JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS Vol. 51, No. 2, Apr. 2016, pp. 685–705

doi:10.1017/S0022109016000193

Systematic Tail Risk

Maarten R. C. van Oordt and Chen Zhou*

Abstract

We test for the presence of a systematic tail risk premium in the cross section of expected returns by applying a measure of the sensitivity of assets to extreme market downturns, the tail beta. Empirically, historical tail betas help predict the future performance of stocks in extreme market downturns. During a market crash, stocks with historically high tail betas suffer losses that are approximately 2 to 3 times larger than their low-tail-beta counterparts. However, we find no evidence of a premium associated with tail betas. The theoretically additive and empirically persistent tail betas can help assess portfolio tail risks.

I. Introduction

Risk managers are concerned with the performance of portfolios in distress events, the so-called tail events in the market. In this article, we investigate the sensitivity of assets to market risk under extremely adverse market conditions (i.e., their loading on systematic tail risk). We estimate an additive measure of sensitivity to systematic tail risk, the "tail beta." We examine whether the estimated loadings on systematic tail risk help explain the cross section of expected returns and discuss their potential application in risk management.

Systematic tail risk may play an important role in asset pricing. One potential reason is that investors may follow a safety-first principle. Arzac and Bawa (AB) (1977) derive an asset pricing theory under the safety-first principle of Telser (1955). They consider investors who maximize their expected return subject to a value-at-risk (VaR) constraint. In their framework, the cross section of expected

*Van Oordt (corresponding author), mvanoordt@bankofcanada.ca, Government of Canada, Bank of Canada, Ottawa, ON K1A 0G9, Canada; and Zhou, c.zhou@dnb.nl, zhou@ese.eur.nl, Economics and Research Division, De Nederlandsche Bank, 1000AB Amsterdam, The Netherlands. The authors thank participants in various conferences and seminars where previous versions of this paper have been presented: the 2012 Swiss Society for Financial Market Research Conference (Zürich), the 2012 International Risk Management Conference (Rome), the 2012 Annual Cambridge Finance–Wharton–Tinbergen Conference (Cambridge), the 2012 Annual Congress of the European Economic Association (Málaga), the 2012 Netherlands Economists Day (Amsterdam), the 2014 Annual Meeting of the Midwest Finance Association (Orlando), and seminars in De Nederlandsche Bank. Especially, the authors want to thank Gerry Abdesaken (discussant), Stephen Brown (the editor), Casper de Vries, Francis Diebold (discussant), Winfried Hallerbach, Mohammad Irani (discussant), and Haim Levy (the referee) for useful comments and suggestions. This research was conducted while the corresponding author was a researcher in the Economics and Research Division of De Nederlandsche Bank. Views expressed are those of the authors and do not necessarily reflect official positions of Bank of Canada or De Nederlandsche Bank.

returns is explained by a "beta" that is different from the regular market beta in the capital asset pricing model (CAPM). If investors are constrained by a VaR with a sufficiently small probability, then our tail beta equals the beta of AB, assuming a linear model under extremely adverse market conditions.

Our empirical results provide evidence that historical tail betas do capture future systematic tail risk: They help predict which stocks take relatively large hits during future market crashes. The evidence is twofold. First, the persistencies of the classification of firms based on tail betas and regular market betas are comparable, even though tail betas are estimated from only a few tail observations. Second, stocks with historically high tail betas suffer losses during future market crashes that are on average 2 to 3 times larger than their low-tail-beta counterparts.

Furthermore, we test whether the estimated tail betas help explain the cross section of expected returns. That is, we test whether stocks with high tail betas are compensated by a risk premium. Surprisingly, from the asset pricing tests we do not observe such a premium for stocks with high tail betas. This finding is not a consequence of losses suffered during the recent financial crisis. The risk premium remains absent if this episode is excluded from our sample. Hence, the role of systematic tail risk in explaining the cross section of expected returns seems to be limited.

These results are not explained by many other factors documented in the asset pricing literature and are robust to methodological changes. The results are established within all size cohorts (Fama and French (2008)) and in the context of both equal- and value-weighted portfolios. They are robust when controlling for downside beta, coskewness, cokurtosis, idiosyncratic risk (Ang, Hodrick, Xing, and Zhang (2006)), and volume (Gervais, Kaniel, and Mingelgrin (2001)). The results are not explained by return characteristics related to short-term reversal (Jegadeesh (1990)), momentum (Carhart (1997)), and long-term reversal (De Bondt and Thaler (1985)).

We focus on tail betas because regular market betas do not necessarily provide an accurate description of the loading on systematic risk under all market conditions. It is a well-known stylized fact that equity returns show higher correlations during periods of high stock market volatility (see, e.g., King and Wadhwani (1990), Longin and Solnik (1995), Karolyi and Stulz (1996), and Ramchand and Susmel (1998)). In addition, correlations increase especially during periods of severe market downturns as reported by Longin and Solnik (2001), Ang and Bekaert (2002), Ang and Chen (2002), and Patton (2004). The changing correlation signals that the comovement of assets with the market depends on market conditions.

Several studies address the changing comovement in the context of a nonlinear relation between asset returns and market risk. In line with the theoretical work of Rubinstein (1973) and Kraus and Litzenberger (1976), one strand of literature estimates the relation between asset returns and market risk with higher order approximations. Among others, Harvey and Siddique (2000) and Dittmar (2002) find that "coskewness" and "cokurtosis" play a role in asset pricing. However, room for extensions with additional higher moments may be limited because the heavy tails observed in stock returns provide evidence that further higher moments may not exist (see Mandelbrot (1963), Jansen and De Vries (1991)).

https://doi.org/10.1017/S0022109016000193 Published online by Cambridge University Press

An alternative strand of literature focuses on the comovement of asset returns with the market under specific market conditions. In line with the theory in Bawa and Lindenberg (1977), several studies investigate the downside beta, which is defined as the market beta conditional on below-average or below-zero market returns (see, e.g., Price, Price, and Nantell (1982), Harlow and Rao (1989)). Our tail beta fits within this latter strand of literature because it focuses on comovement with the market return under specific market conditions. However, in contrast to downside beta, the tail beta measures comovement with the market if market downturns are extreme. The lack of a risk premium for tail beta in our results contrasts with the positive risk premium observed for downside beta (see, e.g., Ang, Chen, and Xing (2006)). This difference suggests that investors are concerned with their exposure to all systematic downside risk rather than focusing only on their loading on systematic tail risk.

Our study is related to the empirical asset pricing literature on tail risks. A few studies focus on the role of tail risk in the cross section of expected returns, irrespective of its relation with market risk (see, e.g., Bali, Demirtas, and Levy (2009), Huang, Liu, Rhee, and Wu (2012), and Cholette and Lu (2011)). Alternatively, Kelly and Jiang (2014) construct an index on the level of tail risks in the market and obtain "tail risk betas" for individual assets by regressing asset returns on innovations in this index. These betas can be considered a tail equivalent of the volatility betas of Ang, Hodrick, Xing, and Zhang (2006), which measure the comovement of stock returns with innovations in market volatility. In contrast to the tail risk beta of Kelly and Jiang, our tail beta can be considered a tail equivalent of the market beta.

The expression tail beta appears in the literature with other meanings. For example, De Jonghe (2010) estimates tail betas by applying a tail dependence measure from Poon, Rockinger, and Tawn (2004) on stock returns. Spitzer (2006) and Chabi-Yo, Ruenzi, and Weigert (2015) examine its asset pricing power. This tail dependence measure is defined as the probability of an extreme downward movement of the asset, conditional on the occurrence of a market crash. Hence, instead of measuring the magnitude of the comovement, this measure has the attributes of a conditional probability. Furthermore, Bali, Cakici, and Whitelaw (2014) estimate "hybrid tail betas." Their aim is to capture the covariance of the asset and the market, given an adverse return on the asset.

Compared to these measures, the tail beta we estimate shares two appealing features with the regular market beta. First, its interpretation as a measure of comovement with the market is in absolute terms. That is, on a day that the market suffers a loss of 10%, an asset with a tail beta of 2 is expected to suffer a downward movement of 20%. Second, the tail beta we estimate is an additive measure of tail risk. In other words, the tail beta of an investment portfolio is the weighted average of the tail betas of the individual assets. Consequently, the estimated tail betas provide a clear insight into how each asset contributes to the systematic tail risk of a portfolio, which is discussed in Section V.

II. Theory

To define the tail beta, we first introduce a linear model that decomposes asset returns under extremely adverse market conditions into a systematic and an idiosyncratic component. We denote the return on asset j and the market portfolio as R_j and R_m. The excess return on asset j and the market are given by $R_j^e = R_j - R_f$ and $R_m^e = R_m - R_f$, where R_f is the risk-free rate. The following model relates the two excess returns under extremely adverse market conditions

$$(1) \qquad R_j^e \quad = \quad \beta_j^T R_m^e + \varepsilon_j, \quad \text{for} \quad R_m^e \quad < \quad -\text{VaR}_m(\bar{p}),$$

where ε_j denotes the idiosyncratic risk that is independent of R_m^e under the condition $R_m^e < -\text{VaR}_m(\bar{p})$ and $E\varepsilon_j = 0$. The $\text{VaR}_m(\bar{p})$ denotes the VaR of the excess market return with some low probability $\bar{p}$ such that $\Pr(R_m^e \leq -\text{VaR}_m(\bar{p})) = \bar{p}$; in other words, it is the loss on the market that is exceeded with probability $\bar{p}$. The tail beta is defined as parameter β_j^T and measures the sensitivity to systematic tail risk.

