The cross-section of tail risks in stock returns

Moore, Kyle and Sun, Pengfei and de Vries, Casper G. and Zhou, Chen

Erasmus University Rotterdam, De Nederlandsche Bank, Chapman University

19 February 2013

The Cross-Section of Tail Risks in Stock Returns*

Kyle Moore†
Erasmus University Rotterdam and Tinbergen Institute

Pengfei Sun‡
Erasmus University Rotterdam and Tinbergen Institute

Casper de Vries§
Erasmus University Rotterdam and Chapman University

Chen Zhou¶
De Nederlandsche Bank and Erasmus University Rotterdam

March 25, 2013

Abstract This paper investigates how the downside tail risk of stock returns is differentiated cross-sectionally. Stock returns follow heavy-tailed distributions with downside tail risk determined by the tail shape and scale. If safety-first investors are concerned with sufficiently large downside losses, i.e. have a sufficiently low risk tolerance, then in the equilibrium, assets traded in the same market share a homogeneous tail shape parameter. Furthermore, if tail shapes are homogeneous, the equilibrium prices of assets are differentiated by the scales.

Keywords: Heavy-tail distribution, safety-first utility, asset pricing

JEL Classification Numbers: G11, G12

*The research of Kyle Moore and Pengfei Sun has received funding from the European Community's Seventh Framework Programme FP7-PEOPLE-ITN-2008. The funding is gratefully acknowledged. Casper is grateful to the ESI at Chapman University for its hospitality.

†Email: kmoore@ese.eur.nl

‡Email: sun@ese.eur.nl

§Email: cdevries@ese.eur.nl

¶Views expressed are those of the authors and do not reflect the official position of De Nederlandsche Bank. Email: c.zhou@dnb.nl, zhou@ese.eur.nl

1 Introduction

Investors supposedly tradeoff expected return against the risk of an asset in composing their portfolios. What constitutes risk can be viewed in different ways. Traditionally, risk is often captured by the variance of the returns. An alternative view focuses on the downside risk. In the banking industry and among professional investors, downside risk has gained popularity through the Value-at-Risk (VaR) concept. In this paper, we study the properties of downside tail risk of assets selected by safety-first investors. Safety-first investors care foremost about meeting a minimum risk constraint before maximizing expected returns.

It is first argued that the tail risk of assets can be represented by a shape and scale parameter in the case that the return distribution is heavy tailed. Given this, we show that safety-first investors with a low risk tolerance only invest in the asset with the largest tail shape parameter, i.e. asset with the thinnest tail. At the same time, assets can differ regarding the tail scale, as this can be compensated for in the expected return. The equilibrium priced derived from the scale heterogeneity is as that in the safety-first analysis of Arzac and Bawa (1977). We are able to achieve more specific results in comparison to methods utilizing the Chebyshev bound, which incorporates a mean-variance analysis.

Empirical research has shown that the tail risk of equity is heavier than that tail risk implied by the normal distribution, see, e.g., Mandelbrot (1963) and Jansen and De Vries (1991). Heavy-tails refer to the fact that the tail region of the distribution function exhibits a power law decay, as opposed to the exponential decay of the Gaussian distribution. Mathematically, denote the return of a financial asset by R, with the distribution function $F(x) = \Pr(R \leq x)$. The distribution function F is heavy-tailed if its left tail can be approximated by a power law as

$$F(-x) = \Pr(R \leq -x) \sim Ax^{-\alpha}, \quad \text{as } x \to \infty, \tag{1.1}$$

where α is the tail shape parameter, commonly referred to as the *tail index*, and A

indicates the *scale* of the distribution.[1] In contrast to the Gaussian distribution, heterogeneity in the downside tail risk of a heavy-tailed distribution is manifest via differences in the tail shape and scale of the distribution.

We develop an asset pricing model based on the safety-first utility to show that when investors exhibit a sufficiently low tolerance for tail risk, the tail index, α, of equity returns traded on the same market are cross-sectionally homogeneous. This is called the *tail index equivalence* hypothesis. We prove this result under a first-order tail risk approximation as in (1.1), then with a further extension under a second-order approximation on the left tail of the distribution function F. Under the second-order approximation, we also show that if investors exhibit a moderate tolerance for tail risk, the shape parameters of the tail distribution may nevertheless differ across different assets. When the tail index equivalence hypothesis holds, the heterogeneity in the downside tail risk of asset returns is then attributed to that in the *s*cale of the tail distribution. In this case, we show that the equilibrium expected return differences compensate for heterogeneity in the scales.

The tail index equivalence hypothesis has been adopted in some theoretical works on portfolio diversification with heavy-tails as a starting assumption. For example, Fama and Miller (1972) consider symmetric stable distributions for modeling the distribution of asset returns. Assuming that the tail shape is invariant, they consider the scale parameter as the only heterogeneous risk parameter in the model. Similarly, Hyung and de Vries (2002) apply the Pareto distribution assuming a constant tail index across assets. More recent works incorporate dependence across assets by assuming a constant tail index cross-sectionally, see, e.g. Hyung and de Vries (2005, 2007); Ibragimov and Walden (2008); Zhou (2010). In all aforementioned studies, there is no theoretical justification on the assumption of tail index equivalence or non-equivalence for that matter. The only weak evidences on the tail index equivalence hypothesis are provided in empirical literature, see, e.g Jansen and De Vries (1991), Loretan and Phillips (1994) and Jondeau and Rockinger (2003). In this paper, we provide an economic argument behind this remarkable equivalence.

[1]The power law approximation belongs to the Fréchet type of domain of attraction as in Extreme Value Theory, see e.g. Embrechts et al. (1997) and De Haan and Ferreira (2006).

The fundamental setup of our model is based on the safety-first utility, which emphasizes the asymmetric preference of investors towards loss aversion over upside potential; see, e.g. Roy (1952)[2] Arzac and Bawa (1977). formalize the safety-first utility in a lexicographic form. Safety-first investors initially minimize the probability that their portfolio return falls below a minimum desired threshold. Only after that is achieved, they seek to maximize expected return.

