Extreme Downside Risk in Asset Returns

Lerby M. Ergun*

May 2020

Abstract

Does extreme downside risk require a risk premium in the pricing of individual assets? Extreme downside risk is a conditional measure for the co-movement of individual stocks with the market, given that the state of the world is extremely bad. It forms an extension of Ang et al. (2006) downside beta framework. This measure, derived from statistical extreme value theory, is non-parametric. Extreme downside risk is used in double-sorted portfolios, where I control for the five Fama-French and various non-linear asset pricing factors. I find that the average annual excess return between high- and low-exposure stocks is around 3.5%.

Bank Topics: Asset Pricing, Econometric and statistical methods.
JEL codes: C14, G12, G11

---

*London School of Economics and Bank of Canada. I want to thank Aaditya Muthukumaran, André Lucas, Bjorn Jorgensen, Casper de Vries, Dirk Schoenmaker, Emil Siriwardane, Jean-Sébastien Fontaine, John Einmahl, Jon Danielsson, Philipp Hartmann, Sermin Gungor, Stijn van Nieuwerburgh and Xavier Gabaix for the valuable feedback and discussions. I also thank the seminar participants at the Tinbergen Institute, Nova SBE, Bank of Canada, Erasmus University, University College London and the 10th Seminar on Risk, Financial Stability and Banking of the Banco Central do Brasil. I am also grateful to Florian Weigert for sharing the lower tail dependence factor data. I thank the Netherlands Organisation for Scientific Research Mozaiek grant [grant number: 017.005.108] for research funding. Also, the support of the Economic and Social Research Council (ESRC) in funding the Systemic Risk Centre is gratefully acknowledged [grant number ES/K002309/1].

1 Introduction

Returns in financial markets are characterized by extreme movements (Mandelbrot, 1963; Fama, 1963; Jansen and De Vries, 1991). It is in these extreme cases that investors are highly concerned about the performance of their portfolio. The extreme movements of the market are not always reflected equally in all individual stocks. Securities which are more sensitive to these extreme negative shocks are undesirable and therefore should sell at a discount, i.e. fetch a risk premium. In this paper, I propose an extreme downside dependency measure, δ, which captures this risk. This non-parametric measure of tail dependency based on extreme value theory (EVT) offers a new approach for capturing extreme risk in asset prices. I find that investors demand a 3.5% risk premium for investing in a high relative to a low δ portfolio.

Prior literature on extreme downside or disaster risk in asset pricing mainly focuses on theoretical models. Part of this literature includes higher moments to account for tail thickness. Samuelson (1970) as well as Harvey and Siddique (2000) and Dittmar (2002) consider skewness and kurtosis as the higher moments. Others, such as Rietz (1988), partially explain the Mehra and Prescott (1985) equity premium puzzle by introducing an 'extreme' bad state to the Arrow-Debreu paradigm. Barro (2006) extends this idea to investigate the impact of extreme risk on asset pricing facts and welfare costs. He finds, as Rietz does, that the equity risk premium and the risk-free rate puzzle can largely be explained by including an extreme bad state. Gabaix (2012) extends these models by adding time variability of disaster risk. His model is able to rationalize ten asset pricing puzzles, including the equity premium puzzle.

Testing theoretical models of extreme downside risk has proven to be a challenge, as extreme events are only rarely observed. Several papers attempt to overcome this challenge by studying different sources of extreme movements in asset prices. Berkman et al. (2011), Bittlingmayer (1998) and Frey and Kucher (2000) use major political crises as a measure of extreme risk. Amihud and Wohl (2004) and Rigobon and Sack (2005) find a link between the stock market and the second Iraq war.

In this paper I consider a novel approach. This approach employs Huang (1991)'s non-parametric count measure to determine the dependence in the tail between individual stocks and the market portfolio. In essence, the measure counts the number of joint excesses of the market return, $R_{m,t}$, and individual stock return, $R_{i,t}$, conditional on $R_{m,t}$ being stressed at time t.

This captures, in a direct way, the dependence given that the world is in an extremely bad state. This measure is directly related to the "recovery rate" or "resilience" of a stock in Gabaix (2012). In his framework, stocks with high resilience command a low-risk premium relative to low-resilience stocks, leading to a cross-sectional risk premium.

The count measure necessitates the choice of a threshold, v and w, to determine the tail region for the joint excesses of $R_{i,t}$ and $R_{m,t}$, respectively. These thresholds should distinguish the extreme behavior, characterized by a power law, from the commonly observed events. Inspired by Bickel and Sakov (2008), Danielsson et al. (2016) propose a methodology for locating the 'start' of the tail by estimating the optimal number of order statistics for the Hill (1975) estimator. To determine the optimal number of extreme order statistics, they use a horizontal distance measure that minimizes the maximum distance between the empirical and the semi-parametric distribution. These optimal thresholds for $R_{i,t}$ and $R_{m,t}$ are univariately determined, and thus in a direct way the multi-variate extreme area for the dependence measure is constructed.

There are currently other empirical approaches that attempt to measure downside risk. To estimate a change in the probability of a tail event, Kelly and Jiang (2014) estimate the conditional thickness of the tail from the cross-section of returns on traded stocks. The month-by-month tail exponent estimates proxy tail risk in the economy. Although this measures the cross-sectional dispersion in the lower tail, it is an indirect measure of extreme downside risk in the economy. Secondly, the use of the estimator of the tail exponent by Hill (1975) in the cross-section violates a necessary independence assumption. The bias caused by violating the independence assumption possibly proxies other latent factors Need citation.

A second approach in the literature uses the information of deep out of the money (OTM) put options to capture tail risk. This approach utilizes the difference between quadratic variation and integrated variance to isolate the risk of jumps. Santa-Clara and Yan (2010) and Bollerslev and Todorov (2011) infer tail risk from the OTM put options on the S&P 500 Index. Bollerslev and Todorov (2011) use EVT to scale up the risk of medium jumps to large jumps. They find that jump risk and fear of jumps accounts for two-thirds of the equity risk premium. Siriwardane (2015) utilizes the difference between OTM put and call options to isolate jump risk for individual stocks. He then sorts these into portfolios according to their jump risk to create a 'high-minus-low' factor.

A third approach focuses on measuring the non-linear risk-return relationship. Harvey and Siddique (2000) develop a measure of conditional skewness in stock returns. As expected, they find that this measure of higher moment covariation demands a negative risk premium. Ang et al. (2006) propose a non-linear market model. They separate the market beta into a downside and upside beta. They find that the conditional downside beta is differently priced from the upside beta, and therefore argue that their conditional betas provide a better risk profile of a stock. Although these measures focus on the asymmetric nature of returns, they focus on the non-extreme part of the return distribution. These measures employ commonly observed returns, which contaminates the information in the tail region of the return distribution. Therefore, extreme downside risk forms a natural extension to their downside risk framework.

A fourth approach, which is most closely related to the measure proposed in this paper, is to measure the tail dependence between stocks and a market index. Van Oordt and Zhou (2016) use the EVT framework to construct a non-parametric tail beta measure, however they do not find a cross-sectional risk-return relationship. Chabi-Yo et al. (2018) use a convex combination of parametric copulas to measure tail dependence between the return of a stock and the market. With their lower tail dependence (LTD) measure they find a cross-sectional risk premia of 4.32%. Extreme downside risk is similar to their approach, in that δ is a limit copula. However, δ is non-parametric and solely relies on tail observations for its estimation. Even though δ and LTD are moderately positively correlated, the empirical analysis suggests they capture different aspects of the risk-return relationship.

An advantage of the approach offered in this paper is that extreme downside risk is a direct and simple measure of the relationship of the state of the world and the pay-off of the financial asset. It is also not diluted by the observations in the center of the return distribution. As EVT shows, the count measure has predictive value at very high but finite levels. Furthermore, I refrain from using deep OTM options, e.g. as Siriwardane (2015) and Bollerslev and Todorov (2011) do. OTM options can suffer from liquidity issues, especially for individual companies.

To investigate whether investors care about extreme downside risk, I sort stocks by their realized measure of extreme dependence. The difference in annualized realized return between the low and high δ quintile portfolios is about 3.5%. This shows that investors want to be compensated for bearing high extreme downside risk. It is possible that extreme downside risk is a proxy for other existing risk factors. In the empirical asset pricing literature, double-sorted portfolios are employed to control for existing risk factors. When controlling for the five factors by Fama and French (2015), momentum (Carhart, 1997), liquidity (Stambaugh and Lubos, 2003), downside beta (Ang et al., 2006), cross-sectional tail risk (Kelly and Jiang, 2014), coskewness, cokurtosis (Harvey and Siddique, 2000), and lower tail dependence measure (Chabi-Yo et al., 2018), the premium on extreme downside risk remains on average 3% and significant. This result is furthermore robust for excluding financial firms, long-lived firms and variation in δ over time.

The positive premium is in line with the results of Kelly and Jiang (2014), Siriwardane (2015) and Santa-Clara and Yan (2010), who also find higher compensation for downside sensitivity. The risk premium of extreme downside risk is in excess of Ang et al. (2006) downside risk beta. This advocates a further non-linearization of their downside beta framework.

Section 2 introduces the extreme dependence measure and the other non-linear asset pricing factors. This is followed by section 3, which describes the data that are used for the empirical analyses. Section 4 presents and discusses the empirical results from the analyses, followed by the conclusion.

2 Methodology

This section consists of three parts. The first two elaborate on how extreme dependence is measured and how I define the start of the tail. The third part provides an overview of other systematic risk measures brought forth by the literature.

