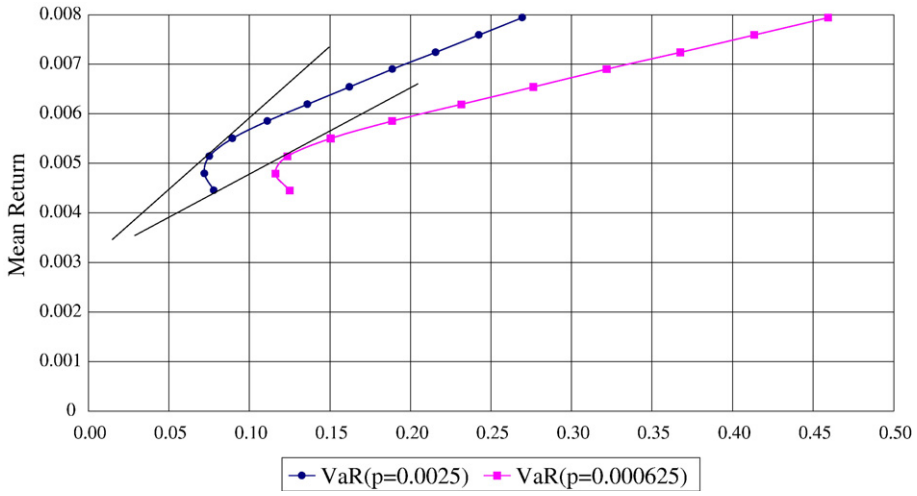

Fig. 1. US stock and bond index.

clusters of high and low volatility. An ARCH process is often used to capture this data feature. Nevertheless, as long as the cross-sectional dependence structure remains time invariant (stationary), the (univariate) time series structure bears no relevance for the (unconditional) cross-sectional portfolio selection rules. Thus we only need to address cross-sectional dependence.

The assumption of cross-sectional independence can be easily weakened. For instance, we can allow for the cross-sectional dependency which arises within the Capital Asset Pricing Model (CAPM). Divide the (excess) return R_i of an asset into the (excess) market return R and the idiosyncratic return Q_i. We apply again Feller's theorem to derive the benefits from cross-

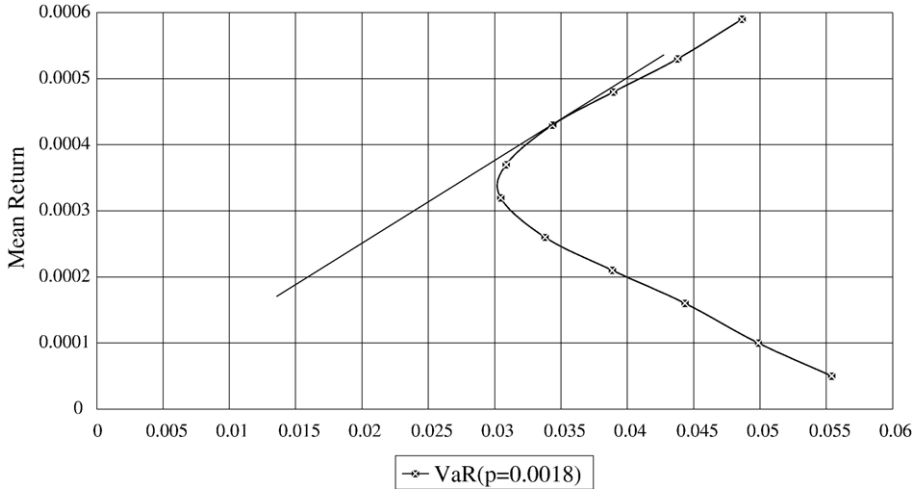

Fig. 2. French stocks.

Table 4
Portfolio selection for daily French stocks

Portfolio	$q_\delta(R)$	$\frac{(R-r)}{(r-q_\delta)}$
		$r=1$
100% L'Oreal	$1-0.048650$	0.01209[a]
90% L'Oreal	$1-0.043786$	0.01218
80% L'Oreal	$1-0.038953$	0.01226
70% L'Oreal	$1-0.034358$	0.01241*
60% L'Oreal	$1-0.030859$	0.01211
50% L'Oreal	$1-0.030450$	0.01037
40% L'Oreal	$1-0.033801$	0.00778
30% L'Oreal	$1-0.038869$	0.00542
20% L'Oreal	$1-0.044338$	0.00352
10% L'Oreal	$1-0.049873$	0.00210
0% L'Oreal	$1-0.055415$	0.00088