Instead of fixing a loss level, we use an equivalent downside risk measure, the VaR by fixing the probability level of an extreme loss. For a given probability level, the VaR is defined as the threshold value such that losses on a portfolio are only exceeded with that probability. For a safety-first investor, to control that the probability of portfolio return falls below a minimum desired threshold, it is equivalent to constrain that the Value-at-Risk (VaR) of the portfolio is below a threshold level. Since the downside tail of a heavy-tailed distribution can be approximated as in (1.1), for low p, the VaR at a probability level p has an explicit approximation as

$$VaR^p(R) \approx (\frac{A}{p})^{\frac{1}{\alpha}}.$$

From the VaR approximation, we observe that once the tail indices differ, for extremely low probability level p, the asset with a lower tail index will always have a higher VaR. In addition, the VaRs corresponding to different tail indices are at incomparable levels when the probability level p tends to zero. This serves as the general intuition for the tail index equivalence hypothesis: by considering rather low probability level p, the differences in VaRs can not be compensated by expected returns. We formalize this intuition into our theoretical analysis.

The paper is organized as follows. In Section 2, we provide the theoretical model on tail index equivalence under first-order and second-order tail approximations. In Section 3, we derive the equilibrium prices of assets under the tail index equivalence hypothe-

[2] The modeling of downside risk aversion was also investigated by Markowitz (1959) where he suggests the use of semi-variance, as opposed to variance, since it captures only the downside losses. Kahneman and Tversky (1979) use a behavioral approach arguing that agents weigh gains and losses differently. More recently, work by Harvey and Siddique (2000) and Ang et al. (2006) explore firm-level downside risk by using co-skewness in equity returns and downside betas, respectively.

sis and show that heterogeneous scale parameters of asset returns are compensated by expected return. Section 4 concludes the paper.

2 Theory

2.1 Investors with Safety-First Utility

Safety-first preference, first introduced by Roy (1952), assumes that an investor seeks to maximize the expected return subject to a downside risk constraint. Arzac and Bawa (1977) formalize this approach by introducing lexicographic safety-first utility as follows. Suppose an investor has total wealth W_t at time t. The investor can borrow or lend at a risk-free rate r_f and invest in a set of assets with prices P_t at time t. Here P_t denotes the vector of asset prices. The investor chooses to construct a portfolio consisting of the risky assets with portfolio holding indicated by ω and the risk-free investment indicated by b, where b can be either positive (lending) or negative (borrowing). Being safety-first, the investor considers the probability that the value of his portfolio at time $t+1$ is below a critical level s. Then, the safety-first investor solves the following utility maximization problem by choosing (ω^T, b):

$$\max\{\pi, \mu\}, \text{ s.t. } \omega^T P_t + b = W_t,$$

where

$$\pi = 1 \text{ if } \zeta = \Pr(\omega^T P_{t+1} + br_f \leq s) \leq p$$

$$\pi = 1 - \zeta, \text{ otherwise,}$$

and

$$\mu = E[\omega^T P_{t+1} + br_f].$$

Here, the lexicographic preference order is captured by $\max\{\pi, \mu\}$ as investors first consider assets such that the probability of having a low wealth at time $t+1$, ζ, is low. If ζ is below the admissible probability level p, then the preference is towards a high expected return μ.

Note that if $W_t r_f > s$, the case $\pi = 1$ can always be achieved by investing sufficiently on the risk-free asset. In that case, the utility maximization problem becomes a maximization problem on the expected return with the constraint that $\zeta \le p$. The constraint can be further written as

$$\begin{aligned} p \ge \Pr(\omega^T P_{t+1} + b r_f \le s) &= \Pr(\frac{\omega^T P_{t+1}}{\omega^T P_t} \le \frac{s - b r_f}{W_t - b}) \\ &= \Pr(R_t \le r_f + \frac{s - W_t r_f}{W_t - b} - 1), \end{aligned} \tag{2.1}$$

where $R_t = \frac{\omega^T P_{t+1}}{\omega^T P_t} - 1$ denotes the return of the risky portfolio. The probability constraint (2.1) is equivalent to a VaR constraint:

$$VaR^p(R) \le -\left(r_f + \frac{s - W_t r_f}{W_t - b} - 1 \right).$$

The utility optimization problem consequently turns to be a maximization problem of the expected return subject to a VaR constraint. Hyung and de Vries (2012) show that the expected return maximization with a VaR constraint is analogous to the mean-variance portfolio selection approach of Markowitz (1959): the VaR replaces the variance as the risk measure.

In the rest of the paper, we consider safety-first investors to be maximizing the expected return under a VaR constraint. Furthermore, we assume that safety-first investors cross-sectionally hold homogeneous concerns on the admissible probability level p that can vary over time.

2.2 Tail Index Equivalence: First-Order Tail Approximation

In this section, we show that the shape parameter that characterizes the tails of asset returns must be equal when p is an extremely low admissible probability level. We start with considering two risky assets only.

Suppose the returns of two assets, R_1 and R_2, both follow heavy-tailed distributions as in (1.1). Denote the tail indices of R_1 and R_2 by α_1 and α_2, with respective scale parameters A_1 and A_2. In the case of $\alpha_1 \neq \alpha_2$, we assume $1 < \alpha_1 < \alpha_2$ without loss of generality[3]. Consider a linear portfolio, R, as

$$R = \omega_1 R_1 + \omega_2 R_2, \quad \omega_1,\ \omega_2 > 0,$$

where ω_1 and ω_2 are the weights assigned to asset 1 and 2, respectively. If R_1 and R_2 are independent, it immediately follows from Embrechts et al. (1997) that if $\alpha_1 < \alpha_2$, the tail index of R is dominated by the lower tail index, α_1. Zhou (2010) shows that this result also holds in case of tail dependence in most finance models, e.g., the Capital Asset Pricing Model (CAPM). The tail distribution of R can be approximated as follows

$$\Pr(R \leq -x) \sim \omega_1^{\alpha_1} A_1 x^{\alpha_1}, \quad \text{as} \quad x \to +\infty.$$