The linear tail model in expression (1) specifies the comovement between the asset and the market excess return only under extremely adverse market conditions. Nevertheless, safety-first investors do not need any further assumptions to value each asset according to the asset pricing theory developed by AB (1977): Given the linear tail model in expression (1), we show that the tail beta determines expected returns in their framework.

The asset pricing theory of AB (1977) builds on the assumption that investors maximize the expected return while limiting the probability of suffering a particularly large loss below a predetermined admissible level p.[1] In other words, investors maximize the expected return under a VaR constraint. Under this objective function, AB prove in a distribution-free setup that the equilibrium price for any asset j is given by

$$(2) \qquad E(R_j^e) \quad = \quad \beta_j^{\text{AB}} E(R_m^e),$$

where the β_j^{AB} is determined by the asset's contribution to the VaR of the market portfolio with a probability level equal to the admissible probability p (see the Appendix), that is,

$$(3) \qquad \beta_j^{\text{AB}} \quad = \quad \frac{E(R_j^e | R_m^e = -\text{VaR}_m(p))}{-\text{VaR}_m(p)}.$$

Given the linear tail model in expression (1), suppose that the investors care about sufficiently large losses, such that the admissible probability p is smaller than $\bar{p}$. We can then express the β_j^{AB} in equation (3) as

$$\beta_j^{\text{AB}} \quad = \quad \frac{E(\beta_j^T R_m^e + \varepsilon_j | R_m^e = -\text{VaR}_m(p))}{-\text{VaR}_m(p)} \quad = \quad \beta_j^T + \frac{E(\varepsilon_j)}{-\text{VaR}_m(p)} \quad = \quad \beta_j^T.$$

Hence, we establish that the tail beta, β_j^T, equals the beta in the AB (1977) asset pricing theory, β_j^{AB}. Consequently, given the linear tail model in expression (1) and the safety-first framework with $p < \bar{p}$, the expected returns of assets depend on their tail betas.

In summary, if the market is populated with safety-first investors who care about extreme losses that occur with a sufficiently low probability, and if tail

[1]The initial safety-first principle introduced by Roy (1952) assumes that agents minimize the probability of suffering a large loss. AB (1977) adapt the formulation by Telser (1955), which assumes that agents intend to limit the probability of suffering a particularly large loss to a prespecified level.

betas can capture future systematic tail risk, then assets with higher tail betas should be compensated with a risk premium in the cross section of expected returns.

III. Methodology

Our objective is to test whether assets with relatively high tail betas perform worse in future market crashes and whether they earn a systematic tail risk premium. For that purpose, we collect daily data on New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks of nonfinancials between July 1963 and Dec. 2010 from the Center for Research in Security Prices (CRSP). In addition, we collect the risk-free rates and the excess returns on the market portfolio from the data library section on Kenneth French's Web site (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/).

We estimate firm-specific tail betas at the start of each month between July 1968 and Dec. 2010. A potential difficulty with the estimation of tail betas is the low number of observations that correspond to extremely adverse market conditions. Researchers often estimate market betas based on the past 60 monthly returns. Such a low number of observations is insufficient for our purpose of estimating tail betas. We therefore use daily returns from the past 60 months in our estimates, which corresponds to approximately 1,260 days.

We estimate tail betas using the estimation methodology based on extreme value theory (EVT) developed by Van Oordt and Zhou (2016). The basic assumption of this approach is that the market and asset returns are heavy-tailed with the following expansion on the tail of their distribution functions:

$$\text{(4)} \qquad \begin{aligned} &\Pr(R_m^e < -u) \sim A_m u^{-\alpha_m} \quad \text{and} \\ &\Pr(R_j^e < -u) \sim A_j u^{-\alpha_j}, \quad \text{as} \quad u \to \infty. \end{aligned}$$

The parameters α_m and α_j are called the *tail indices*, and the parameters A_m and A_j are the *scales*. With the independence between the idiosyncratic risk, ε_j, and the market risk, R_m^e, the linear tail model in (1) induces a dependence structure between extremely adverse market returns and the asset returns. The tail beta is estimated by exploiting the tail dependence structure and using the observations in the tail region only. With the number of observed returns denoted by n, only the k lowest returns are used in the estimation.[2] The estimator of the tail beta is given as

$$\text{(5)} \qquad \hat{\beta}_j^T \quad := \quad \widehat{\tau_j(k/n)}^{1/\hat{\alpha}_m} \frac{\widehat{\text{VaR}_j(k/n)}}{\widehat{\text{VaR}_m(k/n)}},$$

with four components obtained as follows. First, the tail index α_m can be estimated with the so-called *Hill estimator* (see Hill (1975)).[3] Consider the losses

[2]Theoretically, the EVT approach requires $k := k(n)$ to be an intermediate sequence such that $k \to \infty$ and $k/n \to 0$ as $n \to +\infty$. In practice, these conditions on k are not relevant for a finite sample size n. For low values of k, the estimate exhibits a large variance, while for high values of k, it bears a potential bias because observations from relatively normal market conditions are included in the estimation. Practically, we choose $k = 50$ days in each estimation window of 60 months, which corresponds to a k/n-ratio of roughly 4%. The results are robust if tail betas are estimated with $k = 30$.

[3]The EVT approach needs the weak condition that $\alpha_j > \alpha_m/2$. This condition requires a lower bound on the tail index of asset excess returns. Empirical research usually finds that α_m is around 4

$X_t^{(m)} = -R_{m,t}^e$, for $t = 1, \ldots, n$. By ranking them as $X_{n,1}^{(m)} \leq X_{n,2}^{(m)} \leq \ldots \leq X_{n,n}^{(m)}$, the Hill estimator is calculated as

$$(6) \qquad \frac{1}{\hat{\alpha}_m} = \frac{1}{k}\sum_{i=1}^{k} \ln X_{n,n-i+1}^{(m)} - \ln X_{n,n-k}^{(m)}.$$

Second, the $\tau_j(k/n)$ parameter can be estimated nonparametrically as

$$(7) \qquad \widehat{\tau_j(k/n)} := \frac{1}{k}\sum_{t=1}^{n} \mathbf{1}_{\left\{X_t^{(j)} > X_{n,n-k}^{(j)} \text{ and } X_t^{(m)} > X_{n,n-k}^{(m)}\right\}},$$

where $X_{n,n-k}^{(j)}$ is the $(k+1)$th highest loss on the asset, and where $X_t^{(j)} := -R_{j,t}^e$, $t = 1, \ldots, n$ (see Embrechts, De Haan, and Huang (2000)). This parameter characterizes the tail dependence between the market and the asset. Finally, the VaRs of the market and asset return at probability level k/n are estimated by their $(k+1)$th highest losses.[4]

At the end of every estimation window we rank the firms based on their tail betas and construct five portfolios, each of which contains the same number of stocks. To maximize the potential variation after controlling for regular market risk, we also sort stocks based on their tail beta spreads, that is, the spread between tail betas and regular market betas.[5] In the portfolio formation procedure we exclude firms that do not qualify according to the following two conditions. First, stocks should not report zero returns on more than 60% of the trading days in the estimation window. We use this criterion to avoid our results being distorted by daily returns of thinly traded stocks. Second, the stock must be trading at a price above US$5 on the last day of the estimation period. We use this criterion to exclude penny stocks that potentially represent firms in severe financial distress.[6] In summary, portfolio formation occurs at the start of each month using estimates based on daily returns from the past 60 months. The holding period is the first month after the estimation window.

After constructing the portfolios, we calculate daily portfolio returns. The excess return is calculated by averaging the excess returns on individual stocks in each portfolio using both equal and value weights. Furthermore, using several

(see, e.g., Jansen and De Vries (1991), Loretan and Phillips (1994), and Poon et al. (2004)). In line with these results, we observe $\hat{\alpha}_m = 3.5$ as an average estimate for the market. Given these findings, the condition is equivalent to $\alpha_j > 2$, which is satisfied if the excess returns of individual assets have finite variance.

[4]An alternative approach to estimate tail betas involves performing a regression on the observations corresponding to the k largest market losses. For example, Post and Versijp (2007) provide estimates of tail betas from regressions conditional on market returns below -10%. Our results are robust to using the conditional regression approach. However, this approach yields a less persistent ranking of firms over time and a smaller and less significant return difference between high- and low-tail-beta stocks in extreme market downturns.

[5]See, for example, Ang, Chen, and Xing (2006), who sort based on the spread between downside beta and market beta.