Note: * indicates optimal portfolio among available choices on basis of the second order theory, while [a] indicates the optimal choice from Jansen et al. (2000). Portfolio selection is done with $\delta=0.0018$.

sectional portfolio diversification. Under the CAPM the following one factor model for n-assets applies

$$R_i = b_i R + Q_i, i = 1, \ldots, n. \tag{3}$$

Moreover, assume that the Q_i's are independent from each other. Suppose the idiosyncratic risk factors Q_i have distributions with tails

$$P\{Q_i<-s\} = A_i s^{-\alpha_i}[1 + B_i s^{-\beta_i} + o(s^{-\beta_i})], i = 1, \ldots, n, \tag{4}$$

and the market risk has a distribution with a tail

$$P\{R<-s\} = A_R s^{-\alpha_R}[1 + B_R s^{-\beta_R} + o(s^{-\beta_R})], \tag{5}$$

For simplicity we assume[4]

$$2<\alpha_1<\ldots<\alpha_n<\alpha_R.$$

We further assume that for each h, $h=2,\ldots, n$,

$$\alpha_h<\min[(\alpha_1 + \beta_1), \ldots, (\alpha_{h-1} + \beta_{h-1}), (\alpha_1 + 1)],$$

and

$$\alpha_R<\min[(\alpha_1 + \beta_1), \ldots, (\alpha_n + \beta_n), (\alpha_1 + 1)]$$

[4] Similar results follow if one assumes $2<\alpha_R<\alpha_1<\ldots<\alpha_n$ or $2<\alpha_1<\ldots<\alpha_R<\ldots<\alpha_n$.

to ensure that any of the second order tail indices do not appear in the following expression. In the case of positive portfolio weights λ_i, by Feller's theorem

$$P\{\sum_{i=1}^{n} \lambda_i R_i < -s\} \sim \sum_{i=1}^{n} \lambda_i^{\alpha_i} A_i s^{-\alpha_i} + \left(\sum_{i=1}^{n} \lambda_i b_i\right)^{\alpha_R} A_R s^{-\alpha_R} \tag{6}$$

as $s \to \infty$.

Proposition 5. *Suppose α_i, $\alpha_R > 2$, $\lambda_i \geq 0, \Sigma_{i=1}^{n} \lambda_i = 1$ and $b_i > 0$, then the function*

$$f(\lambda) = \sum_{i=1}^{n} \lambda_i^{\alpha_i} A_i s^{-\alpha_i} + \left(\sum_{i=1}^{n} \lambda_i b_i\right)^{\alpha_R} A_R s^{-\alpha_R}$$

is convex.

Proof. In Appendix C we prove that the Hessian $n \times n$ matrix

$$H \equiv \left[\frac{\partial^2 f(\lambda)}{\partial \lambda_i \partial \lambda_j}\right] \text{is positive definite.}$$ □

A corner solution can be ruled out by the following argument. For any choice such that $\lambda_i = 1$ for one i but $\lambda_j = 0$ for all $j \neq i$, the function $f(\lambda)$ is not minimal since $\frac{\partial f(\lambda)}{\partial \lambda_i}|_{\lambda_i = 1, \lambda_j = 0}$ for all $j \neq i = \alpha_i A_i s^{-\alpha_i} + \alpha_R b_i^{\alpha_R} A_R s^{-\alpha_R} > 0$ for all i.

To prove the convexity of VaR, we provide a heuristic explanation instead of doing De Bruijn inversion for multiple assets. By the above convexity result we know that at a given VaR-level s, there exists optimal $\widehat{\lambda}_i$ weights such that

$$p\{\sum_{i=1}^{n} \widehat{\lambda}_i X_i < -s\} < P\{\sum_{i=1}^{n} \lambda_i X_i < -s\}.$$

Denote the associated probability-level as pp

$$\text{pp} = P\{\sum_{i=1}^{n} \widehat{\lambda}_i X_i < -s\}.$$

Given this particular probability-level pp, it must be the case that the VaR levels at pp are increasing if one deviates from the optimal $\widehat{\lambda}_i$ weights and uses any other λ_i. This follows from the fact that the probability for a given set of asset weights λ_i

$$P\{\sum_{i=1}^{n} \lambda_i X_i < -s\} \tag{7}$$

is (weakly) decreasing in the VaR-level s. To bring down the probability (7) with weights $\widehat{\lambda}_i \neq \lambda_i$ to the level pp, we have to raise s since a distribution function is monotonic in s.