This result holds regardless of the dependence structure between R_1 and R_2. Thus,

$$VaR^p(R) \sim \omega_1 \left(\frac{A_1}{p}\right)^{1/\alpha_1} \quad \text{as} \quad p \to 0. \tag{2.2}$$

From a tail risk perspective, (2.2) implies that only the asset with the heavier tail determines the tail risk of the portfolio. We compare this to an alternative investment strategy of only holding the thinner-tailed asset (asset 2). In this case, the VaR is approximated by $(\frac{A_2}{p})^{1/\alpha_2}$. As $p \to 0$, $\frac{\omega_1 (A_1/p)^{1/\alpha_1}}{(A_2/p)^{1/\alpha_2}} \to \infty$. Hence, the difference in tail indices leads to a difference in the asymptotic levels of the VaR. At sufficiently low risk tolerance levels, this will ultimately overshadow the difference in expected returns. In the following theorem we claim that for sufficiently small p, regardless the expected returns, a portfolio which invests a positive fraction in the asset with the heavier tail, even a small fraction, is not sufficiently compensated by the expected return, compared to solely investing on the thinner tailed asset.

Theorem 2.1 *Suppose the distributions of the asset returns R_1, R_2 follow the approximation in (1.1), with respective tail indices $1 < \alpha_1 < \alpha_2$, scale parameters A_1, A_2, and*

[3] The assumption guarantees that the expected returns are finite.

expected returns $\mu_1,\ \mu_2 > r_f > 0$, *where* r_f *is the return on a risk-free asset. Then, as* $p \to 0$, *a safety-first investor will hold a portfolio consisting of only asset 2 and the risk-free asset.*

Proof. We first show that, for any given portfolio consisting of the two risky assets and the risk-free asset, with sufficiently low probability level p, there always exists another portfolio consisting of the thinner-tailed asset (asset 2) and the risk-free asset, such that this alternative portfolio has the equivalent expected return, but a lower VaR level.

The fraction of wealth invested in two risky assets are $\omega_1,\ \omega_2$, where $\omega_1,\ \omega_2 > 0$. The remainder, $1-\omega_1-\omega_2$, is then invested on the risk-free rate. Thus, the return of the safety-first investor's portfolio R is

$$\begin{aligned} R(\omega_1,\omega_2) &= \omega_1 R_1 + \omega_2 R_2 + (1-\omega_1-\omega_2) r_f \\ &= r_f + \omega_1 (R_1 - r_f) + \omega_2 (R_2 - r_f). \end{aligned}$$

Since the heavy-tail property is location-invariant, the excess returns $R_1 - r_f$ and $R_2 - r_f$ follow heavy-tailed distributions with tail indices $\alpha_1,\ \alpha_2$ and scales $A_1,\ A_2$, respectively. From (2.2), for low p, the VaR of R is approximately

$$VaR^p(R) \approx r_f + \omega_1 \left(\frac{A_1}{p}\right)^{1/\alpha_1}.$$

Alternatively, consider a second portfolio, R', which contains only the thinner tailed risky asset along with the risk-free asset with nevertheless the same expected return:

$$\begin{aligned} R'(\omega_1,\omega_2) = &\left(1 - \omega_1 \frac{\mu_1 - r_f}{\mu_2 - r_f} - \omega_2\right) r_f + \left(\omega_1 \frac{\mu_1 - r_f}{\mu_2 - r_f} + \omega_2\right) R_2 \\ = &\, r_f + \left(\omega_1 \frac{\mu_1 - r_f}{\mu_2 - r_f} + \omega_2\right)(R_2 - r_f). \end{aligned}$$

It is straightforward to verify that $E(R') = E(R)$. For low p, the VaR of the alternative portfolio R' can be approximated as

$$VaR^p(R') \approx r_f + \left(\omega_1 \frac{\mu_1 - r_f}{\mu_2 - r_f} + \omega_2\right) \left(\frac{A_2}{p}\right)^{\frac{1}{\alpha_2}}.$$

From $\alpha_1 < \alpha_2$, we get that

$$\lim_{p \to 0} \frac{\omega_1 (A_1/p)^{1/\alpha_1}}{\left(\omega_1 \frac{\mu_1 - r_f}{\mu_2 - r_f} + \omega_2\right)(A_2/p)^{1/\alpha_2}} = \lim_{p \to 0} \frac{\omega_1 A_1^{1/\alpha_1}}{\left(\omega_1 \frac{\mu_1 - r_f}{\mu_2 - r_f} + \omega_2\right) A_2^{1/\alpha_2}} p^{1/\alpha_2 - 1/\alpha_1} = +\infty.$$

Thus, for sufficiently low p, we have that $VaR^p(R') < VaR^p(R)$.

To summarize, for any portfolio based on a linear combination of both heavy-tailed assets and the risk-free asset, $R(\omega_1, \omega_2)$, if the admissible probability p is sufficiently low, there exists an alternative portfolio containing only the thinner tailed asset and the risk-free asset, R', which has equal expected return and lower downside tail risk. Thus the VaR of portfolio R' is strictly below the threshold in the VaR constrain. Since the VaR of R' is a continuous function with respect to the weight on asset 2, we can increase the weight on asset 2 with a small marginal increment to construct a third portfolio R^*, such that R^* still satisfies the VaR constraint. Meanwhile, since $\mu_2 > r_f$, the portfolio R^* has a strictly higher expected return than R', i.e. $E(R^*) > E(R') = E(R)$. Based on the safety-first utility function, the investor will strictly prefer R^* to R. This completes the proof of the theorem. ■

Theorem 2.1 shows that in a market with two risky assets and a risk-free asset, as the admissible probability level in the safety-first utility tends to zero, the asset with a lower tail index will not be traded in any optimal portfolio. Thus, only the one with a higher tail index will be traded. This can be extended to a market with multiple assets: the assets that share the maximum tail index will be traded, whereas other assets with lower tail indices will not be traded. Hence, we conclude that, in an economy populated by safety-first investors exhibiting a low risk tolerance, all assets that are traded must share the same tail index. We call this the "tail index equivalence hypothesis".