2.1 Extreme dependence measures

Investors are interested in the performance of individual stocks relative to their wealth in a particular state of the world. I examine the asset pricing in the extremely bad states of the world. I am interested in observing extreme negative excess stock return at time t, $R_{i,t}$, conditional on the market excess return, $R_{m,t}$, being extremely negative at time t. To measure this relationship, I employ the following count measure:

$$\delta_i = \frac{\sum_{t=1}^{T} I_{\{R_{i,t}<v, R_{m,t}<w\}}}{\sum_{t=1}^{T} I_{\{R_{m,t}<w\}}}, \tag{1}$$

where I is the indicator function that takes value 1 when $R_{i,t} < v$ and $R_{m,t} < w$, and 0 otherwise. The summation in the numerator counts the number of paired observations that fall in the extreme quadrant, the area where both $R_{i,t}$ and $R_{m,t}$ are extreme. Figure 1 gives an illustration of the extreme quadrant. This measure can be viewed as the conditional probability,

$$P\left(R_{i,t} < v \mid R_{m,t} < w\right) = \frac{P\left(R_{m,t} < w \cap R_{i,t} < v\right)}{P\left(R_{m,t} < w\right)}.$$

This measure is a proxy for the level of dependence a stock has on extreme market risk. For v and w going to infinity, the conditional probability tends to the tail dependence measure presented in Hartmann et al. (2004). A thorough derivation is provided in De Haan and Ferreira (2007). Under the self-similarity property, the tail dependence model can be captured by the count measure described above. The event that the count measure seeks to quantify is the following: Given that the market has an extreme downside event, how likely is it that stock i also exhibits an extreme movement in the same direction?

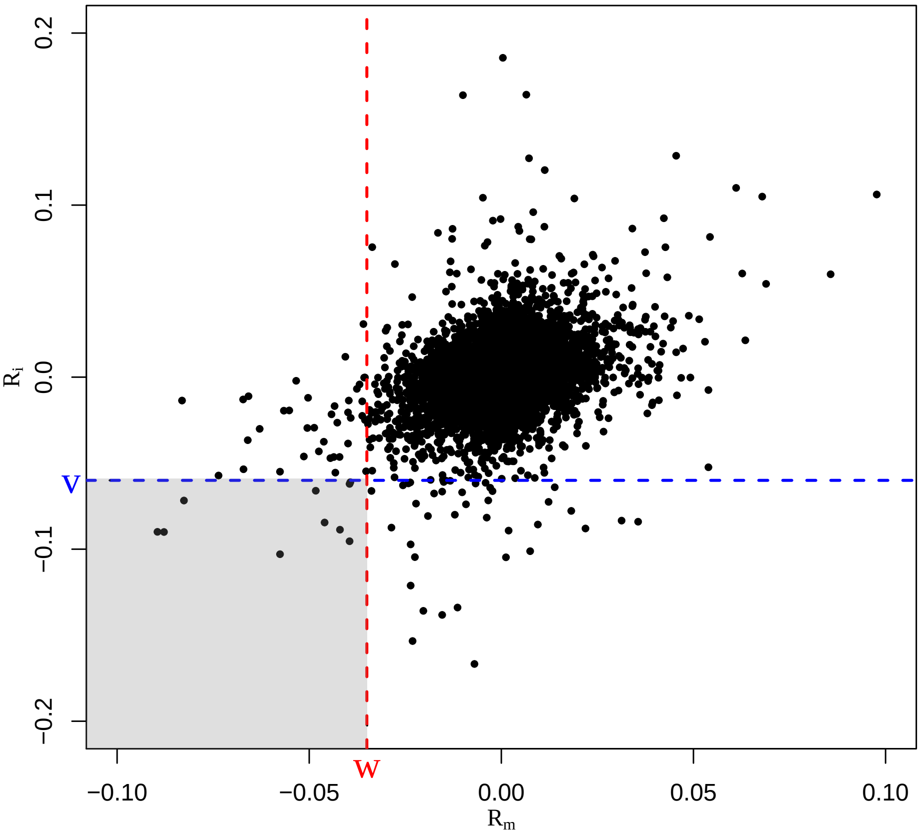
2.2 Where does the tail start?

To measure extreme dependence accurately, it is essential to determine the part of the tail where heavy-tailed behavior holds, i.e. the scaling behavior described by power laws. This paper employs the EVT methodology to locate these points.

In EVT, the $1/\gamma$ in the Pareto distribution, $P(X < x) = 1 - Ax^{-1/\gamma}$, determines the shape of the tail. The power function in the Pareto distribution is often used as an approximation of the tail probability for generic heavy-tailed distributions.[1] In the literature, $1/\gamma$ is often referred to as the tail index. The level of $1/\gamma$ determines how many moments exist and thus how heavy the tail of the distribution is.

[1] Consider the Hall and Welsh (1985) expansion, $1 - F(x) = Ax^{-\alpha}\left[1 + Bx^{-\beta} + o\left(x^{-\beta}\right)\right]$, of a heavy-tailed distribution function. For the Pareto distribution, observe that the Hall expansion perfectly fits the first-order term. All of the standard heavy-tailed distributions, like the Student-t, Pareto, symmetric stable distribution or the unconditional distribution of the stationary solution to a GARCH(1,1) process, satisfy the Hall expansion. Therefore, the Pareto function serves as a good approximation for the tail of most heavy-tailed distributions.

Figure 1: Graphic example δ_i



This graph gives the scatter plot of Allegheny Power Systems Inc. returns and the corresponding market returns. Here w (red) is the illustrative optimal threshold level for the market returns and v (blue) is the illustrative optimal threshold level of stock i for the left tail. The region under v and to the left of w is the extreme quadrant (shaded area) for extreme downside risk.

The most popular estimator for γ is the Hill (1975) estimator,

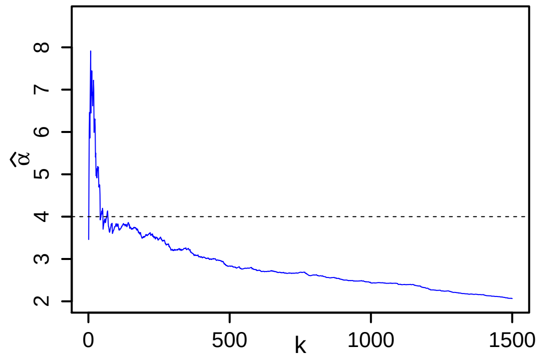
$$\widehat{\gamma} = \frac{1}{k}\sum_{i=1}^{k} \left(\log\left(X_{n-i+1,n}\right) - \log\left(X_{n-k,n}\right)\right), \tag{2}$$

where $X_{n-i+1,n}$ is the i^{th} largest observation (order statistic) out of a sample of size n and k is the number of observations in the tail that are used for estimating γ.[2] As can be seen from Equation (2), one has to choose the nuisance parameter k, which determines how many extreme order statistics

[2]For the left tail the observations are multiplied by -1.

are used in the estimation. Figure 2 shows the change in $1/\hat{\gamma}$ as the number of order statistics included in the estimation increase. To locate k^*, the opti-

Figure 2: Hill plot Student-t(4)



This graph depicts the estimate for $1/\hat{\gamma} = \hat{\alpha}$ for different levels of k. The sample of 10,000 is simulated from a Student-t distribution with 4 degrees of freedom. This graph is often referred to as the Hill plot.

mal number of order statistics for the Hill estimator, Danielsson et al. (2016) introduce a simple method inspired by Bickel and Sakov (2008). Danielsson et al. (2016) use the Kolmogorov-Smirnov metric, but measured in the quantile rather than the probability dimension. The choice of the quantile dimension is motivated by the fact that a probabilistic mistake in the tail of the distribution translates into a disproportionally large quantile mismatch, which is the dimension that economists care about. They furthermore show that this improves the quantile estimates deep in the tail region of the distribution.

In EVT, the Pareto distribution is often utilized to semi-parametrically estimate the extreme quantiles. To fit the tail one only needs estimates for the scale and tail index of the Pareto distribution. Via various simple transformations, Danielsson et al. (2016) arrive at their KS-distance metric,

$$k^* = \underset{k=2,\ldots,K}{\arg\inf} \left[\sup_{j=1,\ldots,K} |X_{n-j,n} - q(j,k)| \right]. \tag{3}$$

The function $q(j,k)$ is the semi-parametric quantile estimate at probability $(n-j)/n$.[3] I limit the area over which the above metric, i.e. KS-distance metric, is measured to $X_{n-K,n} \geq x$. Here $K > k$ is large but is still in the tail.[4] The

[3]See Appendix A.1 for the derivation of the semi-parametric quantile estimator.

[4]For example, 10% of the sample fraction. Danielsson et al. (2016) show that k^* is insensitive to the choice of K, once K is large enough. Alternatively, one can use all the

k that produces the smallest maximum horizontal difference along all the tail observations up to K is chosen as the optimal number of observations to estimate the thickness of the tail. Through the optimal k, I also define the start of the tail.

Here I define k_i^* and k_m^* as the optimal number of order statistics for stock i and the market index, respectively. Once k_i^* and k_m^* are determined, I turn to the multivariate problem of measuring the dependence. From the univariate measures, k_i^* and k_m^*, an extreme dependence region is created, which appears as the shaded area in Figure 1.

The region $(R_{i,t} < v, R_{m,t} < w)$, where v and w correspond to the quantile of the k_i^{*th} and k_m^{*th} order statistics respectively, is appointed as the extreme quadrant. The number of extreme pairs of $R_{i,t}$ and $R_{m,t}$, which fall in this region, relative to the number of extreme market movements, k_m^*, forms the dependence measure in Equation (1).

The measurement of δ_i relies on rarely observed events. To limit the measurement error of extreme downside risk, I use the whole sample of each individual stock to estimate δ_i for the base results. The standard in the empirical asset pricing literature is the use of subsamples (rolling windows) to account for the time variation in systematic risk characteristics of stocks. To account for the possible time variation in δ_i and the introduced look-ahead bias, the analysis is repeated with $\delta_{i,t}$. To estimate $\delta_{i,t}$ we only use data till time t. This alleviates the look-ahead bias and allows for subsample variation in δ_i. These results are presented in the robustness analysis of the empirical results section.

2.3 Non-linear systematic risk

The asset pricing literature suggests several systematic risk factors that capture the asymmetry of the return distribution. In this paper these risk factors are reconstructed and used as control variables.