6. Conclusion

We consider the portfolio problem of choosing the mix between stocks and bonds. Investors often worry about the downside risk features of their portfolio. It is a fact that asset return

distributions exhibit fat tails, i.e. are asymptotic to a Pareto distribution. Typically stocks exhibit fatter tails than bonds corresponding to the greater downside risk of stocks. Downside risk criteria like the safety first criterion therefore often select corner solutions in the sense of a bonds only portfolio. This is due to a focus on the tail of the asset return distributions whereby only the asymptotically dominating first order Pareto term is taken into account. We extend the first order convolution result to a second order asymptotic expansion. Whereas in the first order convolution result only the fattest of the two tails plays a role, in the second order expansion often the tails of both assets play a role. We suggest that with a second order expansion of the downside risk, the portfolio solution may yield a balanced solution, i.e. both assets are held in non-zero proportion, whereas the first order expansion selects the corner solution. The theoretical results were extended to multi asset portfolios with dependent returns.

In the empirical application, we follow up on Jansen et al. (2000), who apply the safety first criterion to a number of portfolio problems. In the cases where Jansen et al. (2000) give a corner solution, our procedure still selects both assets for incorporation in the limited downside risk portfolio. We also briefly addressed another example from the literature.

Acknowledgement

Both authors like to thank the Tinbergen Institute for support. Hyung is gratefully acknowledges support by the Korea Research Foundation Grant (KRF-2003-003-B00103).

Appendix A. Proof of Theorem 1

For the calculation of $P\{X_1+X_2\le s\}$, we divide the area over which we have to integrate into five parts A, B, C, D and E; where $P\{A\}=P\left\{X_1+X_2\le s, X_1>-\frac{s}{2}, X_2>\frac{s}{2}\right\}$, $P\{C\}=P\{X_1\le \frac{s}{2}, X_2\le\frac{s}{2}\}, P\{D\}=P\left\{X_1+X_2\le s, X_1\le-\frac{s}{2}, X_2>\frac{s}{2}\right\}$, and where $P\{B\}$ and $P\{E\}$ are the counterparts of $P\{A\}$ and $P\{D\}$ respectively. We start by $P\{C\}$:

$$\begin{aligned} p\{C\} &= P\left\{X_1\le\frac{s}{2}, X_2\le\frac{s}{2}\right\} = P\left\{X_1\le\frac{s}{2}\right\}P\left\{X_2\le\frac{s}{2}\right\} \\ &= 1-A_1\left(\frac{s}{2}\right)^{-\alpha_1}-A_1B_1\left(\frac{s}{2}\right)^{-\alpha_1-\beta_1}-A_2\left(\frac{s}{2}\right)^{-\alpha_2}-A_2B_2\left(\frac{s}{2}\right)^{-\alpha_2-\beta_2}+o(s^{-\gamma}) \end{aligned}$$

as $s\to\infty$. The terms which are of smaller order than $s^{-\gamma}$, $\gamma=\alpha_1+\beta_1$, α_1+1, $\alpha_2+\beta_2$, or α_2+1, can be ignored throughout this proof. The probability $P\{A\}$ takes more effort

$$\begin{aligned} P\{A\} &= P\left\{X_1+X_2\le s, X_1>-\frac{s}{2}, X_2>\frac{s}{2}\right\} = \int_{-s/2}^{s/2}\left[F_2(s-x)-F_2\left(\frac{s}{2}\right)\right]f_1(x)\mathrm{d}x \\ &= \int_{-s/2}^{s/2}F_2(s-x)f_1(x)\mathrm{d}x-\int_{-s/2}^{s/2}F_2\left(\frac{s}{2}\right)f_1(x)\mathrm{d}x = I-II, \end{aligned}$$

where $f_i(\cdot)$ and $F_i(\cdot)$ denote respectively the density function and distribution function of X_i. For integral I note that a Taylor series around $x=0$ with remainder gives

$$(s-x)^{-\alpha}=s^{-\alpha}+\alpha s^{-\alpha-1}x+\frac{(\alpha+1)\alpha}{2}(s-q)^{-\alpha-2}x^2,$$

where q is some number between $\left[-\frac{s}{2},\frac{s}{2}\right]$. Hence, for large s