2.3 Tail Index Equivalence: Second-Order Tail Approximation

Above, we have established the "tail index equivalence hypothesis" under the assumption that investors' admissible probability tends to zero, i.e. $p \to 0$. This assumption may not reflect the utility of an actual safety-first investor, because investors have downside tail risk concerns with a small, but nevertheless positive, admissible probability. Said otherwise, the risk tolerance may not be so low as to drive all assets with a lower tail index out of the market. To analyze the preferences of investors with a low, but not tending to zero, admissible probability, we need to quantify the accuracy of the first-order approximations. This is achieved by considering the second-order approximation of the left tail of the asset return distribution.

Similar to the proof in Subsection 2.2, we start by considering a market consisting of two risky assets and a risk-free asset. We assume that the left tail region of the distribution functions on the two asset returns follow heavy-tailed distributions as before, but now with a second-order approximation: for $i = 1, 2$,

$$\Pr(R_i \le -x) \sim A_i x^{-\alpha_i}[1 + Bx^{-\gamma_i}], \quad \text{as} \quad x \to +\infty. \tag{2.3}$$

This approximation applies to most of the standard distribution with heavy tails, e.g., the Student-t distribution. Here γ_i and B_i are called the second-order index and the second-order scale. To simplify the second-order approximation on the tail of the distribution function of the portfolio, we first assume that R_1 and R_2 are independent.

The following theorem shows that if the admissible probability p is sufficiently low, the asset with a lower tail index, i.e. the heavier tail, will not be traded in the market, whereas for a sufficiently high probability level, both asset may be included in the optimal portfolio.

Theorem 2.2 *Suppose distributions of the two asset returns R_1, R_2 follow second-order approximation as in* (2.3)*. Denote the expected returns of R_1 and R_2 as $\mu_1, \mu_2 > r_f$. Suppose the tail indices satisfy $1 < \alpha_1 < \alpha_2$. Then, there exists a positive probability $p^* > 0$, such that for $p < p^*$, a safety-first investor with admissible probability level p*

only invests in the asset 2 and the risk-free asset; for $p > p^*$, a safety-first investor with admissible probability level p may invest in both assets.

Proof. We start with a portfolio R that assigns weights ω_1, ω_2 and $1-\omega_1-\omega_2$ to the three assets. The portfolio return can be written as

$$\begin{aligned} R &= \omega_1 R_1 + \omega_2 R_2 + (1-\omega_1-\omega_2) r_f \\ &= r_f + [\omega_1(\mu_1 - r_f) + \omega_2(\mu_2 - r_f)] \left[\omega \frac{R_1 - r_f}{\mu_1 - r_f} + (1-\omega) \frac{R_2 - r_f}{\mu_2 - r_f} \right] \\ &=: r_f + [\omega_1(\mu_1 - r_f) + \omega_2(\mu_2 - r_f)] [\omega X_1 + (1-\omega) X_2] \end{aligned}$$

where $\omega = \frac{\omega_1(\mu_1 - r_f)}{\omega_1(\mu_1 - r_f) + \omega_2(\mu_2 - r_f)} \in [0,1]$. Since $X_1 = \frac{R_1 - r_f}{\mu_1 - r_f}$ and $X_2 = \frac{R_2 - r_f}{\mu_2 - r_f}$ share the same expectation 1, varying the parameter ω will not change the expected return of the portfolio R. Hence, to obtain the optimal portfolio, the safety-first investor can first minimize the VaR of $X = \omega X_1 + (1-\omega) X_2$ with the optimal solution ω^*. The investor can then scale up the VaR of the portfolio by choosing $[\omega_1(\mu_1 - r_f) + \omega_2(\mu_2 - r_f)]$, such that it achieves the boundary of the VaR constraint. This gives the optimal portfolio for maximizing the safety-first utility. Hence, to prove the theorem, it is only necessary to show that there exists a p^* such that for $p < p^*$, $\omega^* = 0$, while for $p > p^*$, $\omega^* > 0$ may occur.

The second-order approximation in (2.3) on R_i leads to a similar approximation on the left tail part of the distribution function of $X_i = \frac{R_i - r_f}{\mu_i - r_f}$. The (first-order) tail index remains as α_i, while the second-order index is $\min(\gamma_i, 1)$ and the first-order and second-order scales A_i and B_i may differ. Without loss of generality, we still use the notations A_i, γ_i and B_i as the first-order scale, the second-order index and scale for X_i.

We use the second-order approximation of the VaR of $X = \omega X_1 + (1-\omega) X_2$ provided in Hyung and de Vries (2007). The approximation depends on the comparison between $\alpha_2 - \alpha_1$ and $\min\{\gamma_1, 1\}$. Here we analyze the case $\alpha_2 - \alpha_1 < \min\{\gamma_1, 1\}$. The other cases are similar, but simpler. If $1 < \alpha_1 < \alpha_2$ and $\alpha_2 - \alpha_1 < \min\{\gamma_1, 1\}$, then for a sufficiently small probability $p > 0$, the VaR of X at a probability level p is given as

$$VaR^p(\omega) \approx \begin{cases} (\frac{A_2}{p})^{1/\alpha_2}[1 + \frac{B_2}{\alpha_2}(\frac{A_2}{p})^{-\gamma_2/\alpha_2}] & \text{if } \omega = 0 \\ \omega(\frac{A_1}{p})^{1/\alpha_1}[1 + (\frac{1-\omega}{\omega})^{\alpha_2}\frac{1}{\alpha_1}\frac{A_2}{A_1^{\alpha_2/\alpha_1}}p^{\alpha_2/\alpha_1 - 1}] & \text{if } \omega \in (0,1) \\ (\frac{A_1}{p})^{1/\alpha_1}[1 + \frac{B_1}{\alpha_1}(\frac{A_1}{p})^{-\gamma_1/\alpha_1}] & \text{if } \omega = 1. \end{cases}$$

Hyung and de Vries (2007) prove that $VaR^p(\omega)$ is a continuous and convex function of ω on $(0,1)$. Hence, there exists a unique interior solution $\tilde{\omega}$ which minimizes $VaR^p(\omega)$ on the set $(0,1)$. Since $\alpha_1 < \alpha_2$, we have that $\frac{VaR^p(0)}{VaR^p(1)} \to 0$ as $p \to 0$, which implies that there exists a threshold $\hat{p}$, such that for $p < \hat{p}$, $VaR^p(0) < VaR^p(1)$. Thus, to check whether $\tilde{\omega}$ minimizes $VaR^p(\omega)$ on $\omega \in [0,1]$, we only need to compare $VaR^p(\tilde{\omega})$ with $VaR^p(0)$. We show that there exists a $p^* \leq \hat{p}$ such that for $p < p^*$, $VaR^p(\tilde{\omega}) > VaR^p(0)$, and for $p > p^*$, the relation $VaR^p(\tilde{\omega}) < VaR^p(0)$ may hold. This will complete the proof of the theorem.