Scott and Horvath (1980) advocate the inclusion of the sensitivity of higher-order moments of the return distribution into the pricing kernel. Harvey and Siddique (2000) use coskewness as a measure of heavy tails, where coskewness

positive observations in the sample.

is defined as

$$coskewness = \frac{E\left[\varepsilon_i \varepsilon_m^2\right]}{\sqrt{E\left[\varepsilon_i^2\right] E\left[\varepsilon_m^2\right]}}.$$

Here ε_i is the residual from regressing the excess return of stock i on the three Fama and French (1996) factors. The variable ε_m is the demeaned excess return on the market portfolio. I also include a measure of cokurtosis:

$$cokurtosis = \frac{E\left[\varepsilon_i \varepsilon_m^3\right]}{\sqrt{E\left[\varepsilon_i^2\right] E\left[\varepsilon_m^3\right]}}.$$

Although coskewness and cokurtosis are not a direct measure of tail dependence, they focus on asymmetry in the risk-return. However, the estimation of both measures requires the full return distribution.

Dekkers and De Haan (1989) show that with EVT, only the tail observations are necessary to provide information about tail risk. Moreover, using the vast number of center observations in the estimation might create a biased measure of tail dependence. An additional problem with the cokurtosis and coskewness measures is that they need the second and third moment to exist. This is not always the case for financial returns.

Ang et al. (2006) propose a non-linear market beta framework. They separate the co-movement of an individual asset conditional on a down movement and up movement of the market. Given that the market is below its average excess return, a beta is estimated. Accordingly, this is also done for the above-average market excess returns. Given the focus of this paper on downside risk and the mixed results for upside beta in Ang et al. (2006), only downside beta is considered. Define downside beta as

$$\beta^- = \frac{cov\left(R_i, R_m | R_m < \mu_m\right)}{var\left(R_m | R_m < \mu_m\right)},$$

where μ_m is the average excess market return.

Ang et al. (2006) use all observations in their upside and downside beta framework. However, the disaster literature to date points towards a separation for the extremely bad states. That information is lost when using the center observations in the downside beta measure. I utilized the information in the tail observations as a further non-linearization of their risk-return framework. By excluding the tail observations from the upside and downside beta, the tail dependence measure can be estimated using these excluded observations. Therefore, the factor proposed in this paper provides a natural extension of their framework.

Kelly and Jiang (2014) develop an approach to estimate the sensitivity of stocks to changes in probability of extreme negative market drops. They estimate the conditional tail index exponent, i.e. $\hat{\alpha}_t^{cs}$, by exploiting the cross-section of individual stock returns. They use pooled daily returns estimate a monthly time series of tail indexes with the Hill estimator. They find that the sensitivity to $\hat{\alpha}_t^{cs}$ carries a positive risk premium. Furthermore, they show that $\hat{\alpha}_t^{cs}$ can predict the excess market return. However, using the Hill estimator for the cross-section of returns can be problematic. Dependencies in the cross-section cause biased estimates. The variation in $\hat{\alpha}_t^{cs}$ could therefore be driven by this bias and consequently proxy other dependencies in the cross-section.

3 Data

The analysis uses US equity market data from 1963 to 2018. Stock market data are obtained from the Center for Research in Security Prices (CRSP). The CRSP database contains individual stock data from the NYSE, AMEX, NASDAQ and NYSE Arca. The five Fama-French factor (Fama and French, 2015) data are provided by the website of Kenneth R. French, as is the momentum factor by Carhart (1997). The data library contains daily and monthly constructed Fama-French and momentum factors from 1963 to 2018. The liquidity factor by Stambaugh and Lubos (2003) is obtained from the website of Lubos Pastor. The book-to-market ratio, which is used as one of the control variables, is obtained from the Compustat database. The Compustat database contains data from 1950 to 2018 on balance sheet items of the respective companies. The lower-tail dependence (LTD) factor is provide by the authors (Chabi-Yo et al., 2018) and the sample is from 1963 till 2012.

In the main analysis, 19,904 stocks are included. The analysis is confined to the period 1963 to 2018.[5] Only stocks with more than 60 months of data are used, as accuracy of EVT estimators typically requires a large total sample size.[6] Only a small fraction is informative for tail estimation. Table 1 gives the descriptive statistics for extreme downside risk measure.

[5]The CRSP database and Fama-French factors dataset provide information going back to 1926. For a detailed description of the construction of the Fama-French factors and the momentum factor, please visit the data library on the website of Kenneth R. French.

[6]Stocks with exchange code -2, -1 or 0 are not included in the analysis. In addition, only stocks with share code 10 and 11 are included in the analysis.

In the Appendix, Figure 3 provides additional details on the distribution of the sample fractions used for the estimation of the count measure. One can see that the shape of the distribution of the k_m^* and k_i^* are different. The distribution of k_i^* shows that in most of the cases, a sample fraction lower than 5% is chosen. This is also the case for k_m^*, but less frequently. The lower left graph shows the difference between k_m^* and k_i^*. The distribution is centered around zero, but in some cases k_m^* is much larger than k_i^*. I perform additional robustness checks by taking a fixed threshold. Using a fixed threshold of 1% of the sample fraction does not change the size of the average excess return for extreme downside risk, but makes the standard errors larger.

Table 1: Descriptive statistics of extreme downside risk

	# Firms	Mean	St Dev	Min	Max
All	19904	0.17	0.20	0.00	0.95
Agriculture, Forestry and Mining	2971	0.12	0.16	0.00	0.90
Contractors and Construction	296	0.14	0.17	0.00	0.90
Manufacturing	7597	0.15	0.18	0.00	0.95
Transport, Communications and Utilities	1656	0.19	0.20	0.00	0.95
Wholesale Trade	975	0.14	0.18	0.00	0.90
Retail Trade	1318	0.15	0.17	0.00	0.90
Finance, Insurance and Real Estate	5980	0.21	0.23	0.00	0.95
Business and Personal Services	2422	0.15	0.18	0.00	0.90
Health Services	437	0.15	0.19	0.00	0.88
Legal, Education and Social Services	134	0.14	0.19	0.00	0.80
Engineering and Accounting Services	496	0.18	0.22	0.00	0.90
Government (Public Administration)	57	0.17	0.20	0.00	0.81

This table displays the summary statistics of the extreme downside risk measures per industry. The industries are arranged according to SIC codes. The first column reports the number of companies that are included in an industry. Columns 2–5 report the mean, standard deviation, maximum and minimum of the extreme downside risk realizations. The data includes all the securities in the CRSP universe from 1963 to 2018.

4 Empirical results

Common practice in the asset pricing literature is to sort stocks in quintile portfolios based on their factor realizations. Subsequently, the direction of the average realized returns of the quintile portfolios are examined for the predicted relationship.

When investigating the relationship between realized factor loadings and average returns, the results should normally hold for equal- and value-weighted portfolios. As pointed out by Ang et al. (2006), previous work finds that the risk due to asymmetries is bigger among smaller stocks. I therefore follow Ang et al. (2006) and Harvey and Siddique (2000) by focusing on equal-weighted portfolios. In the Appendix, tables 6 and 7 report the results for value-weighted portfolios. The results are mostly in the same direction; however, there is more variation in the size of the premium.

Table 2 presents the results for sorting stocks on their realized δ_i. The portfolios sorted on δ_i show an overall increase in the average realized returns. This direction is in line with the risk-return relationship that one expects. Investors want to be compensated, with a higher average return, for holding stocks that perform extremely badly in the extremely bad states of the world.[7] The difference in the annualized average return between the high- and low-risk portfolio is about 3.5%. Furthermore, Table 2 shows that the average β^m in the sorted quantile portfolios is also increasing. It is possible that δ is a proxy for existing risk factors, like the market factor.

Table 2: Single-sorted portfolios

	Return	Mean (δ)	Mean (β^m)
Low δ	8.07	0.01	0.90
2	8.44	0.05	0.91
3	9.24	0.11	0.94
4	10.26	0.22	1.00
High δ	11.51	0.43	1.07
High – Low	3.43	0.43	0.17
t-stat	[6.59]		[18.48]

This table lists the equal-weighted average excess returns and risk characteristics of stocks sorted on extreme downside risk, δ, realizations. Here δ is calculated using daily observations for every individual asset that is listed on NYSE, AMEX, NASDAQ or NYSE Arca between the years 1963 and 2018. Subsequently, the stocks are sorted into quintile portfolios based on their realized extreme downside risk factor. The columns Return, Mean(δ) and Mean(β^m) report the average annualized monthly excess return, the average extreme downside risk measure and the average market β measure of the stocks in the quintile portfolios, respectively. They are measured with a 5-year rolling window. The row 'High - Low' states the average annualized difference in the average realized return in the High and Low portfolio. The last row presents the t-statistic of the difference with Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

To explicitly control for other factors, the asset pricing literature often uses double-sorted portfolios. Furthermore, the double-sorting procedure can reveal non-linearities in the risk-return relationship. In the procedure, stocks

[7] This rationale applies the Arrow-Debreu state-pricing framework.

are independently sorted on their exposure to an existing risk factor and δ_i. They are subsequently allocated to their appropriately ranked portfolio.[8] Table 3 presents the average realized portfolio returns of the double-sorted portfolios where the market, small-minus-big, high-minus-low, momentum and liquidity factors function as controls.

In panel (a), each portfolio in a row has approximately equal exposure to the market factor. In a row, each portfolio along the columns has an increasing level of δ_i. The realized average return among these portfolios is increasing as well. This is true for all five rows. This shows that given the exposure to market risk, stocks also get additionally compensated for their exposure to extreme downside risk. The column 'H-L' shows the difference in the average excess return between the fifth and first quintile portfolio for given levels of market risk. These excess returns show that compensation for bearing extreme downside risk, controlling for market risk, is between 2.12% and 4.06% annually. These premia have a robust t-statistic of 3.73 or higher. The first and fifth row of panel (a) shows that stocks that are described as carrying low and high systematic market risk carry a more sizable risk premium for having a high level of δ_i. This shows that β^m does not fully characterize the systemic risk profile of these stocks and that δ_i provides valuable differentiating information.