[6]Empirically, historical tail betas do not provide much information about the performance of penny stocks in future extreme market downturns. Including penny stocks in the analysis does not qualitatively affect our conclusions.

https://doi.org/10.1017/S0022109016000193 Published online by Cambridge University Press

benchmark models, we calculate risk-adjusted returns for individual stocks as follows:

$$R^*_{j,t} \quad = \quad R_{j,t} - R_{f,t} - \sum_{l=1}^{s} \hat{\beta}_{j,l} F_{l,t}, \tag{8}$$

where $R_{j,t}$ is the daily return on stock j at time t, $R_{f,t}$ is the risk-free rate, and $(F_{1,t}, \ldots, F_{s,t})$ denotes the s risk factors in the benchmark model. We estimate the factor loadings, $\hat{\beta}_{j,k}$, for individual stocks using regressions on daily returns in the 60-month estimation window preceding t. The risk-adjusted returns of the tail beta spread portfolios are calculated by averaging the risk-adjusted returns of the individual stocks in each portfolio. Based on the constructed portfolio returns, we then construct the zero-investment portfolio, which is obtained by taking a long position of US$1 in the portfolio with the 20% highest tail beta spreads, while taking a short position of US$1 in the portfolio with the 20% lowest spreads.

IV. Results

A. Descriptive Statistics

Table 1 reports the descriptive statistics averaged across the stocks in the portfolios sorted on tail betas. Not surprisingly, there is a strong positive relation between the loadings on systematic risk and the loadings on systematic tail risk.

TABLE 1

Descriptive Statistics

At the start of each month t between July 1968 and Dec. 2010 we estimate tail betas of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks by applying the extreme value theory approach in equation (5) to daily returns from the 60 months preceding t. Stocks are sorted into 5 quintiles according to their tail beta estimates. We exclude stocks with more than 60% zero daily returns in the 60 months preceding t, and stocks with a price below US$5 at the end of the month preceding t. The reported numbers are averages for the stocks in each sort. We average first across firms at each t, and then over the 510 months in the sample. The market beta (β), downside beta (β^-), standard deviation, idiosyncratic volatility, tail dependence, skewness, coskewness, excess kurtosis, and cokurtosis are calculated using the daily returns from the 60 months preceding t. We estimate β^- through a regression conditional on below-average market returns. Idiosyncratic volatility is calculated as the standard deviation of the residuals obtained from regressing individual stock returns on the Fama–French (1993) factors. The tail dependence measure, τ, is calculated following the estimator in equation (7) with $k = 50$. Coskewness and cokurtosis are calculated as

$$\beta_j^{\text{SKD}} = \frac{E\left[\epsilon_j \left(R_m^e - \bar{R}_m^e\right)^2\right]}{\sqrt{E\left[\epsilon_j^2\right]} E\left[\left(R_m^e - \bar{R}_m^e\right)^2\right]} \quad \text{and} \quad \beta_j^{\text{KUD}} = \frac{E\left[\epsilon_j \left(R_m^e - \bar{R}_m^e\right)^3\right]}{\sqrt{E\left[\epsilon_j^2\right]} \left(E\left[\left(R_m^e - \bar{R}_m^e\right)^2\right]\right)^{3/2}},$$

where ϵ_j denotes the residual from a regression of R_j^e on R_m^e. We provide trading volume and market capitalization at the end of the month preceding t.

Characteristic	High β^T	4	3	2	Low β^T
Beta (β)	1.28	1.00	0.82	0.67	0.42
Tail beta spread ($\beta^T - \beta$)	1.34	0.95	0.76	0.60	0.46
Downside beta (β^-)	1.46	1.12	0.91	0.73	0.46
Standard deviation	4.29	3.19	2.57	2.12	1.69
Idiosyncratic risk	4.01	2.97	2.38	1.97	1.60
Tail dependence (τ)	0.22	0.21	0.21	0.19	0.15
Skewness	0.80	0.57	0.45	0.42	0.47
Coskewness (β^{SKD})	−0.11	−0.15	−0.19	−0.19	−0.21
Excess kurtosis	10.27	9.05	8.72	9.08	11.23
Cokurtosis (β^{KUD})	−0.79	−0.23	0.39	1.21	2.26
Market capitalization (US$billions)	0.72	1.16	2.11	2.96	2.55
Volume (million shares)	10.00	7.96	7.85	7.55	5.09

On average, stocks with high tail betas tend to have high market betas and high downside betas. There is a clear trend in the (downside) betas across the portfolios. The relation between the tail beta and the market beta stresses the importance of sorting on the tail beta spread when correcting for market risk. Similar patterns can be observed for other risk measures such as volatility, idiosyncratic risk, coskewness, and tail dependence with market risk. These trends indicate that stocks with higher tail betas also tend to have higher values for other potential (systematic) risk measures, with cokurtosis as a notable exception. Finally, high-tail-beta stocks tend to have smaller sizes and higher trading volumes. The observed patterns between tail beta and other risk measures call for an elaborate treatment of these risk measures in robustness checks.

B. Persistence

We verify whether the estimates of tail beta obtained from historical data are persistent over time. In the absence of such persistence, estimating tail betas based on historical data would merely serve a descriptive function and would provide no insight into future comovements during adverse market conditions. To investigate this issue we provide transition matrices based on tail beta estimates and their 60 months lagged estimates in Table 2. We estimate tail betas by using both the EVT approach with $k=50$ and a regression conditional on the 50 sharpest market declines. The table also provides a similar matrix for market betas estimated from a regression with the CAPM as the benchmark model (based on approximately 1,260 observations). Higher numbers along and around the diagonal point to a more persistent sorting.

We observe two patterns from the transition matrices. First, the numbers along the diagonal of the transition matrices with tail betas are higher for the matrix constructed with the EVT approach. This suggests that the EVT approach provides a more persistent classification of firms' sensitivity to systematic tail risk than the conditional regression approach. This is potentially due to the latter approach having a larger standard error. Second, the numbers along and around the diagonal of the transition matrices based on tail betas estimated with the EVT approach are, in general, at a similar level to those based on market betas. This suggests a similar level of persistence in the sorting based on the two betas. Hence, given that historical market betas contain useful information about future comovement with the market, there seems to be no reason to worry about the persistence in tail betas estimated with the EVT approach, even though they are estimated from fewer observations.

C. Expected Returns

Table 3 reports the excess returns of the portfolios sorted on tail beta. The table provides unconditional averages and averages conditional on days with a sharp market decline defined as $R^e_{m,t} < -2\%$. Such a market loss occurred on 282 days, or approximately 2.5% of all days in our sample. These conditional averages represent the losses in the portfolios under extremely adverse market conditions. They provide direct evidence on whether historical tail betas capture future sensitivity toward systematic tail risk.

TABLE 2

Transition Matrices Based on 60 Months Lagged Values

Table 2 provides transition matrices for tail betas and regular market betas based on 60 months lagged values. At the start of each month *t* between July 1968 and Dec. 2005 we estimate tail betas of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks by applying the extreme value theory (EVT) approach in equation (5) to daily returns from the 60 months preceding *t*. Stocks are sorted into 5 quintiles according to their tail beta estimates. We exclude stocks with more than 60% zero daily returns in the 60 months preceding *t*, and stocks with a price below US$5 at the end of the month preceding *t*. We also determine the allocation of each firm in the tail beta quintiles based on tail betas estimated at $t+60$ months. For each tail beta quintile at time *t*, we calculate the percentage of surviving firms allocated in each tail beta quintile at $t + 60$ months. We repeat the procedure for market betas and tail betas obtained from, respectively, an unconditional regression and a conditional regression (based on the 50 worst market returns) on the daily returns using the capital asset pricing model as benchmark model. The numbers in the transition matrices are averages over time. Higher numbers along and around the diagonal of the transition matrices point toward a more persistent sorting.

	$t + 60$ Months				
Panel A. Tail Beta					
t	High β^T	4	3	2	Low β^T
EVT Approach					
High β^T	49	29	14	6	3
4	21	29	27	17	6
3	9	18	29	29	13
2	4	11	23	35	26
Low β^T	2	5	11	22	61
Conditional Regression					
High β^T	28	22	18	16	15
4	21	22	22	19	16
3	16	20	22	22	19
2	13	18	22	24	23
Low β^T	13	16	20	23	28
Panel B. Market Beta					
t	High β	4	3	2	Low β
High β	53	28	13	5	2
4	23	33	26	13	5
3	11	24	31	24	10
2	5	13	24	33	25
Low β	2	4	10	25	59

The unconditional averages on the zero-investment portfolio are not significantly different from 0 for either the value- or the equal-weighted portfolios. Nevertheless, the zero-investment portfolios suffered large losses on days with sharp market declines. For example, the high-tail-beta portfolio incurred on average a value-weighted loss of 4.69%, whereas the low-tail-beta portfolio incurred a corresponding loss of only 1.81%. A similar pattern is observed for equal-weighted portfolios. Hence, shorting the high-tail-beta portfolio while taking a long position in the low-tail-beta portfolio would have provided significant protection against systematic tail risk, without bearing a cost in the long run. In fact, such a strategy would have led to a positive, albeit insignificant, average return.