We first show the second half of the statement. Note that $VaR^p(1^-) := \lim_{\omega \to 1^-} VaR^p(\omega) = (\frac{A_1}{p})^{1/\alpha_1}$. Since the parameters determining $VaR^p(0)$ and $VaR^p(1^-)$ are unrelated, with a proper combination of the parameters, $VaR^p(1^-) < VaR^p(0)$ may hold for some $\tilde{p}$, which implies by convexity that $VaR^p(0) > VaR^p(1^-) > VaR^p(\tilde{\omega})$. Since $\frac{VaR^p(0)}{VaR^p(1^-)}$ is an increasing function of p, we conclude that if $VaR^p(0) > VaR^p(\tilde{\omega})$ holds for some $\tilde{p}$, it holds for any $p > \tilde{p}$.

Next, we prove the first half of the statement by finding the threshold p^* such that for $p < p^*$, $VaR^p(\tilde{\omega}) > VaR^p(0)$. We start by studying some properties of the function $\tilde{\omega} = \tilde{\omega}(p)$. Since $\tilde{\omega}$ is the optimal solution for the minimization of $VaR^p(\omega)$ on (0,1), it satisfies that $\frac{\partial VaR^p(\omega)}{\partial \omega}|_{\omega=\tilde{\omega}} = 0$. Write

$$VaR^p(\omega) = \omega \left(\frac{A_1}{p}\right)^{1/\alpha_1} [1 + f(\omega, p)],$$

where $f(\omega, p)$ is defined as follows

$$f(\omega, p) = (\frac{1-\omega}{\omega})^{\alpha_2}\frac{1}{\alpha_1}\frac{A_2}{A_1^{\alpha_2/\alpha_1}}p^{\alpha_2/\alpha_1 - 1}.$$

Then, the first-order condition is equivalent to

$$1 + f(\tilde{\omega}, p) + \tilde{\omega} \frac{\partial f}{\partial \omega} |_{\omega=\tilde{\omega}} = 0,$$

which can be further simplified to

$$f(\tilde{\omega}, p) \left[\frac{\alpha_2}{1-\tilde{\omega}} - 1 \right] = 1. \tag{2.4}$$

We first show that as $p \to 0$, $\tilde{\omega}(p) \to 0$. This is proved by contradiction. Suppose $\limsup_{p\to 0} \tilde{\omega}(p) = c > 0$. It implies that there is a sequence $p_m \to 0$, $\tilde{\omega}(p_m) \to c$, as $m \to \infty$. From the definition of f and given $1 < \alpha_1 < \alpha_2$, we get that as $m \to \infty$, $f(\tilde{\omega}(p_m), p_m) \to 0$ and $\frac{1}{1-\tilde{\omega}(p_m)} f(\tilde{\omega}(p_m), p_m) \to 0$. This is in contradiction with (2.4). Hence, as $p \to 0$, necessarily $\tilde{\omega}(p) \to 0$.

Next, we compare $VaR^p(\tilde{\omega})$ with $VaR^p(0)$ as

$$\begin{aligned}
\Theta(p) := \frac{VaR^p(\tilde{\omega})}{VaR^p(0)} &= \tilde{\omega} \frac{(A_1/p)^{1/\alpha_1}}{(A_2/p)^{1/\alpha_2}} \frac{1 + f(\tilde{\omega}, p)}{1 + \frac{B_2}{\alpha_2} (\frac{A_2}{p})^{-\gamma_2/\alpha_2}}, \\
&= \frac{1-\tilde{\omega}}{\alpha_1^{1/\alpha_2}} \left(\alpha_1^{1/\alpha_2} \frac{\tilde{\omega}}{1-\tilde{\omega}} \frac{(A_1/p)^{1/\alpha_1}}{(A_2/p)^{1/\alpha_2}} \right) \frac{1 + f(\tilde{\omega}, p)}{1 + \frac{B_2}{\alpha_2} (\frac{A_2}{p})^{-\gamma_2/\alpha_2}}, \\
&= \frac{1-\tilde{\omega}}{\alpha_1^{1/\alpha_2}} (f(\tilde{\omega}, p))^{-1/\alpha_2} \frac{1 + f(\tilde{\omega}, p)}{1 + \frac{B_2}{\alpha_2} (\frac{A_2}{p})^{-\gamma_2/\alpha_2}}.
\end{aligned}$$

Together with equation (2.4), we get that

$$\Theta(p) = \frac{\alpha_2}{\alpha_1^{1/\alpha_2}} \left(\frac{\alpha_2}{1-\tilde{\omega}} - 1 \right)^{1/\alpha_2 - 1} \frac{1}{1 + \frac{B_2}{\alpha_2} (\frac{A_2}{p})^{-\gamma_2/\alpha_2}}.$$

By taking $p \to 0$, together with $\tilde{\omega}(p) \to 0$, we get that

$$\lim_{p\to 0} \Theta(p) = \frac{\alpha_2}{(\alpha_2 - 1)^{1-1/\alpha_2}} \frac{1}{\alpha_1^{1/\alpha_2}} = \left(\frac{\alpha_2 - 1}{\alpha_1} \right)^{1/\alpha_2 - 1} \frac{\alpha_2}{\alpha_1} > 1.$$

Therefore, there exists a p^* such that for $p < p^*$, $\Theta > 1$, which implies that $\omega^* = 0$. This corresponds to the conclusion that the optimal portfolio for a safety-first consists of the risky asset 2 and the risk-free rate only. ■

Theorem 2.2 shows that there is threshold p^* such that if safety-first investors consider an admissible probability $p < p^*$, they only invest on the risky asset with the higher tail index. Conversely, for $p > p^*$, the optimal portfolio may consist of both risky assets. By extending Theorem 2.2 to a multiple-asset market, with the same reasoning as in Subsection 2.2, we conclude that with a sufficiently low admissible probability level, all assets traded in the market must share the same tail index, i.e. the "tail index equivalence hypothesis" remains valid. However, if the admissible probability level is at a more moderate level, the "tail index equivalence hypothesis" may not apply. In such a case, the investment universe may comprise assets with different tail indices. This would allow for different asset classes such as stocks and bonds.