Panels (b) and (c) show the average 'H-L' realized return for δ controlling for the small-minus-big and high-minus-low factors. The average 'H-L' excess returns are between 2.30% and 4.12% and significant. Furthermore, one can also wonder if stocks that have a high extreme dependence are sensitive to liquidity issues. Stambaugh and Lubos (2003) find that stocks with high sensitivity to liquidity have a higher expected return. This is especially a concern when the market experiences an extreme downward movement. Panel (d) shows that controlling for the liquidity factor does not explain the risk premium for extreme downside dependence. Furhermore, controlling for Carhart's (1997) momentum factor, in panel (e), does not significantly influ-

[8] The double-sorting results in this paper are acquired by an independent double-sorting procedure. A concern with the independent double-sort procedure is the sparse number of stocks in some of the sorting portfolios. In contrast to the dependent-sorting procedure, with the independent sort one circumvents the issue of correlated factor loadings. For a detailed description of the double-sorting procedure used in this paper, see Appendix A.2. The results with a dependent-sorting procedure are quantitatively similar. In the Appendix, Table 13 presents the cross-sectional correlation between the factor loadings for each systemic risk factor. The market factor shares the highest correlation with δ, but this correlation is relatively low. Therefore, it is unlikely that δ proxies for one of the other factors.

Table 3: Double-sorted portfolios

	Low δ	2	3	4	High δ	H-L	t-stat
Low β^m	5.34	5.92	6.99	8.28	9.40	**4.06**	[8.58]
2	7.24	7.79	7.91	8.49	9.55	**2.31**	[4.52]
3	8.24	8.59	8.56	9.50	10.55	**2.31**	[4.32]
4	9.43	9.42	10.10	10.75	11.55	**2.12**	[3.73]
High β^m	11.43	11.70	13.28	13.44	14.40	**2.98**	[4.29]
panel (a)							
Low β^{SMB}	7.22	7.76	8.83	9.36	10.79	**3.57**	[5.33]
2	7.65	8.38	8.67	9.61	10.70	**3.04**	[6.03]
3	8.17	8.99	9.01	9.85	10.92	**2.74**	[5.63]
4	8.54	8.71	9.29	10.04	11.69	**3.16**	[6.16]
High β^{SMB}	9.13	8.62	10.24	12.07	13.25	**4.12**	[7.21]
panel (b)							
Low β^{HML}	11.49	11.83	12.87	13.62	15.57	**4.08**	[5.43]
2	8.18	8.57	9.49	10.35	11.36	**3.18**	[6.21]
3	7.18	8.05	8.25	9.37	10.04	**2.86**	[6.18]
4	7.35	7.67	8.01	8.67	9.65	**2.30**	[4.90]
High β^{HML}	6.52	6.67	7.80	8.86	9.95	**3.43**	[5.37]
panel (c)							
Low β^{Liq}	8.97	9.14	10.18	11.75	13.00	**4.04**	[6.16]
2	8.16	8.51	8.99	9.83	11.03	**2.88**	[5.16]
3	7.81	8.35	8.91	9.25	10.22	**2.41**	[5.85]
4	7.75	8.18	8.82	9.40	10.52	**2.77**	[5.68]
High β^{Liq}	7.86	8.19	9.33	10.89	12.42	**4.56**	[6.61]
panel (d)							
Low β^{Mom}	5.10	5.71	6.64	7.34	8.76	**3.66**	[5.10]
2	6.38	6.86	7.73	8.64	8.89	**2.50**	[4.68]
3	7.76	8.07	8.65	9.80	10.65	**2.89**	[5.67]
4	9.51	9.62	10.10	10.95	12.22	**2.71**	[5.59]
High β^{Mom}	11.83	11.93	12.92	14.50	16.49	**4.66**	[8.05]
panel (e)							

This table lists the equal-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^m is the market beta. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996) respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ_i, is estimated over the whole sample period of an asset. I use the lowest 1% of the market excess return to estimate δ_i and the remaining observations to estimate the control risk-factor exposures. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ_i portfolios. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

ence the previously found results. This implies that δ is not likely functioning as a proxy for other tested risk factors.[9]

Given that extreme downside risk is focused on the asymmetric risk-return relationship, one needs to control for factors that attempt the same. Table 4 displays the average realized returns for the double-sorted portfolios with the extreme downside risk and other alternative downside risk measures. To integrate δ with Ang et al. (2006) framework, the months corresponding to the lowest 1% market excess return observations are used to measure δ_i. The remaining observations are used to estimate the alternative downside risk measures, which are used as control variables. In this manner, δ_i forms a natural extension of downside beta.

Panel (a) controls for downside beta. The average excess returns for 'H-L' portfolios are positive and significant at a 1% confidence level. The marginally lower t-statistics alludes to δ_i being more crudely measured by taking a fixed threshold at 1% of the sample.[10] As in Ang et al. (2006), I also control for the exposure to downside beta relative to the market beta. The average excess returns, in panel (b), are positive and significant at a 1% confidence level. This shows that the incremental information that is contained in β^- over the unconditional β^m is different from the information contained in δ_i.

By taking a higher-order Taylor approximation of an investor's utility function, Scott and Horvath (1980) show that the higher-order moments in the return distribution play a role in describing a stock's systematic risk characteristics. After controlling for coskewness in panel (c), the average excess return on the 'H-L' portfolios for extreme downside risk is still positive and significant. Due to the use of the numerous central observations, coskewness lacks the focus to capture the information contained in the very extreme observations. As expected, in panel (d) LTD shows an increase in the average return along the rows. Along the columns, δ displays a risk premium for each level of LTD exposure. The premium is highest for low LTD stocks, demonstrating the differentiating information contained in δ. I also control

[9] In Table 10 in the Appendix, I also control for the conservative-minus-aggressive and robust-minus-weak factors by Fama and French (2015). Kelly and Jiang (2014) cross-sectional tail risk measure is included in the analysis as well. Additionally, I control for the book-to-market ratio and the variance of the asset. The average realized 'H-L' excess returns for δ are approximately the same size and significant.

[10] The results where the extreme quadrant is determined with the KS-distance are quantitatively similar.

Table 4: Double-sorted portfolios – non-linear factors

	Low δ	2	3	4	High δ	H-L	t-stat
Low β^-	5.14	5.70	6.32	7.21	8.96	**3.82**	[5.72]
2	7.40	7.32	7.74	8.42	8.85	**1.45**	[2.76]
3	8.25	9.05	9.02	9.10	9.52	**1.26**	[2.19]
4	9.29	10.02	10.32	10.44	11.07	**1.78**	[2.68]
High β^-	13.26	12.66	13.54	14.23	15.61	**2.36**	[3.97]
panel (a)							
Low $\beta^- - \beta^m$	7.35	8.36	8.47	9.64	10.12	**2.76**	[4.35]
2	7.59	8.47	8.77	9.51	10.52	**2.94**	[5.64]
3	7.79	8.87	9.36	9.74	11.19	**3.40**	[6.55]
4	8.80	9.00	9.97	10.65	12.32	**3.53**	[6.58]
High $\beta^- - \beta^m$	9.11	8.01	9.68	11.54	13.08	**3.97**	[6.11]
panel (b)							
Low Coskewness	8.51	8.27	9.32	10.53	11.55	**3.04**	[6.02]
2	7.98	8.20	9.59	10.61	11.56	**3.58**	[7.80]
3	8.07	8.48	8.99	10.34	11.47	**3.40**	[5.83]
4	7.64	8.39	8.69	9.84	11.23	**3.59**	[5.84]
High Coskewness	8.07	8.96	9.77	9.72	11.34	**3.27**	[4.91]
panel (c)							
Low LTD	5.45	5.68	7.13	8.72	10.13	**4.68**	[6.03]
2	6.53	7.26	7.80	8.58	10.04	**3.51**	[6.13]
3	7.67	7.96	8.60	9.18	10.10	**2.42**	[5.03]
4	8.24	8.66	8.80	9.97	10.56	**2.31**	[5.90]
High LTD	10.66	10.68	11.09	11.45	13.28	**2.62**	[4.07]
panel (d)							
Low α^{cs}	8.92	9.11	9.44	11.41	12.98	**4.07**	[6.93]
2	8.17	8.37	8.72	9.63	10.81	**2.64**	[5.62]
3	7.71	8.19	8.46	9.31	10.19	**2.48**	[5.74]
4	8.11	8.05	8.63	9.56	10.44	**2.33**	[4.93]
High α^{cs}	7.88	8.79	10.86	10.84	12.50	**4.62**	[6.66]
panel (e)							

This table lists the equal-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^- is the downside beta by Ang et al. (2006). *coskewness* is the risk factor, by Harvey and Siddique (2000), which measure third moment co-movement. Furthermore, α^{cs} is Kelly and Jiang (2014) cross-sectional tail risk measure and LTD is the lower tail dependence measure by Chabi-Yo et al. (2018). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ_i, is estimated over the whole sample period of an asset. I use the lowest 1% of the market excess return to estimate δ_i and the remaining observations to estimate the control risk-factor exposures. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ portfolios. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

for Kelly and Jiang (2014) cross-sectional tail index, α^{cs} in panel (e). There is a positive risk premium for δ over α^{cs}. That δ and α^{cs} capture different aspects of the risk-return relationship can also be deduced from the low correlation in the factor loading between α^{cs} and δ_i in Table 13.

Fama-MacBeth regressions

The difference in the average return in the high δ_i and low δ_i portfolios does not directly convey the risk premia for extreme downside risk for the average stock. Furthermore, double-sorting stocks into portfolios does not allow one to control for multiple factors simultaneously. To address these issues, I use Fama and MacBeth (1973) regressions at the firm level.