A potential reason for not observing higher average returns for high-tail-beta stocks is the inclusion of the perhaps unusually large losses that high-tail-beta stocks suffered during the recent financial crisis. If this was the explanation for our results, we would observe a risk premium on systematic tail risk in the period preceding the financial crisis. However, the data do not support this supposition. The risk premium remains absent when excluding the period from the start of 2007 onward, as shown in the lower panel of Table 3. Moreover, the significance of the difference in performance among the portfolios during sharp market declines does not depend on including or excluding the financial crisis.

TABLE 3

Returns of Stocks Sorted by Past Tail Betas

At the start of each month t between July 1968 and Dec. 2010 we estimate tail betas of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks by applying the extreme value theory approach in equation (5) to daily returns from the 60 months preceding t. We form 5 equal-weighted (EW) and value-weighted (VW) portfolios by sorting on the tail beta, and construct a zero-investment portfolio (High − Low). We exclude stocks with more than 60% zero daily returns in the 60 months preceding t, and stocks with a price below US$5 at the end of the month preceding t. Table 3 reports average excess returns for each portfolio. We report the unconditional average (all days) and averages conditional on days that the market factor lost at least 2% of its value (adverse), that is, $\bar{R}_p^e | R_m^e < -2\%$. The precrisis results exclude observations from 2007 onward. The Newey–West (1994) corrected t-statistics for unconditional averages and standard t-statistics for conditional averages are reported in parentheses.

Weights	$\bar{R}_p^e$	No. of Obs.	High β^T	4	3	2	Low β^T	High − Low
Panel A. Full Sample								
EW	All days	10,704	0.03	0.04	0.04	0.04	0.03	0.00
			(1.7)	(2.8)	(3.3)	(3.5)	(4.1)	(−0.1)
	Adverse	282	−3.96	−3.04	−2.60	−2.18	−1.49	−2.47
			(−31.7)	(−29.3)	(−28.0)	(−27.2)	(−23.5)	(−27.7)
VW	All days	10,704	0.02	0.02	0.02	0.02	0.02	−0.01
			(0.8)	(1.7)	(1.8)	(2.0)	(3.2)	(−0.6)
	Adverse	282	−4.69	−3.83	−3.25	−2.67	−1.81	−2.88
			(−33.6)	(−33.7)	(−31.5)	(−28.1)	(−21.5)	(−24.0)
Panel B. Precrisis								
EW	All days	9,697	0.03	0.04	0.04	0.04	0.03	0.00
			(1.5)	(2.8)	(3.2)	(3.5)	(4.3)	(−0.4)
	Adverse	191	−3.55	−2.58	−2.18	−1.84	−1.21	−2.34
			(−26.2)	(−23.9)	(−21.9)	(−21.0)	(−18.8)	(−22.3)
VW	All days	9,697	0.01	0.02	0.02	0.02	0.03	−0.01
			(0.7)	(1.6)	(1.7)	(2.1)	(3.2)	(−0.7)
	Adverse	191	−4.46	−3.65	−3.01	−2.44	−1.63	−2.83
			(−28.5)	(−28.4)	(−25.7)	(−21.1)	(−15.2)	(−19.4)

The reported results do not guarantee that tail betas provide additional information on the sensitivity to systematic tail risk over and above the information provided by regular market betas. In other words, the differences in losses may also be a consequence of underlying differences in market betas across the portfolios. Although such an explanation would be surprising given the absence of a premium in Table 3, we check whether the differences in losses remain significant when controlling for standard risk factors. We sort the stocks based on their tail beta spreads and calculate risk-adjusted returns using the Fama and French (1993) model (FF3) as the benchmark model in equation (8). The risk-adjusted returns reported in the first two rows in Table 4 follow a similar pattern as before. After controlling for market risk, the losses on the equal- and value-weighted, zero-investment portfolios during market crashes are, respectively, 0.39% and 0.47% (both highly significant). Hence, tail betas do provide information on the sensitivity to systematic tail risk that is not captured by regular market betas.

Table 4 also reports the results on the potential premium when controlling for the FF3 risk factors. The average risk-adjusted returns on the value-weighted portfolios provide no evidence of an additional risk premium for loading on systematic tail risk. Only the equal-weighted, zero-investment portfolio earned a positive average return (borderline significant). The difference between value and equal weighting is related to differences in firm size within each portfolio.

https://doi.org/10.1017/S0022109016000193 Published online by Cambridge University Press

TABLE 4

Risk-Adjusted Returns of Stocks Sorted by Past Tail Betas

At the start of each month between July 1968 and Dec. 2010 we estimate tail betas of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks by applying the extreme value theory approach in equation (5) to daily returns from the 60 months preceding t. We form 5 equal-weighted (EW) and value-weighted (VW) portfolios by sorting on the spread between the tail beta and the market beta and construct a zero-investment portfolio (High − Low). We exclude stocks with more than 60% zero daily returns in the 60 months preceding t, and stocks with a price below US$5 at the end of the month preceding t. We calculate risk-adjusted returns by applying equation (8) on daily stock returns at time t, where the loadings on the risk factors in the benchmark model are estimated for each stock by an ordinary least squares regression on daily returns from the 60 months preceding t. Table 4 reports the average Fama–French (1993) adjusted portfolio return, $\bar{R}^*_{\mathrm{FF3},p}$. We report the unconditional average (all days) and averages conditional on days that the market factor lost at least 2% of its value (adverse), that is, $\bar{R}^*_{\mathrm{FF3},p}|R^e_m < -2\%$. The third and fourth sets report the averages after first presorting the stocks in 5 size cohorts and then sorting on the tail beta spread within each size cohort, where size is measured by market capitalization at the end of month preceding t. The Newey–West (1994) corrected t-statistics for unconditional averages and the standard t-statistics for conditional averages are reported in parentheses.

Weights	$\bar{R}^*_{\mathrm{FF3},p}$	Presort	High $\beta^T - \beta$	4	3	2	Low $\beta^T - \beta$	High − Low
1. EW	All days	—	0.01 (3.4)	0.01 (4.4)	0.01 (3.6)	0.01 (3.0)	0.00 (2.4)	0.01 (2.0)
	Adverse	—	−0.26 (−6.2)	−0.03 (−1.1)	0.01 (0.5)	0.07 (3.3)	0.13 (6.5)	−0.39 (−8.0)
2. VW	All days	—	−0.01 (−1.0)	0.00 (−0.7)	0.00 (−0.2)	0.00 (0.7)	0.00 (1.1)	−0.01 (−1.2)
	Adverse	—	−0.39 (−4.5)	−0.16 (−2.9)	−0.06 (−1.5)	0.02 (0.8)	0.07 (3.7)	−0.47 (−4.9)
3. VW	All days	Small	−0.02 (−2.6)	0.00 (0.2)	0.01 (1.2)	0.01 (1.4)	0.01 (3.1)	−0.03 (−4.6)
		2	−0.02 (−2.7)	0.00 (−1.1)	0.00 (1.1)	0.00 (0.7)	0.01 (1.7)	−0.02 (−3.2)
		3	−0.01 (−1.6)	0.00 (0.4)	0.00 (0.4)	0.00 (0.1)	0.01 (2.1)	−0.01 (−2.3)
		4	0.00 (−0.8)	0.00 (−0.1)	0.00 (−0.7)	0.00 (0.4)	0.00 (0.7)	−0.01 (−1.1)
		Large	0.00 (0.2)	0.00 (−0.5)	0.00 (−0.5)	0.00 (1.8)	0.00 (0.9)	0.00 (−0.1)
		Average	−0.01 (−2.2)	0.00 (−0.4)	0.00 (0.7)	0.00 (1.4)	0.01 (3.1)	−0.01 (−3.2)
4. VW	Adverse	Small	−0.34 (−5.4)	−0.13 (−2.4)	0.01 (0.2)	−0.01 (−0.2)	0.08 (2.0)	−0.42 (−6.5)
		2	−0.52 (−9.8)	−0.13 (−3.5)	0.05 (1.2)	0.04 (1.3)	0.13 (4.2)	−0.65 (−10.4)
		3	−0.38 (−7.3)	−0.01 (−0.4)	0.06 (1.9)	0.11 (3.5)	0.19 (6.1)	−0.57 (−8.7)
		4	−0.18 (−3.3)	0.01 (0.1)	0.04 (1.3)	0.07 (2.2)	0.13 (4.5)	−0.31 (−5.5)
		Large	−0.19 (−2.9)	−0.09 (−2.3)	0.04 (1.4)	0.10 (3.6)	0.05 (1.6)	−0.24 (−3.1)
		Average	−0.32 (−7.5)	−0.07 (−2.6)	0.04 (1.7)	0.06 (2.8)	0.12 (6.0)	−0.44 (−9.0)

Consequently, we apply a double-sorting procedure to evaluate the patterns among similar sized firms. First, we presort stocks into five size cohorts based on market capitalization at the end of the month preceding t. Subsequently, we sort on the tail beta spread within each size cohort. We report the value-weighted, risk-adjusted returns of each portfolio within each size cohort in the third and fourth sets of Table 4.[7] The last line is the average within each tail beta spread quintile

[7]The unconditional average risk-adjusted return of the zero-investment strategy is also negative but insignificant within each size cohort in equal-weighted portfolios (unreported).

across the different size cohorts. The portfolios constructed from high-tail-beta spreads report significantly larger risk-adjusted losses during market crashes and lower average risk-adjusted returns over the entire period within each size cohort. Hence, contrary to the theoretical prediction, high tail betas seem to be associated with lower, rather than higher, average returns. Note that the difference in losses on adverse market days between high- and low-tail-beta spreads are strongest within the three smallest size cohorts. Accordingly, one would expect the theoretical prediction to be strongest within these size cohorts. However, these size cohorts provide the strongest refutation of the predicted presence of a positive systematic tail risk premium by reporting negative and significant risk-adjusted returns on the zero-investment portfolio over the entire period.