We compare the two theorems obtained under the first-order and second-order tail approximation. The general conclusions are similar. However, the main difference is that the theorem under the second-order approximation is proved for a range of low p rather than requiring $p \to 0$. This leads to different economic interpretations. The result in Theorem 2.1 shows that under the first-order approximation, any portfolio based on two risky assets is dominated, in the case of safety-first utility, by another portfolio consisting of the asset with the higher tail index and the risk-free rate only for sufficiently low p. Note that the admissible probability level p ensures such a statement varies according to the initial portfolio based on the two risky assets. Hence, based on this theorem, we conclude the "tail index equivalence hypothesis" only by taking $p \to 0$. Per contrast, in Theorem 2.2, we prove that by first fixing a low admissible probability level p as $p < p^*$, any portfolio based on the two risky assets will be dominated by the "corner solution". In other words, the "corner solution" dominates other portfolios uniformly as opposed to the point-wise dominance in Theorem 2.1. Therefore, although we also require the admissible probability level to be low, we do not have to consider the limit case in order to establish the "tail index equivalence" hypothesis. The second theorem is thus more appealing in characterizing the real situation in the market.

3 Heterogeneity in the scale parameter

In Section 2, we show that if safety-first investors have a low admissible probability, they only invest in assets that share the highest value of tail index in their optimal portfolio. In other words, the "tail index equivalence" hypothesis holds: assets traded in the same market must share the same tail index. Suppose the tail index equivalence hypothesis holds. The next step in studying the cross-section of downside tail risk then regards the scale parameters. Is there a "scale equivalence hypothesis" as well? In this section, we show that even if the safety-first investors have a low admissible probability, it is still possible to have heterogeneity in the scales. Such heterogeneity can be compensated by differences in expected returns. We further derive the equilibrium prices of asset returns with heterogeneous scale parameters. The result is in accordance with that in Arzac and Bawa (1977).

Suppose all assets share the same tail index, any portfolio constructed from these assets will then have the same tail index, regardless the dependence structure among the asset returns, see Zhou (2010). We first show that when safety-first investors construct their optimal portfolios the relative proportion of risky assets is invariant with respect to investors' wealth and downside risk concerns. The relative proportion only depends on the characteristics of asset returns, more specifically, the scales and the dependence structure among them. This is the first step to derive the equilibrium prices as in Arzac and Bawa (1977).

Consider an investor who constructs a weighted portfolio from N assets. Let R_i denote the return of the asset i with expected return μ_i. Suppose R_i follows a heavy-tailed distribution as in (1.1) with tail index α and scale A_i, $i = 1, ..., N$. As discussed in Subsection 2.1, a safety-first investor seeks to maximize the expected portfolio returns with a downside risk constraint. Given an admissible probability of failure p, the optimization problem can be stated as

$$\max_{\{\omega_i\}} \sum_{i=1}^{N} \omega_i \mu_i + (1 - \sum_{i=1}^{N} \omega_i) r_f$$

s.t.

$$VaR^p(P) \leq T,$$

$$\omega_i \geq 0, \text{ for } i = 1, ..., N,$$

where $VaR^p(P)$ is the VaR of the portfolio $P = \sum_{i=1}^{N} \omega_i R_i + (1 - \sum_{i=1}^{N} \omega_i) r_f$. To calculate the VaR, we apply the theory on aggregating heavy-tailed distributions with the same tail index.

Given that all assets share the same tail index, Zhou (2010) shows the portfolio return P follows a heavy-tailed distribution with tail index α and scale $A_P =: A(\omega_1, ..., \omega_N)$, where $A(\omega_1, ..., \omega_N)$ is a function of the weights $\omega_i \geq 0$, $i = 1, ..., N$, the scales of asset returns and the dependence structure among the asset returns. Moreover, it is a homogeneous function with degree α, i.e. for any constant c,

$$A(c\omega_1, ..., c\omega_N) = c^\alpha A(\omega_1, ..., \omega_N).$$

Then the VaR of the portfolio P can be approximated as

$$VaR^p(P) = \left(\frac{A_P}{p}\right)^{\frac{1}{\alpha}} = \left(\frac{A(\omega_1, ..., \omega_N)}{p}\right)^{\frac{1}{\alpha}}.$$

We solve the optimal portfolio problem in two steps as in Subsection 2.3. First, we solve a scale minimization problem for portfolios with a fixed expected return. We show that the solution exists and is unique. Then we prove that the optimal portfolio for a safety-first investor can be obtained by scaling up the weights of the optimal portfolio solved from the scale minimization problem.

Without loss of generality we consider portfolios with a fixed expected return C, i.e.