For the Fama and MacBeth (1973) regressions I construct a factor based on the δ_i measures. At the start of each year I sort all stocks with a dollar value higher than 5 dollars, based on their measured δ_i. I construct a high δ_i quintile portfolio and a low δ_i quintile portfolio. For each month in that year the average return of the high quintile portfolio is subtracted from the average return of the low quintile portfolio. This time series of 'H-L' monthly returns forms the F^{δ} factor. Subsequently, the factor is used to estimate the factor loadings to extreme downside risk for each stock.

The first two columns in Table 5 display the price of risk for the existing factors pushed forward in the asset pricing literature. These two models serve as a benchmark for the price of risk of F^{δ}. Column 3 displays that the risk premium for a unit exposure to F^{δ} is about 1.99% annually. Given the 1.65 cross-sectional standard deviation in $\beta^{F^{\delta}}$, a one standard deviation shift in exposure leads to a 3.5% increase in risk premium.

The risk premium for F^{δ} is robust to the inclusion of other existing factors, as shown in columns 4 to 7.[11] Of the factors focused on the asymmetry in the risk-return relationship, in column 2 only LTD has a significant premium. After including F^{δ}, the risk premium diminishes and becomes insignificant. This can be attributed to the simplicity of δ and the exclusive use of tail observations to cleanly capture the non-linear tail relationship.

The last row in Table 5 reports the average R^2 for each model. The exposure to F^{δ} explains about 6% of the cross-sectional variation in the average re-

[11]Cokurtosis is excluded from the Fama-MacBeth regressions due to near perfect correlation with coskewness in various subsamples.

Table 5: Fama-MacBeth regressions

	1	2	3	4	5	6	7
Constant	4.17	4.41	6.32	5.13	5.48	5.10	4.50
	(0)***	(0)***	(0)***	(0)***	(0)***	(0)***	(0)***
β^m	5.10	5.31		4.79	4.37	4.98	5.20
	(0)***	(0)***		(0)***	(0)***	(0)***	(0)***
β^{SMB}		1.22			1.03	1.30	1.28
		(0.02)**			(0.05)**	(0.01)***	(0.02)**
β^{HML}		-2.87			-2.68	-2.62	-2.89
		(0)***			(0)***	(0)***	(0)***
β^{F^δ}			1.99	2.11	1.99	2.01	2.00
			(0)***	(0)***	(0)***	(0)***	(0)***
β^{Mom}		6.08				5.96	6.33
		(0)***				(0)***	(0)***
β^{Liq}		2.14				1.69	1.84
		(0)***				(0.02)**	(0)***
$\beta^{\alpha^{cs}}$		1.24					1.45
		(0.7)					(0.59)
β^{LTD}		0.89					0.34
		(0)***					(0.23)
Coskewness		-9.76					-0.97
		(0.51)					(0.09)*
R^2	7.44	18.65	6.17	10.71	13.86	17.55	19.57

This table shows the results of the second stage of the Fama and MacBeth (1973) regressions of 60-month excess returns on realized-risk characteristics. An overlapping 60-month rolling window is employed on assets that are listed on the NYSE, AMEX, NASDAQ or NYSE Arca. β^m is the market beta. β^{F^δ} is the beta based on the factor of extreme downside risk, δ. The factor is constructed by subtracting the average return of the highest 20% minus the lowest 20% δ stocks. The portfolios are reconstructed at the start of each year. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996), respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). *coskewness* is the risk factor, by Harvey and Siddique (2000), which measures third moment co-movement. Furthermore, α^{cs} is Kelly and Jiang (2014) cross-sectional tail risk measure and LTD is the lower-tail dependence measure by Chabi-Yo et al. (2018). The sample period is from 1963 to 2018. The p-values for the overlapping Fama-MacBeth regression are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors. The last row reports the average R^2 over the rolling windows of the second-stage regressions. *, **, *** are indicators for the significance level at 10%, 5% and 1% of the coefficients, respectively.

turn. This is 1% below the R^2 for the market model. Including both factors increases the R^2 relative to the market model by 3.3%. There is also an increase in the R^2 when comparing models 2 and 7. This implies that F^δ is a valuable addition in explaining the cross-section of the average return on stocks.

5 Robustness analyses

The results in Table 3 indicate that on average a positive premium for extreme downside risk is demanded. A possibility is that the sensitivity to the exposure is not equal among all firms. Table 8 shows the average excess returns for stocks that are listed for longer than 180 months on one of the stock exchanges.[12] These stocks have a slightly lower premium for extreme downside dependence. This is an indication that the exposure to the extreme dependence risk is important for firms that are relatively young.

Financial firms, such as banks, are highly leveraged. The high leverage might mean these firms are very sensitive during market turmoil. Table 1 shows that for the financial sector, the average δ is the highest among the different sectors. This could imply that the results are solely driven by financial firms. In Table 9 I exclude financial firms and find that the results of the average excess returns are comparable in size and significance.

In the analysis, δ is measured over the whole sample period. This implies that the firm's average δ is used in the analysis. To account for the possible variability in a firm's δ over time, in the rolling window analysis I use returns till time t to estimate δ_t. Table 11 reports the average excess return on the 'H-L' portfolio. The excess returns are positive, significant and are on average 2% to 2.5%. This is lower than the base results. This is also reflected in the estimation of the risk premia via the Fama and MacBeth (1973) procedure, displayed in Table 12. The estimated risk premium is approximately 0.65% and mostly significant. When including all non-linear risk factors, the need for long time-series to measure δ is apparent. The risk premium for LTD increases and remains significant, where the risk premium for δ diminishes and becomes insignificant.

6 Conclusion

The dependence of a stock on the extreme movements of the market is an essential part of understanding the compensation investors demand for bearing risk. In these infrequent and extreme cases, investors care most about the performance of their own portfolios. In this paper, a measure for the dependence of stock returns on the extreme downward movements of the market is created. This measure is derived from statistical extreme value theory.

[12] Here, time of being listed on one of the exchanges functions as a proxy for the maturity of the firm.

The measure of extreme downside risk is subsequently used in a portfolio double sort and Fama and MacBeth (1973) regressions to explain the cross-section of average excess returns. This reveals whether investors care about this extreme dependence on top of other risk factors and whether extreme risk fetches a premium or a discount. The results from the cross-sectional analysis show that extreme downside risk carries a premium, as one would expect. The difference in the average realized return between a quintile portfolio that has high extreme downside risk and one that has low extreme downside risk is around 3.5% per annum and remains significant in various robustness checks. These results are in line with the literature to date. The disaster/jump risk literature finds that investors indeed require a premium for stocks that have high returns when tail risk is high. Ang et al. (2006) find in their downside beta framework that there is a premium for the downside beta in excess to the market factor. In line with the bisection of the beta model, I proposed a further division of the risk sensitivity of the market model. I find that adding extreme downside risk to Ang et al. (2006)'s downside beta forms a significant and natural extension to their framework.

References

Amihud, Y., Wohl, A., 2004. Political news and stock prices: The case of Saddam Hussein contracts. Journal of Banking and Finance 28, 1185–1200.

Ang, A., Chen, J., Xing, Y., 2006. Downside risk. Review of Financial Studies 19, 1191–1239.

Barro, R. J., 2006. Rare disasters and asset markets in the twentieth century. The Quarterly Journal of Economics 121, 823–866.

Berkman, H., Jacobsen, B., Lee, J. B., 2011. Time-varying rare disaster risk and stock returns. Journal of Financial Economics 101, 313–332.

Bickel, P. J., Sakov, A., 2008. On the choice of m in the m out of n bootstrap and confidence bounds for extrema. Statistica Sinica 18, 967–985.

Bittlingmayer, G., 1998. Output, stock volatility, and political uncertainty in a natural experiment: Germany, 1880–1940. The Journal of Finance 53, 2243–2257.

Bollerslev, T., Todorov, V., 2011. Tails, fears, and risk premia. The Journal of Finance 66, 2165–2211.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of Finance 52, 57–82.

Chabi-Yo, F., Ruenzi, S., Weigert, F., 2018. Crash sensitivity and the cross section of expected stock returns. Journal of Financial and Quantitative Analysis 53, 1059–1100.

Danielsson, J., Ergun, L. M., De Haan, L., De Vries, C., 2016. Tail index estimation: Quantile driven threshold selection. BoC Staff Working Paper .

De Haan, L., Ferreira, A., 2007. Extreme value theory: An introduction. Springer, New York.

Dekkers, A. L., De Haan, L., 1989. On the estimation of the extreme-value index and large quantile estimation. The Annals of Statistics 17, 1795–1832.

Dittmar, R. F., 2002. Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. The Journal of Finance 57, 369–403.

Fama, E. F., 1963. Mandelbrot and the stable paretian hypothesis. The Journal of Business 36, 420–429.

Fama, E. F., French, K. R., 1996. Multifactor explanations of asset pricing anomalies. The Journal of Finance 51, 55–84.

Fama, E. F., French, K. R., 2015. A five-factor asset pricing model. Journal of Financial Economics 116, 1–22.

Fama, E. F., MacBeth, J. D., 1973. Risk, return, and equilibrium: Empirical tests. Journal of Political Economy 81, 607–636.

Frey, B. S., Kucher, M., 2000. History as reflected in capital markets: The case of World War II. The Journal of Economic History 60, 468–496.

Gabaix, X., 2012. Variable rare disasters: An exactly solved framework for ten puzzles in macro-finance. The Quarterly Journal of Economics 127, 645–700.

Hall, P., Welsh, A., 1985. Adaptive estimates of parameters of regular variation. The Annals of Statistics 13, 331–341.

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

Harvey, C. R., Siddique, A., 2000. Conditional skewness in asset pricing tests. The Journal of Finance 55, 1263–1295.

Hill, B. M., 1975. A simple general approach to inference about the tail of a distribution. The Annals of Statistics 3, 1163–1174.