In summary, our results confirm that estimated tail betas help predict losses in future stock market crashes; that is, they capture future systematic tail risk. Nevertheless, we find no evidence that investing in high tail betas earns a positive and significant premium. These results suggest that the safety-first framework may not be suitable for explaining the cross section of expected returns or, at least, not with the assumption that investors are concerned only with extreme losses that occur with some very small probability.

D. Robustness

We perform robustness checks in several directions. Following Daniel and Titman (1997), we apply a double-sorting procedure to examine whether our results are also captured by other stock characteristics, namely, downside beta (spread); (co)skewness; (co)kurtosis; (idiosyncratic) volatility; short-, medium-, and long-term past performance; and trading volume. Table 5 reports the results on FF3-adjusted returns after averaging within each tail beta quintile across the cohorts. The larger losses of stocks with higher tail betas remain significant after presorting on these characteristics. The absence of a positive systematic tail risk premium is also robust (unreported). Furthermore, our results remain qualitatively unchanged if the FF3 benchmark model is extended to include other asset pricing factors, such as momentum and short- and long-term reversal factors, or factors constructed from downside beta (spread), coskewness, and cokurtosis.

We also consider several methodological variations. Successively, we restrict the sample to the period after the entrance of NASDAQ firms to the CRSP data in Jan. 1973; we restrict our sample to NYSE firms only; we decrease the number of worst days, k, used in the estimation of tail beta from 50 to 30; we replace the EVT approach by the conditional regression approach; and we use monthly returns instead of daily returns to estimate market betas and other factor loadings in the benchmark model. Finally, we use alternative thresholds to define extremely adverse market conditions, that is, excess market return below -5% or 0%. The results remain qualitatively unchanged under all these methodological changes.

E. Downside Beta and Tail Beta

The lack of a premium for high-tail-beta stocks among all firms and the (limited) evidence of a negative premium among small and medium-sized firms contrast

TABLE 5

Presorting and Risk-Adjusted Returns under Adverse Market Conditions

At the start of each month between July 1968 and Dec. 2010 we estimate tail betas of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks by applying the extreme value theory approach in equation (5) to daily returns from the 60 months preceding t. We first presort the equities into 5 cohorts according to the stock characteristic specified in the first column. We exclude stocks with more than 60% zero daily returns in the 60 months preceding t, and stocks with a price below US\$5 at the end of the month preceding t. Within each cohort we form 5 value-weighted portfolios by sorting on the spread between tail beta and market beta and construct a zero-investment portfolio (High − Low). We calculate risk-adjusted returns by applying equation (8) to daily stock returns in the month containing t, where the loadings on the risk factors in the benchmark model are estimated for each stock by an ordinary least squares regression on daily returns from the 60 months preceding t. The reported numbers are the risk-adjusted returns averaged within each tail beta quintile across the cohorts based on the presorting characteristic. Table 5 reports the average Fama–French (1993) adjusted portfolio return on days that the market factor lost at least 2% of its value, that is, $\bar{R}^{*}_{\text{FF3},p}|R^{e}_{m} < -2\%$. Standard t-statistics are reported in parentheses.

Presorting Characteristic	High $\beta^T - \beta$	4	3	2	Low $\beta^T - \beta$	High − Low
Size	−0.32	−0.07	0.04	0.06	0.12	−0.44
	(−7.5)	(−2.6)	(1.7)	(2.8)	(6.0)	(−9.0)
Downside beta (β^-)	−0.38	−0.17	−0.06	−0.03	0.10	−0.49
	(−5.2)	(−3.4)	(−1.7)	(−1.3)	(5.0)	(−6.2)
Downside beta spread ($\beta^- - \beta$)	−0.36	−0.05	−0.05	0.05	0.01	−0.37
	(−5.6)	(−1.4)	(−1.4)	(1.9)	(0.4)	(−5.4)
Coskewness	−0.39	−0.09	−0.03	−0.01	0.03	−0.42
	(−5.7)	(−2.3)	(−0.8)	(−0.5)	(1.3)	(−5.8)
Cokurtosis	−0.36	−0.10	−0.04	0.00	0.06	−0.42
	(−5.0)	(−2.0)	(−1.2)	(0.1)	(3.1)	(−5.3)
Idiosyncratic volatility	−0.37	−0.21	−0.03	−0.04	0.18	−0.55
	(−5.2)	(−4.5)	(−0.9)	(−1.2)	(4.4)	(−6.4)
Past 1 month performance	−0.32	−0.10	−0.09	0.00	0.07	−0.39
	(−4.6)	(−2.0)	(−2.4)	(0.0)	(3.4)	(−5.2)
Past 2–12 months performance	−0.30	−0.10	−0.05	0.03	0.10	−0.40
	(−4.3)	(−2.2)	(−1.5)	(1.2)	(4.7)	(−5.4)
Past 13–60 months performance	−0.37	−0.10	−0.03	0.01	0.10	−0.48
	(−5.5)	(−2.2)	(−1.0)	(0.3)	(5.2)	(−6.4)
Volatility	−0.37	−0.22	−0.05	−0.02	0.19	−0.56
	(−5.2)	(−4.7)	(−1.4)	(−0.7)	(4.5)	(−6.4)
Skewness	−0.39	−0.12	−0.06	0.01	0.05	−0.44
	(−5.6)	(−2.7)	(−1.9)	(0.4)	(2.4)	(−5.8)
Kurtosis	−0.37	−0.16	−0.06	0.02	0.04	−0.42
	(−6.0)	(−3.5)	(−1.9)	(0.9)	(2.1)	(−5.9)
Volume	−0.30	−0.06	0.00	0.06	0.10	−0.40
	(−6.5)	(−1.7)	(−0.1)	(2.6)	(5.0)	(−8.1)

with the positive premium associated with downside beta (see, e.g., Ang, Chen, and Xing (2006)). Although both measures evaluate the comovement of stocks with downward movements of the market portfolio, there is a conceptual difference between the two. The downside beta measures stocks' comovement in the more general scenario of a below-average or zero market return, whereas the tail beta focuses on the comovement if market downturns are extreme. This difference also shows up in the robustness checks in Table 5, where the ability of tail beta to predict which firms suffer larger FF3-adjusted losses in extreme market downturns remains after presorting stocks on downside beta (spread).

Ang, Chen, and Xing (2006) report a positive in-sample risk premium for holding high-downside-beta stocks with downside betas estimated from 12 months of daily returns conditional on below-average market returns. To contrast their results with ours, we follow their procedure on our sample. At the end of each month between June 1964 and Dec. 2009, we estimate regular market betas, downside betas, and tail betas from daily returns over the next 12 months.[8] We construct equal-weighted portfolios from sorts based on downside beta (spread) and tail beta (spread). Table 6 reports the annualized average excess returns of these portfolios, calculated over the same period as that used to compute the betas.

TABLE 6

Returns of Stocks Sorted by Realized Betas

Table 6 lists the equal-weighted average excess returns and risk characteristics of stocks sorted by realized betas. For each month t between June 1963 and Dec. 2009 we estimate β, β^-, and β^T of New York Stock Exchange (NYSE), American Stock Exchange (AMEX), and National Association of Securities Dealers Automated Quotations (NASDAQ) stocks from daily returns over the next 12 months. β^- is obtained from a regression conditional on below-average market returns, and β^T is obtained by applying the extreme value theory approach in equation (5). We exclude stocks with more than 60% zero daily returns in the 12 months following month t, and stocks with a price below US$5 at the end of month t. Following Ang, Chen, and Xing ((2006), Table 1), we form 5 equal-weighted portfolios by sorting on the corresponding risk characteristics at the end of month t and construct a zero-investment portfolio (High − Low). $\bar{R}_p^e$ is the annualized (daily compounded) average return in excess of the 1-month Treasury bill rate over the next 12 months (the same 12-month period as the period used to compute stock betas). The entry labeled "t-stat" is the t-statistic (in parentheses) for testing against the null hypothesis of a zero average return on the High − Low portfolio, computed using Newey–West (1987) heteroskedasticity- and autocorrelation-robust standard errors with 12 lags. The columns labeled "β," "β^-," and "β^T" report the cross-sectional average of the stock betas in each portfolio. We first average across firms at each t, and then over the 559 months in the sample.