$$\begin{aligned} I &= [1-A_2 s^{-\alpha_2}-A_2B_2 s^{-\alpha_2-\beta_2}+o(s^{-\alpha_2-\beta_2})]\int_{-s/2}^{s/2} f_1(x)\mathrm{d}x-[\alpha_2 A_2 s^{-\alpha_2-1}+o(s^{-\alpha_2-1})] \\ &\quad\times\int_{-s/2}^{s/2} x f_1(x)\mathrm{d}x+o(s^{-\gamma})=[1-A_2 s^{-\alpha_2}-A_2B_2 s^{-\alpha_2-\beta_2}] \\ &\quad\times\left[1-2A_1\left(\frac{s}{2}\right)^{-\alpha_1}-2A_1B_1\left(\frac{s}{2}\right)^{-\alpha_1-\beta_1}+o(s^{-\alpha_1-\beta_1})\right] \\ &\quad-\alpha_2 A_2 s^{-\alpha_2-1}[E[X_1]+O(s^{-\alpha_1+1})]+o(s^{-\gamma}). \end{aligned}$$

Note that $(s-q)^{-\alpha-2}\le\left(\frac{3}{2}s\right)^{-\alpha-2}$ for $q\in\left[-\frac{s}{2},\frac{s}{2}\right]$. Thus

$$\frac{(\alpha_2+1)\alpha_2}{2}A_2(s-q)^{-\alpha_2-2}\int_{-s/2}^{s/2}x^2 f_1(x)dx\le\frac{(\alpha_2+1)\alpha_2}{2}A_2\left(\frac{3}{2}s\right)^{-\alpha_2-2}E[X_1^2]=o(s^{-\alpha_2-1}).$$

Hence the $o(s^{-\gamma})$ in the integral I expression.

And for part II

$$\begin{aligned} II &= F_2\left(\frac{s}{2}\right)\int_{-s/2}^{s/2} f_1(x)\mathrm{d}x=\left[1-A_2\left(\frac{s}{2}\right)^{-\alpha_2}-A_2B_2\left(\frac{s}{2}\right)^{-\alpha_2-\beta_2}+o(s^{-\alpha_2-\beta_2})\right] \\ &\quad\times\left[1-2A_1\left(\frac{s}{2}\right)^{-\alpha_1}-2A_1B_1\left(\frac{s}{2}\right)^{-\alpha_1-\beta_1}+o(s^{-\alpha_1-\beta_1})\right] \end{aligned}$$

Combine the two parts to obtain $P\{A\}$.

$$\begin{aligned} P\{A\} &= I-II=-A_2 s^{-\alpha_2}-A_2B_2 s^{-\alpha_2-\beta_2}+A_2\left(\frac{s}{2}\right)^{-\alpha_2}+A_2B_2\left(\frac{s}{2}\right)^{-\alpha_2-\beta_2} \\ &\quad-\alpha_2 A_2 s^{-\alpha_2-1}E[X_1]+o(s^{-\gamma}). \end{aligned}$$

The probability $P\{D\}$ is

$$P\{D\}=P\left\{X_1+X_2\le s, X_1\le-\frac{s}{2}, X_2>\frac{s}{2}\right\}=\int_{-\infty}^{-s/2}\left[F_2(s-x)-F_2\left(\frac{s}{2}\right)\right]f_1(x)\mathrm{d}x=o(s^{-\gamma})$$

Similar expressions hold for $P\{B\}$and $P\{E\}$. □

Appendix B. Bias in $\hat{\alpha}$

Suppose that the tails of the distributions of X satisfy $\mathrm{P}\{X>s\}=As^{-\alpha}\ [1+Bs^{-\beta}+o\ (s^{-\beta})]$ as $s\to\infty$, where $\alpha>0$, $A>0$, $\beta>0$ and B is a real number. The asymptotic bias for the Hill estimator $\widehat{1/\alpha}$ is

$$E\left[\widehat{1/\alpha}-1/\alpha\right]=\frac{-B\beta}{\alpha(\alpha+\beta)}s^{-\beta}+o(s^{-\beta})$$

as $s \to \infty$ in Goldie and Smith (1987). For the portfolio from case I in Theorem 1, the asymptotic bias of the Hill estimator is

$$\text{Bias}(\widehat{1/\alpha}) = -\frac{(1-\omega)^{\alpha_2}}{\omega^{\alpha_1}}\frac{A_1}{A_2}\frac{(\alpha_2-\alpha_1)}{\alpha_1\alpha_2}s^{-(\alpha_2-\alpha_1)} + o\left(s^{-(\alpha_2-\alpha_1)}\right)$$

where

$$-\frac{(1-\omega)^{\omega_2}}{\omega^{\alpha_1}}\frac{A_1}{A_2}\frac{(\alpha_2-\alpha_1)}{\alpha_1\alpha_2}s^{-(\alpha_2-\alpha_1)}<0$$

which proves the upward bias in the tail estimator $\hat{\alpha}$.