$\sum_{i=1}^{N} \omega_i \mu_i + (1 - \sum_{i=1}^{N} \omega_i) r_f = C$, and try to solve the following optimization problem:

$$\min_{\omega_i} A(\omega_1, ..., \omega_N)$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \omega_i \mu_i + (1 - \sum_{i=1}^{N} \omega_i) r_f = C.$$

Zhou (2010) shows that the function $A(\omega_1, \cdots, \omega_N)$ is strictly convex in $[0,1]^N$, $i = 1, ..., N$. By choosing a low level of C, for example $C < \min_{1 \le i \le N} \{\mu_i - r_f\}$, we get that the area $\sum_{i=1}^{N} \omega_i \mu_i + (1 - \sum_{i=1}^{N} \omega_i) r_f = C$ is a convex subset of $[0,1]^N$. From the convex optimization theory, there exists a unique interior solution, $(\omega_1^*, \cdots, \omega_N^*)$, which solves the scale minimization problem. More specifically, the solution $(\omega_1^*, \cdots, \omega_N^*)$ satisfies the first-order conditions as follows:

$$\frac{\partial}{\partial \omega_i} \left(A(\omega_1, ..., \omega_N) - \lambda \left(\sum_{i=1}^{N} \omega_i \mu_i + (1 - \sum_{i=1}^{N} \omega_i) r_f - C \right) \right) = 0,$$

for $i = 1, 2, \cdots, N$, where λ is the Lagrange multiplier. Denote the partial derivatives as $a_i(\omega_1, ..., \omega_N) = \frac{\partial}{\partial \omega_i} A(\omega_1, ..., \omega_N)$, for $i = 1, 2, \cdots, N$. The first-order condition, for any $1 \le i \le N$, is then written as

$$a_i(\omega_1^*, \cdots, \omega_N^*) = \lambda(\mu_i - r_f). \tag{3.1}$$

It implies that for any i and j,

$$\frac{a_i(\omega_1^*, \cdots, \omega_N^*)}{a_j(\omega_1^*, \cdots, \omega_N^*)} = \frac{\mu_i - r_f}{\mu_j - r_f}, \tag{3.2}$$

which is independent of λ. Because the function $A(\omega_1, \ldots, \omega_d)$ is homogeneous of degree α, its partial derivatives, a_i, are thus homogeneous functions with degree $\alpha - 1$. Hence the relation (3.2) determines the relative proportion among ω_i^*.

Next, we go back to the original optimization problem for safety-first investors. To construct the optimal portfolio, a safety-first investor will assign weights $(\tilde{\omega}_1, \cdots, \tilde{\omega}_N)$

to the risky assets, while borrowing or lending $1 - \sum_{i=1}^{N} \tilde{\omega}_i$ in the risk-free asset. Then $(\tilde{\omega}_1, \cdots, \tilde{\omega}_N)$ must satisfy the following first-order conditions:

$$\frac{\partial}{\partial \omega_i} \left(\sum_{i=1}^{N} \omega_i \mu_i + (1 - \sum_{i=1}^{N} \omega_i) r_f - \eta \left(\left(\frac{A(\omega_1, ..., \omega_N)}{p} \right)^{1/\alpha} - T \right) \right) = 0,$$

for $i = 1, \cdots, N$, where η is the Lagrange multiplier. Together with the VaR constraint, we simplify the first-order conditions to get that

$$\frac{a_i(\tilde{\omega}_1, \cdots, \tilde{\omega}_N)}{A(\tilde{\omega}_1, \cdots, \tilde{\omega}_N)} = \frac{\alpha}{\eta T} \cdot (\mu_i - rf).$$

Hence, for any i and j, $\frac{a_i(\tilde{\omega}_1, \cdots, \tilde{\omega}_N)}{a_j(\tilde{\omega}_1, \cdots, \tilde{\omega}_N)} = \frac{\mu_i - r_f}{\mu_j - r_f}$, which is independent of η and T. Because these conditions are identical to (3.2), we conclude that the relative proportion among ω_i^* is equivalent to that of $\tilde{\omega}_i$, i.e.

$$(\tilde{\omega}_1, \cdots, \tilde{\omega}_N) = l(\omega_1^*, \cdots, \omega_N^*).$$

The constant l is determined by the VaR constraint as $l = \left(\frac{A(\omega_1^*, \ldots, \omega_N^*)}{p} \right)^{-1/\alpha} T$.

To summarize, we have shown that the optimal portfolio for a safety-first investor can be obtained by scaling up the weights of the optimal portfolio solved from a scale minimization problem. Notice that $(\omega_1^*, \cdots, \omega_N^*)$ only depends on the characteristics of the assets. Hence, the relative proportion of assets held in any portfolio of safety-first investor are homogeneous, even if the admissible probabilities and VaR constraints are heterogeneous. In other words, denote $R_m = \frac{\sum_{i=1}^{N} \omega_i^* R_i}{\sum_{i=1}^{N} \omega_i^*}$ as a market portfolio. Any optimal portfolio for a safety-first investor is obtained by scaling the market portfolio with proper borrowing and lending. This is in accordance with the asset pricing theory in Arzac and Bawa (1977).

Next, we derive the equilibrium price for each risky asset R_i. Firstly, for the market portfolio, we have that

$$E(R_m) = \frac{\sum_{i=1}^{N} \omega_i^* \mu_i}{\sum_{i=1}^{N} \omega_i^*} = \frac{C - (1 - \sum_{i=1}^{N} \omega_i^*) r_f}{\sum_{i=1}^{N} \omega_i^*} = \frac{C - r_f}{\sum_{i=1}^{N} \omega_i^*} + r_f.$$

For each individual asset, the expected return is derived from (3.1). Because the function A is homogeneous with degree α, we have that

$$\sum_{i=1}^{N} \omega_i a_i = \alpha A.$$

Combining with the first-order conditions in (3.1), we get that

$$\lambda \sum_{i=1}^{N} \omega_i^*(\mu_i - r_f) = \alpha A(\omega_1^*, \cdots, \omega_N^*),$$

where $(\omega_1^*, \cdots, \omega_N^*)$ is the solution of the scale minimization problem. Together with the expected return constraint, we get that $\lambda = \frac{\alpha A(\omega_1^*, \cdots, \omega_N^*)}{C - r_f}$. Together with (3.1), we relate the expected returns of each asset with the optimal portfolio weights as

$$a_i(\omega_1^*, \cdots, \omega_N^*) = \alpha A(\omega_1^*, \cdots, \omega_N^*) \frac{\mu_i - r_f}{C - r_f}.$$

This gives the equilibrium price of R_i as

$$\mu_i - r_f = \frac{a_i(\omega_1^*, \cdots, \omega_N^*)}{\alpha A(\omega_1^*, \cdots, \omega_N^*)}(C - r_f) = \beta_i^*(E(R_m) - r_f).$$

where $\beta_i^* = \frac{a_i}{\alpha A}|_{\omega_i = \omega_i^*} \cdot \sum_{i=1}^{n} \omega_i^*$.