Portfolio	$\bar{R}_p^e$	β	β^-	β^T	Portfolio	$\bar{R}_p^e$	β	β^-	β^T
Panel A. Stocks Sorted by Realized β^-					*Panel B. Stocks Sorted by Realized β^T*				
1 Low β^-	1.35	0.34	0.11	1.05	1 Low β^T	5.50	0.41	0.39	0.54
2	3.81	0.61	0.60	1.30	2	9.20	0.66	0.72	1.23
3	4.88	0.84	0.93	1.59	3	8.33	0.87	0.97	1.63
4	6.15	1.11	1.32	1.96	4	5.80	1.10	1.26	2.11
5 High β^-	9.15	1.60	2.11	2.63	5 High β^T	−5.15	1.46	1.75	3.01
High − Low	7.80	1.26	2.00	1.58	High − Low	−10.65	1.06	1.36	2.46
t-stat.	(2.18)				t-stat.	(−2.42)			
Panel C. Stocks Sorted by Realized $\beta^- - \beta$					*Panel D. Stocks Sorted by Realized $\beta^T - \beta$*				
1 Low $\beta^- - \beta$	0.49	1.02	0.55	1.65	1 Low $\beta^T - \beta$	9.61	0.92	0.83	0.88
2	5.44	0.84	0.75	1.45	2	11.81	0.91	0.94	1.37
3	6.67	0.81	0.90	1.50	3	9.31	0.91	1.01	1.63
4	7.76	0.85	1.15	1.70	4	4.94	0.92	1.11	1.98
5 High $\beta^- - \beta$	7.01	0.99	1.73	2.23	5 High $\beta^T - \beta$	−10.27	0.85	1.18	2.67
High − Low	6.52	−0.03	1.18	0.58	High − Low	−19.88	−0.07	0.35	1.79
t-stat.	(5.53)				t-stat.	(−8.42)			

Panels A and B of Table 6 report the results after sorting on downside beta and tail beta, respectively. Portfolios with high downside betas and high tail betas also tend to have high regular market betas. In line with Ang, Chen, and Xing (2006), high-downside-beta stocks are associated with significantly higher in-sample returns. However, high-tail-beta stocks are associated with significantly lower in-sample returns. This contrast between downside beta and tail beta becomes more pronounced when sorting on the spreads in Panels C and D. The average regular market beta across the sorts based on spreads is almost flat in both panels. The zero-investment portfolio based on downside beta spread earns an annual premium of 6.5% (t-statistic = 5.5), whereas the zero-investment portfolio

[8]With on average about 252 daily observations in each estimation window of 12 months, we choose $k = 15$ for estimating the tail beta, or $k/n \approx 8\%$. The relatively short estimation window is a disadvantage for obtaining an accurate estimate of tail beta. Nevertheless, the in-sample results from the estimation window of 12 months are in line with our other results based on estimates from longer estimation windows.

based on tail beta spread takes an annual loss of on average 19.9% (t-statistic = −8.4). Naturally, these in-sample results may be partly due to a mechanical relation. For example, large losses in market crashes may result directly in both high tail betas and low in-sample average returns. However, such a mechanical relation cannot explain the negative premium for loading on tail beta observed in the out-of-sample results for small and medium-sized firms in Table 4.

In summary, different from downside beta, we observe no positive and significant in-sample premium associated with tail beta. The combination of a positive premium for downside beta and the lack of such a premium for tail beta in our results suggests that investors are predominantly concerned with their exposure to general downside risk in their portfolio choice rather than focusing only on the potential extreme losses that occur with some very low probability.

V. Risk Management

Because historical tail betas can capture future losses under extremely adverse market conditions, tail betas may help investors assess the tail risk of portfolios. As an additive measure of loading on systematic tail risk, the tail beta is a useful measure in the context of managing the tail risk of portfolios. We discuss this application in the current section.

We consider a portfolio consisting of d assets, following the linear tail model in equation (1) with nonnegative tail betas, $\beta_1^T, \ldots, \beta_d^T$. Under extremely adverse market conditions, the excess return of a portfolio with nonnegative investment weights, $w_1, \ldots, w_d$, can be written as

$$(9) \qquad R_P^e = \left(\sum_{j=1}^{d} w_j \beta_j^T \right) R_m^e + \sum_{j=1}^{d} w_j \varepsilon_j, \quad \text{for} \quad R_m^e < -\text{VaR}_m(\bar{p}).$$

Hence, the portfolio return also follows a linear tail model with a portfolio tail beta equal to the weighted average of the tail beta of the individual assets, that is, $\beta_P^T = \sum_{j=1}^{d} w_j \beta_j^T$, and an idiosyncratic component that is given by $\varepsilon_P = \sum_{j=1}^{d} w_j \varepsilon_j$.

To evaluate the tail risk of a portfolio, it is necessary to aggregate the systematic and idiosyncratic tail risks. We start by discussing the aggregation for a single asset. Suppose the linear tail model in equation (1) and the heavy-tailed setup in equation (4) hold for a larger area, $\min(R_m^e, R_j^e) < -\text{VaR}_m(\bar{p})$. It then follows that the probability of a loss on asset j larger than u can be approximated by

$$(10) \qquad \Pr(R_j^e < -u) \sim \Pr\left(\beta_j^T R_m^e < -u\right) + \Pr(\varepsilon_j < -u), \quad \text{as} \quad u \to \infty.$$

This approximation follows from Feller's (1971) convolution theorem on aggregating risk factors, which states that the probability that the sum of independent heavy-tailed risk factors is above a high threshold can be approximated by the sum of the probabilities of each risk factor being above that threshold.[9] Suppose the idiosyncratic risk, ε_j, follows a heavy-tailed distribution with tail index α_{ε_j} and

[9]Embrechts, Klüppelberg, and Mikosch ((1997), Lemma 1.3.1) provides the proof for the case $\alpha_m = \alpha_{\varepsilon_j}$. Along the same lines of proof, one can obtain that this relation holds for $\alpha_m \neq \alpha_{\varepsilon_j}$.

https://doi.org/10.1017/S0022109016000193 Published online by Cambridge University Press

scale A_{ε_j}.[10] If $\alpha_{\varepsilon_j} > \alpha_m$, then the systematic tail risk dominates the idiosyncratic tail risk, that is, $\Pr(\varepsilon_j < -u) = o(\Pr(\beta_j^T R_m^e < -u))$ as $u \to \infty$. Consequently, the downside tail distribution of the excess asset return, R_j^e, follows a heavy-tailed distribution with tail index $\alpha_j = \alpha_m$ and scale $A_j = (\beta_j^T)^{\alpha_m} A_m$. In contrast, if $\alpha_{\varepsilon_j} < \alpha_m$, then the idiosyncratic risk dominates the tail risk of the asset, and we have $\alpha_j = \alpha_{\varepsilon_j}$ and $A_j = A_{\varepsilon_j}$. In the case $\alpha_{\varepsilon_j} = \alpha_m$, both of the two components contribute to the tail risk of the asset, and we have $A_j = (\beta_j^T)^{\alpha_m} A_m + A_{\varepsilon_j}$.

In the portfolio context, we first consider the case $\alpha_{\varepsilon_1} = \cdots = \alpha_{\varepsilon_d} = \alpha_m$. Suppose the idiosyncratic tail risks are independent with scales $A_{\varepsilon_1}, \ldots, A_{\varepsilon_d}$. Following Feller's (1971) convolution theorem, the downside tail of the portfolio follows a heavy-tailed distribution with tail index $\alpha_P = \alpha_m$ and scale

$$A_P \quad = \quad \left(\beta_P^T\right)^{\alpha_m} A_m + \sum_{j=1}^{d} w_j^{\alpha_m} A_{\varepsilon_j}. \tag{11}$$

In practice, all parameters in equation (11) can be statistically estimated. In particular, the tail beta of the portfolio, β_P^T, can be obtained by taking a weighted average of the tail beta estimates of the individual assets, the $\hat{\beta}_j^T$s. Furthermore, the scales of the idiosyncratic tail risks, A_{ε_j}, can be obtained from

$$\hat{A}_{\varepsilon_j} \quad = \quad \hat{A}_j - \left(\hat{\beta}_j^T\right)^{\hat{\alpha}_m} \hat{A}_m, \tag{12}$$

where the scales of the market return and the asset return, A_m and A_j, can be estimated by univariate EVT analysis (see, e.g., Hill (1975)). With equation (11), we thus obtain the estimate of the scale of a portfolio. Subsequently, the VaR of the portfolio for some low probability level p can be calculated from the approximation

$$\text{VaR}_P(p) \quad \approx \quad \left(\frac{A_P}{p}\right)^{1/\alpha_m}. \tag{13}$$

Next, consider the case in which some assets in the portfolio correspond to $\alpha_{\varepsilon_j} > \alpha_m$. The idiosyncratic tail risks of those assets are dominated by their systematic tail risk and do not contribute to the tail risk of the portfolio. Hence, it is still possible to evaluate the scale of the portfolio with equation (11) by omitting the idiosyncratic tail risks of those assets. However, it is not necessary to identify those assets or to modify the estimation procedure from equations (11) and (12). Assets with $\alpha_{\varepsilon_j} > \alpha_m$ exhibit complete tail dependence with the market return, that is, $\tau_j = 1$ and $A_j = (\beta_j^T)^{\alpha_m} A_m$. Therefore, the estimator on A_{ε_j} in equation (12) converges to zero under the EVT approach. Including the estimate of A_{ε_j} for such assets in equation (11) will not contaminate the estimate of the portfolio scale. In summary, equation (11) can be applied to any portfolio consisting of assets with $\alpha_{\varepsilon_j} \geq \alpha_m$.