Huang, X., 1991. Statistics of bivariate extreme values. Thesis Publishers, Rotterdam.

Jansen, D. W., De Vries, C. G., 1991. On the frequency of large stock returns: Putting booms and busts into perspective. The Review of Economics and Statistics 73, 18–24.

Kelly, B., Jiang, H., 2014. Tail risk and asset prices. Review of Financial Studies 27, 2841–2871.

Mandelbrot, B., 1963. The variation of certain speculative prices. The Journal of Business 36, 394–419.

Mehra, R., Prescott, E. C., 1985. The equity premium: A puzzle. Journal of Monetary Economics 15, 145–161.

Newey, W. K., West, K. D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708.

Rietz, T. A., 1988. The equity risk premium: A solution. Journal of Monetary Economics 22, 117–131.

Rigobon, R., Sack, B., 2005. The effects of war risk on US financial markets. Journal of Banking and Finance 29, 1769–1789.

Samuelson, P. A., 1970. The fundamental approximation theorem of portfolio analysis in terms of means, variances, and higher moments. Review of Economic Studies 37, 537–542.

Santa-Clara, P., Yan, S., 2010. Crashes, volatility, and the equity premium: Lessons from S&P 500 options. The Review of Economics and Statistics 92, 435–451.

Scott, R. C., Horvath, P. A., 1980. On the direction of preference for moments of higher order than the variance. The Journal of Finance 35, 915–919.

Siriwardane, E., 2015. The probability of rare disasters: Estimation and implications. Harvard Business School Finance Working Paper .

Stambaugh, R. F., Lubos, P., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642–685.

Van Oordt, M. R., Zhou, C., 2016. Systematic tail risk. Journal of Financial and Quantitative Analysis 51, 685–705.

Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. Journal of the American Statistical Association 73, 812–815.

A Appendix

A.1 Quantile estimator

The starting point for the quantile estimators is the first-order term in (4):

$$1 - \mathrm{F}(x) = Ax^{-1/\gamma}\left[1 + o\left(x^{-\alpha}\right)\right]. \tag{4}$$

This function is identical to a Pareto distribution if the higher-order terms are ignored. By inverting (4), one gets the quantile function

$$x \approx \left[\frac{\mathrm{P}(X \geq x)}{A}\right]^{-1/\alpha}. \tag{5}$$

To turn the quantile function into an estimator, the empirical probability j/n is substituted for $\mathrm{P}(X \geq x)$. A is replaced with the Weissman (1978) estimator $\frac{k}{n}(X_{n-k+1,n})^{\alpha}$, and α is estimated by the Hill estimator.[13] The quantile is thus estimated by

$$q(j,k) = X_{n-k+1,n}\left(\frac{k}{j}\right)^{1/\widehat{\alpha}_k}. \tag{6}$$

Here j indicates that the quantile estimate is measured at probability $(n-j)/n$.

[13]The estimate of A is obtained by inverting $P = Ax^{-\alpha}$ at threshold $X_{n-k+1,n}$.

A.2 Double-sorted portfolio

To explicitly control for other factors, the asset pricing literature often uses double-sorted portfolios. For time t, in months, we have excess return $r_{t,i}$ for stock i, as the return in excess of the risk-free rate. To construct the double-sorted portfolios, we first estimate the loading for factor F^j, $\beta_{i,t}^{Fj}$, for the n_t stocks. I use returns from time t to $t+60$ to estimate:

$$r_{t,i} = \beta_{i,\tau}^{c} + \sum\nolimits_{j=1}^{J} \beta_{i,\tau}^{Fj} F_t^j + \varepsilon_{i,t}. \tag{7}$$

Here, J is the total number of factors included in the regression for stock i. Given that I repeat the to-be-described procedure for different time windows, I use τ to indicate the particular window. In the basic setup, the factors included are the market, SMB, HML, momentum and liquidity factors. To carry out the double sorting, we also need to estimate the extreme downside risk, δ_i, for each stock. To estimate δ_i we use the whole available sample,

$$\delta_i = \frac{\sum\nolimits_{t^*=1}^{T_i} I_{\left\{r_{i,t^*}<v, R_{m,t^*}<w\right\}}}{\sum\nolimits_{t^*=1}^{T_i} I_{\left\{r_{m,t^*}<w\right\}}}, \tag{8}$$

where t^* are daily returns. Here, T_i is the total sample size of stock i, and I is the indicator function. Furthermore, r_{m,t^*} is the excess return on the market portfolio.

Dependent sorting

Given the loadings $\hat{\beta}_{i,\tau}^{Fj}$ and $\hat{\delta}_i$, the stocks are first sorted by $\hat{\beta}_{i,\tau}^{Fj}$ and put into quintile portfolios. Stocks with rank smaller than $^{n_t}/_5$ are allocated to the first quantile portfolio. Within each quintile portfolio, I subsequently sort on $\hat{\delta}_i$ and allocate an equal number of stocks to five sub-quintile portfolios. This totals 25 $P^{b,d}$ portfolios. Here b is the portfolio rank for $\hat{\beta}_{\tau}^{Fj}$ and d for $\hat{\delta}$, ranging from 1 to 5 in both cases. For each $P^{b,d}$ the 5-year average excess return, $\sum\limits_{i=t}^{t+59} r_t^{b,d}$, is calculated. For a fixed b we subtract the average excess return of $P^{b,5}$ from $P^{b,1}$, giving $E[r_t^{H-L}]_b$. This shows for a given level of $\hat{\beta}_{\tau}^{Fj}$ the premium of holding high $\hat{\delta}$ stocks relative to low $\hat{\delta}$ stocks.

The dependent sorting has a disadvantage. If the loadings $\hat{\beta}_{\tau}^{Fj}$ and $\hat{\delta}_i$ are correlated, then the average $\hat{\delta}$ for given d over different b portfolios might differ. This makes it difficult to interpret the different $E[r_t^{H-L}]_b$. To circumvent this problem one can independently sort the stocks.

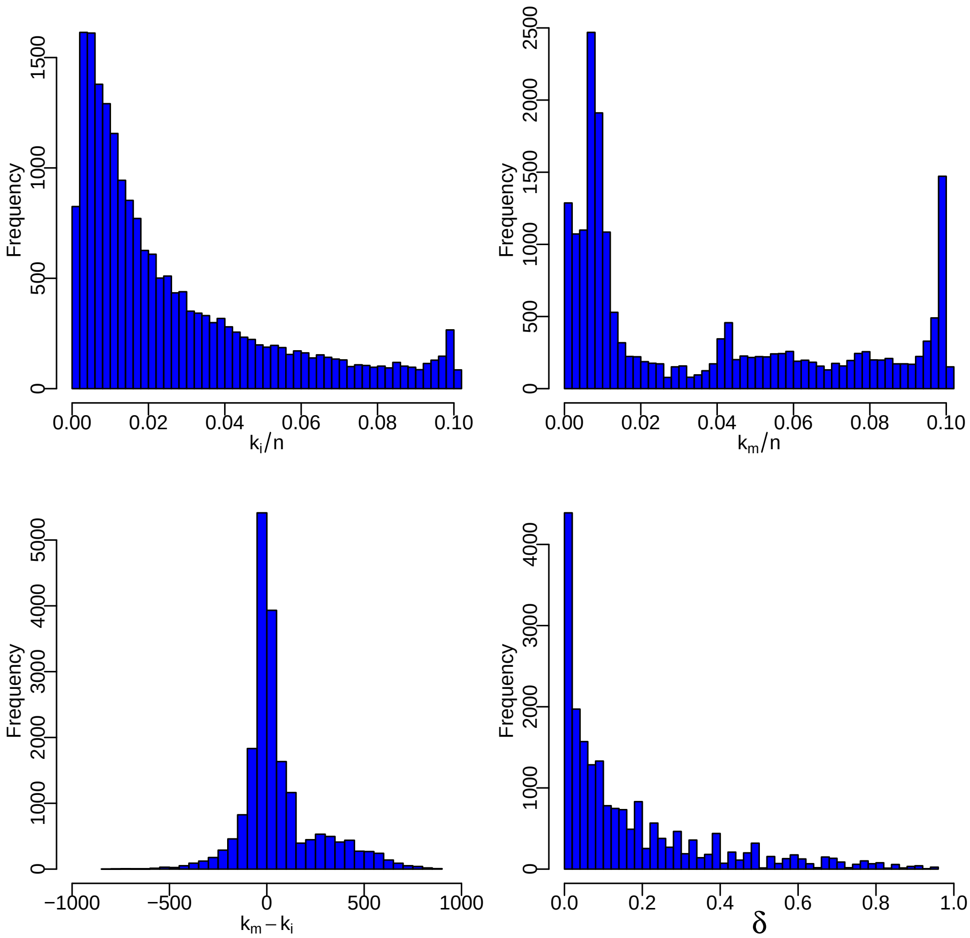
Independent sorting

For the independent sort, stocks receive a rank for both factors independently. This is in contrast to the sequential sorting in the dependent sort. Given the rank for both factors, stocks with ranks for $\hat{\beta}_{i,\tau}^{Fj}$ and $\hat{\delta}_i$ smaller than $n_t/5$ are allocated to portfolio $P^{1,1}$. Stocks with rank $\hat{\beta}_{i,\tau}^{Fj}$ smaller than $n_t/5$ and a rank for $\hat{\delta}_i$ between $n_t/5$ and $2n_t/5$ are allocated to $P^{1,2}$, and so on. This guarantees that the average $\hat{\beta}_{i,\tau}^{Fj}$ is equal across the different d portfolios for a given b and that the average $\hat{\delta}$ is equal across the different b portfolio for given d. The disadvantage is that in some $P^{b,d}$ portfolios the number of stocks can be limited.