Appendix C. Proof of positive definiteness of *H*

The elements of H are

$$H_{ii} \equiv \frac{\partial^2 f(\lambda)}{\partial \lambda_i^2} = c_i + b_i^2 c_R$$
$$H_{ij} \equiv \frac{\partial^2 f(\lambda)}{\partial \lambda_i \partial \lambda_j} = b_i b_j c_R,$$

where b_i is in Eq. (3) and

$$c_i = \alpha_i(\alpha_i-1)\lambda_i^{\alpha_i-2}A_i s^{-\alpha_i}>0$$
$$c_R = \alpha_R(\alpha_R-1)\left(\sum_{i=1}^{n}\lambda_i b_i\right)^{\alpha_R-2}A_R s^{-\alpha_R}>0,$$

given the assumptions of α_i, $\alpha_R>2$, $\lambda_i \geq 0$, and $b_i>0$ for all $i=1,\ldots, n$.

Then matrix H can be decomposed as

$$H = C + c_R BB',$$

where C is a diagonal matrix of $(c_1,\ldots, c_n)$, $B'=[b_1,\ldots, b_n]$. We show $x'Hx>0$ for any non-zero vector x. Note that the quadratic expression can be split into two parts

$$x'Hx = x'(C + c_R BB')x = x'Cx + c_R x'BB'x$$

Since C is positive definite as all diagonal elements $c_i>0$ and the other part is a quadratic expression $x'BB'x=(x'B)^2>0$, it follows that H is positive definite.

References

Arzac, E.R., Bawa, V.S., 1977. Portfolio choice and equilibrium in capital markets with safety-first investors. Journal of Financial Economics 4, 277–288.

Bingham, N.H., Goldie, C.M., Teugels, J.L., 1987. Regular Variation. Cambridge University Press, Cambridge.

Dacarogna, M.M., Müller, U.A., Pictet, O.V., de Vries, C.G., 2001. Extremal foreign exchange returns in extremely large data sets. Extremes 4 (2), 105–127.

Danielsson, J., de Haan, L., Peng, L., de Vries, C.G., 2000. Using a bootstrap method to choose the sample fraction in tail index estimation. Journal of Multivariate Analysis 76, 226–248.

de Haan, L., Stadtmüller, U., 1996. Generalized regular variation of second order. Journal of the Australian Mathematical Society 61, 381–395 (series A).

Feller, W., 1971. An Introduction to Probability Theory and Its Applications, vol. II. New York, Wiley.

Geluk, J., de Haan, L., 1987. Regular variation, extensions and Tauberian theorems. CWI Tract, vol. 40. CWI, Amsterdam.

Goldie, C.M., Smith, R.L., 1987. Slow varition with remainder: theory and application. Quarterly Journal of Mathematics 38, 45–71 (Oxford 2nd series).

Gourieroux, C., Laurent, J.P., Scaillet, O., 2000. Sensitivity analysis of values at risk. Journal of Empirical Finance 7, 225–246.

Hartmann, P., Straetmans, S., de Vries, C.G., 2004. Asset market linkages in crisis periods. Review of Economics and Statistics 86, 313–326.

Jansen, D., Limited downside risk in portfolio selection among U.S. and Pacific Basin equities. mimeo 2000.

Jansen, D., Koedijk, K.G., de Vries, C.G., 2000. Portfolio selection with limited down-side risk. Journal of Empirical Finance 7, 247–269.

Poon, S., Rockinger, M., Tawn, J., 2003. Nonparametric extreme value dependence measures and finance applications. Statistica Sinica 13, 929–953.

Roy, A.D., 1952. Safety first and the holding of assets. Econometrica 20, 431–449.

Susmel, R., 2001. Extreme observations and diversification in Latin American emerging equity markets. Journal of International Money and Finance 20, 971–986.

Straetmans, S., 1998. Extreme Financial Returns and their Comovements. Ph.D. thesis: Tinbergen Institute Research Series; Erasmus University Rotterdam.