We show that the equilibrium price is in accordance with the result in Arzac and Bawa (1977). Under the safety-first utility, Arzac and Bawa (1977) show that

$$E(R_i) = r_f + \beta_i(E(R_m) - r_f),$$

where $\beta_i = \frac{\frac{\partial VaR^p(R_m)}{\partial \omega_i}|_{\omega_i = \omega_{m,i}}}{VaR^p(R_m)}$, $i = 1, \cdots, N$. and $\omega_{m,i}$ are the weights for the market portfolio. In our case $\omega_{m,i} = \frac{\omega_i^*}{\sum_{i=1}^{N} \omega_i^*}$.

We prove that, $\beta_i^* = \beta_i$. From the VaR formula, we get that $\log VaR^p(R) = \frac{1}{\alpha}(\log A -$

$\log p$). By taking partial derivative on both sides and let $\omega_i = \frac{\omega_i^*}{\sum_{i=1}^N \omega_i^*}$, we get that

$$\beta_i = \frac{\frac{\partial VaR^p}{\partial \omega_i}|_{\omega_i=\omega_{m,i}}}{VaR^p(R_m)} = \frac{a_i}{\alpha A}|_{\omega_i=\omega_{m,i}}.$$

Because the functions A and a_i are both homogeneous function, but with different degree α and $\alpha - 1$, we finally get that

$$\beta_i = \frac{a_i}{\alpha A}|_{\omega_i=\omega_{m,i}} = \frac{a_i}{\alpha A}|_{\omega_i=\omega_i^*} \frac{\left(\sum_{i=1}^N \omega_i^*\right)^{-(\alpha-1)}}{\left(\sum_{i=1}^N \omega_i^*\right)^{-\alpha}} = \beta_i^*.$$

To conclude, we have shown that under the tail index equivalence hypothesis, scale heterogeneity is possible across assets traded in the same market. The heterogeneity in scales is priced by the expected return. Although expected returns may not compensate for the differences in the tail index, these can be compensated for differences in the scales.

4 Conclusion

This paper addresses the question of how the downside tail risk of stock returns are differentiated cross-sectionally. For stock returns with a heavy-tailed distribution, the downside tail risk is determined by two parameters: the tail index and scale. We provide a theoretical model to show that if safety-first investors consider sufficiently large downside losses, then the distributions of asset returns share a homogeneous shape parameter. We call this the "tail index equivalence" hypothesis. When the tail index equivalence hypothesis holds, the equilibrium price of assets is related to the cross-section of tail risks indicated by the scales. In other words, there is no "scale equivalence hypothesis". Conversely, we show that tail index equivalence hypothesis may fail, if investors consider moderate downside losses only.

A direct consequence of our results is on portfolio diversification with heavy-tailed assets. Once the tail index equivalence hypothesis holds, investors are able to diversify their downside tail risk based on the scale of the distributions of stock returns. Compared with the asset pricing theory based on the mean-variance utility, our theory gives a similar result: investors optimally choose a diversified portfolio to minimize the downside tail risk measured by the portfolio scale. On the other hand, when the tail index equivalence hypothesis fails, investors are not able to diversify away the tail risks in assets with heavier tails. This differs from the classic asset pricing theory under the Gaussian framework.

Based on our theoretical result, there remains an empirical exercise to test the tail index equivalence hypothesis. In addition, when the tail index equivalence does hold, one may test whether the scale equivalence hypothesis fails as predicted by the theory. Further, if the scales are heterogeneous, what are the firm-level determinants that can differentiate such heterogeneity? These questions are left for future research.

References

A. Ang, R.J. Hodrick, Y. Xing, and X. Zhang. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.

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

L. De Haan and A. Ferreira. *Extreme value theory: an introduction*. Springer Verlag, 2006.

P. Embrechts, C. Klüppelberg, and T. Mikosch. *Modelling extremal events for insurance and finance*, volume 33. Springer Verlag, 1997.

E.F. Fama and M.H. Miller. *The Theory of Finance*. Holt, Rinehart and Winston, New York, 1972.

C.R. Harvey and A. Siddique. Conditional skewness in asset pricing tests. *The Journal of Finance*, 55(3):1263–1295, 2000.

N. Hyung and C.G. de Vries. Portfolio diversification effects and regular variation in financial data. *Allgemeines Statistiches Archiv*, 86:69–82, 2002.

N. Hyung and C.G. de Vries. Portfolio diversification effects of downside risk. *Journal of Financial Econometrics*, 3:107–125, 2005.

N. Hyung and C.G. de Vries. Portfolio selection with heavy tails. *Journal of Empirical Finance*, 14:383–400, 2007.

N. Hyung and C.G. de Vries. Simulating and calibrating diversification against black swans. *Journal of Economic Dynamics and Control*, 2012.

R. Ibragimov and J. Walden. Portfolio diversification under local and moderate deviations from power laws. *Insurance: Mathematics and Economics*, 42:594–599, 2008.

D.W. Jansen and C.G. De Vries. On the frequency of large stock returns: Putting booms and busts into perspective. *The Review of Economics and Statistics*, pages 18–24, 1991.

E. Jondeau and M. Rockinger. Testing for differences in the tails of stock-market returns. *Journal of Empirical Finance*, 10(5):559–581, 2003.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, pages 263–291, 1979.

M. Loretan and P.C.B. Phillips. Testing the covariance stationarity of heavy-tailed time series: An overview of the theory with applications to several financial datasets. *Journal of empirical finance*, 1(2):211–248, 1994.

B. Mandelbrot. The variation of certain speculative prices. *The journal of business*, 36 (4):394–419, 1963.

H. Markowitz. Portfolio selection: efficient diversification of investments. *New Haven, CT: Cowles Foundation*, 94, 1959.

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

C. Zhou. Dependence structure of risk factors and diversification effects. *Insurance: Mathematics and Economics*, 46(3):531–540, 2010.