Given either form of sorting, this procedure is repeated for $\tau = 1, ..., T - 60$. The $E\left[r_t^p\right]$ and $E\left[r_t^{\text{‘}H-L'}\right]_b$ are averaged over the different time periods and reported. The time series of $E\left[r_t^{\text{‘H-L’}}\right]_b$ are utilized to estimate the Newey and West (1987) heteroskedastic and autocorrelation robust t-statistics for the $E\left[r_t^{\text{‘}H-L'}\right]_b$ excess returns, which are reported in the last column of the tables.

A.3 Tables and figures

Figure 3: Characteristics δ_i



These graphs depict the distribution of the different characteristics of the extreme downside risk measure. The upper-left picture depicts the sample fraction of the total data used to define the extreme negative region of the stock. On the right you see this for the market. The lower-left picture depicts the difference in the number of observations applied in the count measure for the market and stock i. The picture on the bottom right gives the distribution of the extreme downside risk measure.

Table 6: Double-sorted value-weighted portfolios

	Low δ	2	3	4	High δ	H-L	t-stat
Low β^{m}	5.77	5.96	5.72	5.36	6.54	**0.77**	[1.18]
2	6.67	7.09	6.23	6.65	7.12	**0.45**	[0.98]
3	7.18	6.33	6.71	6.75	7.85	**0.68**	[1.19]
4	5.91	5.36	7.44	6.84	7.36	**1.45**	[3.31]
High β^{m}	4.34	4.66	7.16	7.64	7.92	**3.58**	[4.34]
			panel (a)				
Low β^{SMB}	6.83	5.10	7.11	6.96	7.42	**0.58**	[1.42]
2	6.06	6.40	6.77	6.68	7.06	**1.00**	[2.51]
3	5.77	7.12	6.13	6.78	6.67	**0.90**	[1.63]
4	4.72	6.49	6.78	6.02	7.55	**2.83**	[5.02]
High β^{SMB}	4.13	3.94	5.76	6.40	8.48	**4.35**	[7.07]
			panel (b)				
Low β^{HML}	8.58	8.34	8.75	9.44	9.90	**1.31**	[1.23]
2	6.41	6.31	8.03	7.25	7.17	**0.76**	[1.66]
3	5.87	5.50	5.79	6.15	6.45	**0.58**	[1.31]
4	5.85	6.11	5.96	6.09	6.71	**0.85**	[1.91]
High β^{HML}	3.72	4.25	5.94	5.65	6.93	**3.21**	[4.66]
			panel (c)				
Low β^{Liq}	6.28	6.39	6.32	7.15	8.42	**2.14**	[2.71]
2	6.59	5.16	6.54	7.04	7.95	**1.36**	[2.97]
3	6.05	5.96	7.25	6.62	7.49	**1.44**	[2.92]
4	7.28	6.36	6.48	6.92	6.73	**-0.55**	[-1.27]
High β^{Liq}	3.75	6.11	8.22	5.98	6.89	**3.14**	[4.10]
			panel (d)				
Low β^{Mom}	0.80	2.01	3.26	3.44	2.55	**1.75**	[2.35]
2	4.76	3.06	4.83	5.10	5.50	**0.74**	[1.33]
3	7.12	5.88	6.65	6.53	7.68	**0.55**	[1.17]
4	8.66	8.88	8.46	8.81	9.17	**0.50**	[0.83]
High β^{Mom}	10.29	11.15	11.68	11.06	12.77	**2.47**	[3.77]
			panel (e)				

This table lists the value-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^{m} is the market beta. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996) respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ, is estimated over the whole sample period of an asset. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ portfolios. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

Table 7: Non-linear factors – value-weighted portfolios

	Low δ	2	3	4	High δ	H-L	t-stat
Low β^-	4.99	4.93	5.24	5.23	6.11	**1.11**	[1.63]
2	5.66	5.41	5.36	5.77	6.81	**1.15**	[2.07]
3	4.43	5.84	6.27	6.62	7.22	**2.79**	[5.03]
4	7.84	6.58	8.58	7.98	8.15	**0.31**	[0.75]
High β^-	8.09	9.18	9.58	9.97	10.78	**2.70**	[3.32]
panel (a)							
Low $\beta^- - \beta^m$	5.77	5.76	6.58	7.00	6.73	**0.95**	[2.13]
2	5.92	4.65	6.32	6.37	7.02	**1.10**	[2.63]
3	5.55	5.02	6.28	6.24	7.29	**1.75**	[3.75]
4	6.19	6.69	6.49	6.57	8.35	**2.16**	[3.75]
High $\beta^- - \beta^m$	7.25	6.08	6.42	7.17	8.56	**1.31**	[1.45]
panel (b)							
Low Coskewness	6.76	6.94	7.27	7.08	7.32	**0.56**	[1.08]
2	5.55	6.30	7.16	6.83	7.89	**2.34**	[4.46]
3	5.63	5.92	6.36	6.52	7.65	**2.01**	[4.23]
4	5.76	5.17	6.42	6.92	7.37	**1.61**	[2.98]
High Coskewness	6.55	5.49	6.78	6.96	7.01	**0.46**	[0.72]
panel (c)							
Low LTD	1.96	3.08	4.75	3.44	5.25	**3.29**	[2.88]
2	5.48	3.83	5.91	5.69	6.38	**0.89**	[1.48]
3	6.15	4.92	5.70	6.27	6.51	**0.36**	[0.85]
4	6.16	6.18	6.16	6.90	7.09	**0.93**	[1.98]
High LTD	6.77	6.50	8.07	6.26	8.14	**1.37**	[2.07]
panel (d)							
Low α^{cs}	4.88	6.53	6.92	7.25	8.75	**3.87**	[5.07]
2	7.19	6.67	7.29	6.70	7.69	**0.51**	[0.94]
3	6.63	6.14	7.18	7.04	7.35	**0.72**	[1.54]
4	5.91	4.96	6.63	6.55	6.53	**0.62**	[1.44]
High α^{cs}	5.07	4.08	5.41	6.50	7.49	**2.42**	[3.00]
panel (e)							

This table lists the value-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^m is the market beta. β^- is the downside beta by Ang et al. (2006).*coskewness* is the risk factor, by Harvey and Siddique (2000), which measure third moment co-movement. Furthermore, α^{cs} is Kelly and Jiang (2014) cross-sectional tail risk measure and LTD is the Lower tail dependence measure by Chabi-Yo et al. (2018). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ, is estimated over the whole sample period of an asset. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ portfolios. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

Table 8: Double-sorted portfolios – traded more than 15 years

	Low δ	2	3	4	High δ	H-L	t-stat
Low β^{m}	6.16	6.71	7.35	8.70	9.81	**3.66**	[7.01]
2	7.70	8.32	8.20	8.62	9.70	**2.00**	[4.06]
3	8.78	9.27	8.91	9.60	10.85	**2.07**	[3.92]
4	10.29	10.05	10.53	10.94	11.42	**1.13**	[2.16]
High β^{m}	12.24	12.92	13.75	14.10	14.43	**2.19**	[3.26]
			panel (a)				
Low β^{SMB}	8.05	8.42	9.06	9.28	10.73	**2.68**	[4.39]
2	8.22	8.76	8.83	9.65	10.58	**2.37**	[4.98]
3	8.58	9.58	9.40	9.92	10.79	**2.21**	[4.28]
4	9.16	9.44	9.75	10.46	11.92	**2.77**	[5.85]
High β^{SMB}	10.23	9.89	11.01	13.35	14.08	**3.86**	[7.80]
			panel (b)				
Low β^{HML}	12.11	13.06	13.49	14.19	15.39	**3.28**	[4.10]
2	8.97	9.32	9.86	10.61	11.34	**2.37**	[4.73]
3	7.79	8.61	8.51	9.54	10.07	**2.27**	[4.73]
4	7.80	8.19	8.48	8.70	9.79	**1.99**	[4.18]
High β^{HML}	7.55	7.64	8.14	9.22	10.27	**2.72**	[4.78]
			panel (c)				
Low β^{Liq}	10.14	10.18	10.95	12.38	13.46	**3.32**	[5.25]
2	8.88	9.07	9.16	9.85	11.02	**2.15**	[3.67]
3	8.27	8.98	8.98	9.29	10.03	**1.76**	[4.23]
4	8.39	8.64	9.06	9.51	10.30	**1.90**	[4.12]
High β^{Liq}	8.50	9.42	10.10	11.46	12.66	**4.16**	[6.11]
			panel (d)				
Low β^{Mom}	6.06	6.57	7.53	7.83	8.50	**2.44**	[3.51]
2	7.05	7.47	8.03	8.76	8.92	**1.88**	[3.54]
3	8.29	8.65	8.98	9.89	10.76	**2.46**	[4.72]
4	10.02	10.26	10.23	11.20	12.33	**2.32**	[4.68]
High β^{Mom}	12.82	13.15	13.49	14.67	16.90	**4.08**	[7.36]
			panel (e)				

This table lists the equal-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^{m} is the market beta. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996) respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ, is estimated over the whole sample period of an asset. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ portfolios. The assets have to be listed consecutively on one of the exchanges for at least 180 months. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