Finally, we discuss the case in which some assets correspond to $\alpha_{\varepsilon_j} < \alpha_m$. Theoretically, the downside tail risk of the portfolio would be dominated by the idiosyncratic risk of the asset with the lowest tail index. However, in practice this may not be the case. The reason is that the return on many assets is in fact bounded

[10]A thin-tailed idiosyncratic risk could be thought of as having $\alpha_{\varepsilon_j} = \infty$ in the following discussion.

https://doi.org/10.1017/S0022109016000193 Published online by Cambridge University Press

from below by -100%.[11] Such an asset j with investment weight w_j can generate a maximum loss of w_j. Therefore, in a well-diversified portfolio with a sufficiently large number of assets, the idiosyncratic tail risks do not contribute to the tail risk of the portfolio under the condition that their returns have a lower bound. This is achieved even if some assets correspond to the case $\alpha_{\varepsilon_j} < \alpha_m$.[12] In contrast to the idiosyncratic risks, the systematic tail risk cannot be diversified away by investing in a large number of assets, because the tail beta of a portfolio is the weighted average of those of the individual assets. Hence, for any well-diversified portfolio consisting of a sufficiently large number of assets with lower bounded returns, the scale of its downside tail distribution can be approximated by

$$A_P = \left(\sum_{j=1}^{d} w_j \beta_j^T\right)^{\alpha_m} A_m.$$

Subsequently, the VaR can be calculated from equation (13).

VI. Concluding Remarks

This article investigates whether systematic tail risk is compensated in the cross section of expected returns. Asset pricing theory based on an equilibrium framework with safety-first investors suggests that higher loadings on systematic tail risk should be associated with a positive risk premium if investors are predominantly concerned with extreme losses in low-probability events. Theoretically, the risk premium is proportional to its tail beta, which measures the sensitivity to systematic tail risk. Based on an EVT approach, we estimate tail betas and test empirically whether high-tail-beta stocks receive higher average returns.

We find that assets with higher tail betas are associated with significantly larger losses during future extreme market downturns. Hence, historical tail betas are able to capture the sensitivity to future systematic tail risk. Furthermore, the asset pricing tests do not report a positive and significant premium for high-tail-beta stocks over the entire historical sample. One potential reason for this result is that there are measurement issues, such as time variation in the actual tail betas. However, because our historical estimates perform well in differentiating future losses under extremely adverse market conditions, such an explanation is satisfactory only if the risk premium for loading on systematic tail risk is low. This suggests that room for a positive systematic tail risk premium in the cross section of expected returns is limited. In addition, we find some evidence of a negative and significant premium among small and medium-sized firms, which further contradicts the presence of a positive premium for high-tail-beta stocks.

Our results suggest that investors are concerned with the general downside risk of their portfolio rather than considering only the potential extreme losses

[11] Examples of assets in which the returns have a lower bound are long positions in stocks and bonds. Counterexamples are short positions in currencies and stocks.

[12] The lower bound of equity returns is not accounted for in the heavy tail approximation, as in equation (4). Instead, one could consider truncated heavy-tailed distributions. Ibragimov and Walden (2007) prove the diversification effects of bounded risk factors from truncated heavy-tailed distributions provided that the number of risk factors is sufficiently large.

that materialize with a very low probability. Alternative explanations are that investors are insufficiently aware of the cross-sectional differences in the loadings on systematic tail risk or that fund managers are less concerned with extreme tail risk than their clients are.

Another explanation for our results may be that systematic tail risk should be investigated at a different investment horizon. Any difference between the theoretical investment horizon and the frequency used for estimation causes a systematic bias in the linear beta coefficients.[13] Our result on the lack of a systematic tail risk premium may change if tail betas are estimated from low-frequency data. However, analyzing tail betas at a lower frequency, such as monthly or annual, remains a difficult and interesting issue, which is beyond the scope of the current study.

Parallel to the tail beta that measures assets' sensitivity to extreme market downturns, individual assets may also exhibit differences in their comovement with large market booms. The methodology to estimate downside tail betas can also be applied to estimate upside tail betas. In the same vein as the discussion on (downside) risk management, such upside tail betas may provide information on portfolio profits in a hypothetical large boom, where the upside tail beta of a portfolio is a weighted average of the upside tail betas of the individual assets. In the safety-first framework, which focuses on downside risk only, upside tail betas are irrelevant for the cross section of expected returns. The relevance of upside tail betas in other asset pricing frameworks is left for future research.

Appendix. Proof of Equation (3)

We start by introducing the notation of Arzac and Bawa (1977). Let the initial and future market value of asset j be denoted by V_j and X_j. Each asset j generates a return $R_j = X_j / V_j$. The market portfolio has the initial and future value $V_m = \sum_j V_j$ and $X_m = \sum_j X_j$. Hence, the market return is defined as

$$R_m \quad = \quad \frac{\sum_j X_j}{\sum_j V_j} \quad = \quad \sum_j w_j^* R_j,$$

with weights $w_j^* = V_j / (\sum_j V_j)$.

Investor i holds a portfolio $(\gamma_{i,1}, \gamma_{i,2}, \ldots)$, which generates a future value as $\sum_j \gamma_{i,j} X_j = \sum_j \gamma_{i,j} V_j R_j$. Denote the p-quantile of this future value and the market portfolio as Q^i and Q_m, respectively. Then, the p-quantile of the market return is $q_m = Q_m / V_m$.

Arzac and Bawa ((1977), eq. (14)) derive the formula to calculate the parameter β_j^{AB} as

$$\beta_j^{\text{AB}} \quad = \quad \frac{q_j - r_f}{q_m - r_f}, \quad \text{with} \quad q_j \quad := \quad \frac{\left.\dfrac{\partial Q^i}{\partial \gamma_{i,j}}\right|_{(\gamma_{i,j}) = (\gamma_i)}}{V_j},$$

[13]We thank the referee for pointing out this issue. Tail betas estimated at a lower frequency may deviate from their high-frequency counterparts for two reasons. The first reason is the mathematical bias discussed by, for example, Levhari and Levy (1977), which applies to all betas estimated from discrete returns. The second reason is that conditioning on extremely adverse market conditions may refer to different observations in case of different frequencies; for example, many small daily losses may add up to a relatively large monthly loss without a single extreme market downturn.

where (γ_i) is the optimal portfolio holding for investor i on all assets. The right-hand side is the same across all investors. Because $q_m - r_f$ is the p-quantile of the market excess return, we have that $q_m - r_f = -\mathrm{VaR}_m(p)$. Therefore, to prove the equality in equation (3), it is necessary to prove only that $q_j = E(R_j|R_m = Q_m(p))$, where $Q_m(p) = q_m$.

To relate the quantile of the future value of investors' portfolio to that of the market return, we define for any positive investments $(u_1, u_2, \ldots)$ the p-quantile of $\sum_j u_j R_j$ as $f(u_1, u_2, \ldots)$. Notice that $Q_m = f(V_1, V_2, \ldots)$, $Q^i = f(\gamma_{i,1}V_1, \gamma_{i,2}V_2, \ldots)$. We calculate q_j as

$$q_j = \frac{\left.\dfrac{\partial f(\gamma_{i,1}V_1, \gamma_{i,2}V_2, \ldots)}{\partial \gamma_{i,j}}\right|_{(\gamma_{i,j})=(\gamma_i)}}{V_j} = \frac{V_j \left.\dfrac{\partial f}{\partial u_j}\right|_{(u_j)=(\gamma_i V_j)}}{V_j} = \left.\frac{\partial f}{\partial u_j}\right|_{(u_j)=(\gamma_i V_j)}.$$

The function f is homogeneous with degree 1, which implies that its partial derivative $\partial f/\partial u_j$ is a homogeneous function with degree 0. Consequently, we have

$$\left.\frac{\partial f}{\partial u_j}\right|_{(u_j)=(\gamma_i V_j)} = \left.\frac{\partial f}{\partial u_j}\right|_{(u_j)=(w_j^*)}.$$

To derive the partial derivative of the f function, we use the expression that

$$f(u_1, u_2, \ldots) = E\left(\sum_j u_j R_j \middle| \sum_j u_j R_j = f(u_1, u_2, \ldots)\right).$$

Thus,

$$\begin{aligned}
\left.\frac{\partial f}{\partial u_j}\right|_{(u_j)=(w_j^*)} &= \left.\frac{\partial}{\partial u_j} E\left(\sum_j u_j R_j \middle| \sum_j u_j R_j = f(u_1, u_2, \ldots)\right)\right|_{(u_j)=(w_j^*)} \\
&= \left.\frac{\partial}{\partial u_j} \sum_j u_j E\left(R_j \middle| \sum_j u_j R_j = f(u_1, u_2, \ldots)\right)\right|_{(u_j)=(w_j^*)} \\
&= \left. E\left(R_j \middle| \sum_j u_j R_j = f(u_1, u_2, \ldots)\right)\right|_{(u_j)=(w_j^*)} \\
&= E(R_j|R_m = Q_m(p)).
\end{aligned}$$

The last equality follows from the fact that $R_m = \sum_j w_j^* R_j$ and $Q_m(p) = f(w_1^*, w_2^*, \ldots)$. The q_j thus quantifies the contribution of asset j to the p-quantile of the market return. □

References

Ang, A., and G. Bekaert. "International Asset Allocation with Regime Shifts." *Review of Financial Studies*, 15 (2002), 1137–1187.

Ang, A., and J. Chen. "Asymmetric Correlations of Equity Portfolios." *Journal of Financial Economics*, 63 (2002), 443–494.

Ang, A.; J. Chen; and Y. Xing. "Downside Risk." *Review of Financial Studies*, 19 (2006), 1191–1239.

Ang, A.; R. Hodrick; Y. Xing; and X. Zhang. "The Cross-Section of Volatility and Expected Returns." *Journal of Finance*, 61 (2006), 259–299.

Arzac, E., and V. Bawa. "Portfolio Choice and Equilibrium in Capital Markets with Safety-First Investors." *Journal of Financial Economics*, 4 (1977), 277–288.

Bali, T.; N. Cakici; and R. Whitelaw. "Hybrid Tail Risk and Expected Stock Returns." *Review of Asset Pricing Studies*, 4 (2014), 206–246.

Bali, T.; K. Demirtas; and H. Levy. "Is There an Intertemporal Relation between Downside Risk and Expected Returns?" *Journal of Financial and Quantitative Analysis*, 44 (2009), 883–909.

Bawa, V., and E. Lindenberg. "Capital Market Equilibrium in a Mean-Lower Partial Moment Framework." *Journal of Financial Economics*, 5 (1977), 189–200.

Carhart, M. "On Persistence in Mutual Fund Performance." *Journal of Finance*, 52 (1997), 57–82.

Chabi-Yo, F.; S. Ruenzi; and F. Weigert. "Crash Sensitivity and the Cross-Section of Expected Stock Returns." Working Paper, University of St. Gallen (2015).

Cholette, L., and C. Lu. "The Market Premium for Dynamic Tail Risk." Working Paper, University of Stavanager and National Chengchi University (2011).

Daniel, K., and S. Titman. "Evidence on the Characteristics of Cross Sectional Variation in Stock Returns." *Journal of Finance*, 52 (1997), 1–33.

De Bondt, W., and R. Thaler. "Does the Stock Market Overreact?" *Journal of Finance*, 40 (1985), 793–805.

De Jonghe, O. "Back to the Basics in Banking? A Micro-Analysis of Banking System Stability." *Journal of Financial Intermediation*, 19 (2010), 387–417.

Dittmar, R. "Nonlinear Pricing Kernels, Kurtosis Preference, and Evidence from the Cross Section of Equity Returns." *Journal of Finance*, 57 (2002), 369–403.

Embrechts, P.; L. De Haan; and X. Huang. "Modelling Multivariate Extremes." In *Extremes and Integrated Risk Management*, P. Embrechts, ed. London, UK: RISK Books (2000), 59–67.

Embrechts, P.; C. Klüppelberg; and T. Mikosch. *Modelling Extremal Events: For Insurance and Finance*. Heidelberg, DE: Springer (1997).

Fama, E., and K. French. "Common Risk Factors in the Returns on Stocks and Bonds." *Journal of Financial Economics*, 33 (1993), 3–56.

Fama, E., and K. French. "Dissecting Anomalies." *Journal of Finance*, 63 (2008), 1653–1678.

Feller, W. *An Introduction to Probability Theory and Its Applications*, Vol. 2. New York, NY: Wiley & Sons (1971).

Gervais, S.; R. Kaniel; and D. Mingelgrin. "The High-Volume Return Premium." *Journal of Finance*, 56 (2001), 877–919.

Harlow, W., and R. Rao. "Asset Pricing in a Generalized Mean-Lower Partial Moment Framework: Theory and Evidence." *Journal of Financial and Quantitative Analysis*, 24 (1989), 285–311.

Harvey, C., and A. Siddique. "Conditional Skewness in Asset Pricing Tests." *Journal of Finance*, 55 (2000), 1263–1295.

Hill, B. "A Simple General Approach to Inference about the Tail of a Distribution." *Annals of Statistics*, 3 (1975), 1163–1174.

Huang, W.; Q. Liu; G. Rhee; and F. Wu. "Extreme Downside Risk and Expected Stock Returns." *Journal of Banking & Finance*, 36 (2012), 1492–1502.

Ibragimov, R., and J. Walden. "The Limits of Diversification When Losses May Be Large." *Journal of Banking & Finance*, 31 (2007), 2551–2569.

Jansen, D., and C. de Vries. "On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective." *Review of Economics and Statistics*, 73 (1991), 18–24.

Jegadeesh, N. "Evidence of Predictable Behavior of Security Returns." *Journal of Finance*, 45 (1990), 881–898.

Karolyi, G., and R. Stulz. "Why Do Markets Move Together? An Investigation of U.S.-Japan Stock Return Comovements." *Journal of Finance*, 51 (1996), 951–986.

Kelly, B., and H. Jiang. "Tail Risk and Asset Prices." *Review of Financial Studies*, 27 (2014), 2841–2871.

King, M., and S. Wadhwani. "Transmission of Volatility between Stock Markets." *Review of Financial Studies*, 3 (1990), 5–33.

Kraus, A., and R. Litzenberger. "Skewness Preference and the Valuation of Risk Assets." *Journal of Finance*, 31 (1976), 1085–1100.

Levhari, D., and H. Levy. "The Capital Asset Pricing Model and the Investment Horizon." *Review of Economics and Statistics*, 59 (1977), 92–104.

Longin, F., and B. Solnik. "Is the Correlation in International Equity Returns Constant: 1960–1990?" *Journal of International Money and Finance*, 14 (1995), 3–26.

Longin, F., and B. Solnik. "Extreme Correlation of International Equity Markets." *Journal of Finance*, 56 (2001), 649–676.

Loretan, M., and P. 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 (1994), 211–248.

Mandelbrot, B. "The Variation of Certain Speculative Prices." *Journal of Business*, 36 (1963), 394–419.

Newey, W., and K. West. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (1987), 703–708.

Newey, W., and K. West. "Automatic Lag Selection in Covariance Matrix Estimation." *Review of Economic Studies*, 61 (1994), 631–653.

Patton, A. "On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation." *Journal of Financial Econometrics*, 2 (2004), 130–168.

Poon, S.; M. Rockinger; and J. Tawn. "Extreme Value Dependence in Financial Markets: Diagnostics, Models, and Financial Implications." *Review of Financial Studies*, 17 (2004), 581–610.

Post, T., and P. Versijp. "Multivariate Tests for Stochastic Dominance Efficiency of a Given Portfolio." *Journal of Financial and Quantitative Analysis*, 42 (2007), 489–515.

Price, K.; B. Price; and T. Nantell. "Variance and Lower Partial Moment Measures of Systematic Risk: Some Analytical and Empirical Results." *Journal of Finance*, 37 (1982), 843–855.

Ramchand, L., and R. Susmel. "Volatility and Cross Correlation across Major Stock Markets." *Journal of Empirical Finance*, 5 (1998), 397–416.

Roy, A. "Safety First and the Holding of Assets." *Econometrica*, 20 (1952), 431–449.

Rubinstein, M. "The Fundamental Theorem of Parameter-Preference Security Valuation." *Journal of Financial and Quantitative Analysis*, 8 (1973), 61–69.

Spitzer, J. "Market Crashes, Market Booms and the Cross-Section of Expected Returns." Working Paper, New York University (2006).

Telser, L. "Safety First and Hedging." *Review of Economic Studies*, 23 (1955), 1–16.

Van Oordt, M., and C. Zhou. "Estimating Systematic Risk under Extremely Adverse Market Conditions." Bank of Canada Staff Working Paper Series No. 2016-22 (2016).