Table 9: Double-sorted portfolios – non-financial firms

	Low δ	2	3	4	High δ	H-L	t-stat
Low β^m	6.01	6.26	7.57	8.55	9.54	**3.53**	[6.78]
2	7.34	7.98	8.15	8.52	9.88	**2.54**	[5.10]
3	8.39	8.53	8.80	9.77	10.64	**2.26**	[4.09]
4	9.47	9.48	10.29	10.88	11.82	**2.35**	[4.16]
High β^m	11.89	12.12	13.22	13.71	14.50	**2.60**	[3.71]
panel (a)							
Low β^{SMB}	7.84	8.14	9.10	9.42	10.99	**3.15**	[4.80]
2	7.96	8.58	8.90	9.65	10.73	**2.78**	[5.30]
3	8.32	9.07	9.29	10.01	11.06	**2.74**	[5.08]
4	8.64	8.85	9.43	10.36	11.98	**3.33**	[6.11]
High β^{SMB}	9.46	8.92	10.60	12.46	13.55	**4.09**	[7.62]
panel (b)							
Low β^{HML}	12.06	12.43	13.33	14.07	15.97	**3.91**	[5.09]
2	8.70	8.98	9.99	10.76	11.84	**3.14**	[5.67]
3	7.35	8.05	8.51	9.50	10.15	**2.80**	[5.59]
4	7.51	7.79	8.07	8.61	9.53	**2.02**	[4.52]
High β^{HML}	6.77	6.82	7.76	8.84	9.92	**3.15**	[5.32]
panel (c)							
Low β^{Liq}	9.32	9.45	10.52	12.05	13.29	**3.97**	[6.10]
2	8.26	8.60	9.11	9.92	11.09	**2.83**	[4.96]
3	7.99	8.46	9.07	9.41	10.31	**2.32**	[5.43]
4	8.10	8.41	9.10	9.58	10.81	**2.71**	[5.48]
High β^{Liq}	8.37	8.54	9.70	11.21	12.52	**4.16**	[5.95]
panel (d)							
Low β^{Mom}	5.47	5.99	6.82	7.73	8.94	**3.47**	[4.76]
2	6.75	7.12	7.88	8.73	8.84	**2.08**	[3.89]
3	7.98	8.22	8.86	9.82	10.71	**2.72**	[5.09]
4	9.75	9.79	10.42	11.05	12.47	**2.72**	[5.58]
High β^{Mom}	12.23	12.17	13.41	14.92	16.81	**4.57**	[7.60]
panel (e)							

This table lists the equal-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^m is the market beta. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996) respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ, is estimated over the whole sample period of an asset. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ portfolios. Assets with SIC codes between 6000 and 6200, i.e. financial firms, are excluded. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

Table 10: Double-sorted portfolios – additional factors

	Low δ	2	3	4	High δ	H-L	t-stat
Low CMA	5.56	9.37	10.92	12.57	13.62	**8.06**	[8.64]
2	6.18	8.60	10.00	10.82	10.90	**4.72**	[7.23]
3	6.03	8.01	9.90	10.28	10.34	**4.31**	[6.21]
4	6.25	7.88	9.96	10.33	10.18	**3.93**	[4.83]
High CMA	6.76	9.41	11.10	11.55	10.82	**4.06**	[4.09]
panel (a)							
Low RMW	5.55	8.99	11.06	11.78	11.39	**5.84**	[5.51]
2	5.46	7.50	9.27	9.70	9.88	**4.42**	[5.80]
3	5.70	7.69	9.39	10.13	10.41	**4.71**	[7.44]
4	6.56	8.47	10.16	10.86	11.09	**4.53**	[6.47]
High RMW	7.53	10.46	11.96	12.94	13.15	**5.61**	[5.99]
panel (b)							
Low Cokurtosis	6.77	9.82	11.09	11.65	10.72	**3.95**	[3.65]
2	5.98	9.08	10.68	11.22	10.92	**4.94**	[5.42]
3	6.01	8.51	10.62	11.06	11.48	**5.48**	[7.21]
4	5.49	8.00	10.24	10.83	11.63	**6.14**	[7.98]
High Cokurtosis	6.11	7.92	9.65	10.87	11.45	**5.34**	[8.96]
panel (c)							
Low Bk-Mkt	9.33	11.42	13.54	15.05	15.83	**6.50**	[15.86]
2	6.93	8.96	11.22	11.86	12.17	**5.24**	[9.33]
3	5.72	8.45	9.94	10.56	10.34	**4.62**	[6.67]
4	4.60	8.43	9.86	10.24	9.43	**4.84**	[4.98]
High Bk-Mkt	3.33	6.52	8.45	9.14	9.25	**5.92**	[4.42]
panel (d)							
Low σ_{R_i}	6.47	7.47	7.95	8.02	8.47	**2.00**	[3.77]
2	7.40	8.06	9.14	9.65	10.14	**2.74**	[4.63]
3	7.76	8.87	10.18	11.50	12.24	**4.49**	[8.14]
4	8.45	10.35	11.43	13.60	14.87	**6.42**	[8.47]
High σ_{R_i}	3.75	8.23	11.86	14.00	16.35	**12.60**	[12.47]
panel (e)							

This table lists the equal-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ. β^{CMA} and β^{RMW} are respectively the conservative-minus-aggressive and robust-minus-weak betas by Fama and French (2015). *cokurtosis* is the risk factor which measure fourth moment co-movement (used in Ang et al. (2006)). Bk-Mkt is the book-to-market ratio, and σ_{R_i} is the variance of the asset return. The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ, is estimated over the whole sample period of an asset. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ portfolios. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

Table 11: Double-sorted portfolios – conditional δ

	Low δ_t	2	3	4	High δ_t	H-L	t-stat
Low β^m	5.91	6.59	7.03	8.27	8.49	**2.58**	[5.19]
2	7.38	8.08	7.96	8.53	9.04	**1.66**	[2.60]
3	8.18	9.18	9.11	9.44	9.60	**1.42**	[1.99]
4	9.13	10.22	10.45	10.90	10.41	**1.29**	[1.77]
High β^m	10.78	12.78	13.25	13.41	13.27	**2.66**	[2.89]
panel (a)							
Low β^{SMB}	7.12	8.37	8.60	9.87	9.91	**2.79**	[3.59]
2	7.60	8.75	8.93	9.74	9.91	**2.31**	[4.74]
3	8.12	9.26	9.35	10.11	10.22	**2.10**	[4.36]
4	8.11	9.32	9.59	10.37	10.75	**2.65**	[4.46]
High β^{SMB}	8.85	10.12	10.82	11.29	11.75	**2.96**	[4.47]
panel (b)							
Low β^{HML}	11.10	12.80	12.91	14.52	14.29	**3.19**	[3.94]
2	8.02	8.72	9.70	10.80	10.72	**2.70**	[4.80]
3	7.17	8.55	8.56	9.24	9.61	**2.44**	[4.46]
4	7.20	8.47	8.22	8.64	8.97	**1.77**	[3.09]
High β^{HML}	6.81	7.53	7.86	8.14	8.90	**2.10**	[2.93]
panel (c)							
Low β^{Liq}	8.43	10.46	10.59	11.66	12.23	**3.80**	[5.27]
2	7.97	8.96	9.30	10.14	10.18	**2.21**	[3.91]
3	7.76	8.61	8.87	9.59	9.77	**2.00**	[3.57]
4	7.49	8.87	8.95	9.50	9.79	**2.30**	[3.74]
High β^{Liq}	8.11	9.10	9.60	10.53	10.71	**2.61**	[4.51]
panel (d)							
Low β^{Mom}	5.28	6.64	7.06	7.51	6.90	**1.62**	[2.15]
2	6.34	7.54	7.85	8.63	8.21	**1.87**	[3.32]
3	7.66	8.87	8.94	9.72	9.90	**2.23**	[4.14]
4	9.10	10.04	10.31	11.03	11.61	**2.51**	[4.29]
High β^{Mom}	11.40	12.50	13.06	14.55	15.44	**4.05**	[6.32]
panel (e)							

This table lists the equal-weighted average excess returns of double-sorted portfolios. The assets are independently sorted by their exposure to a risk factor, mentioned in the first column, and δ_t. β^m is the market beta. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996) respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). The factor exposures are estimated with monthly return data each month with a 5-year rolling window. The extreme dependence measure, δ_t, is estimated with information till time t. The average realized return of each formed portfolio is measured over the 5-year horizon used to estimate the factor loadings. The seventh column provides the average difference between the realized return on the high and low δ_t portfolios. The sample period is from 1963 to 2018. The t-statistics in the last column are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors.

Table 12: Price of risk – conditional δ

	1	2	3	4	5	6	7
Constant	4.17	4.41	4.89	9.02	5.62	5.48	5.03
	(0)***	(0)***	(0)***	(0)***	(0)***	(0)***	(0)***
β^{m}	5.10	5.31	5.03		4.81	5.22	5.23
	(0)***	(0)***	(0)***		(0)***	(0)***	(0)***
β^{SMB}		1.22			0.83	1.00	1.00
		(0.02)**			(0.12)	(0.04)**	(0.05)*
β^{HML}		-2.87			-2.82	-2.73	-2.96
		(0)***			(0)***	(0)***	(0)***
$\beta^{F^{\delta}}$			0.66	0.46	0.64	0.64	0.34
			(0.03)**	(0.18)	(0.02)**	(0.02)**	(0.25)
β^{Mom}		6.08				5.37	5.28
		(0)***				(0)***	(0)***
β^{Liq}		2.14				0.48	0.98
		(0)***				(0.35)	(0.02)**
$\beta^{\alpha^{cs}}$		1.24					-2.79
		(0.7)					(0.17)
β^{LTD}		0.89					1.19
		(0)***					(0)***
Coskewness		-9.76					-0.25
		(0.51)					(0.62)
R^2	7.44	18.65	9.35	3.01	12.93	15.71	17.72

This table shows the results of the second stage of the Fama and MacBeth (1973) regressions of 60-month excess returns on realized-risk characteristics. An overlapping 60-month rolling window is employed on assets that are listed on the NYSE, AMEX, NASDAQ or NYSE Arca. β^{m} is the market beta. $\beta^{F^{\delta_t}}$ is the beta based on the factor of extreme downside risk, δ_t. δ_t is estimated with information up to time t. The factor is constructed by subtracting the average return of the highest 20% minus the lowest 20% δ_t stocks. The portfolios are reconstructed at the start of each year. β^{HML} and β^{SMB} are high-minus-low and small-minus-big betas (Fama and French, 1996), respectively. β^{Liq} is the liquidity beta by Stambaugh and Lubos (2003). β^{Mom} is the momentum beta created by Carhart (1997). *coskewness* is the risk factor, by Harvey and Siddique (2000), which measure third moment co-movement. Furthermore, α^{cs} is Kelly and Jiang (2014) cross-sectional tail risk measure and LTD is the lower-tail dependence measure by Chabi-Yo et al. (2018). The sample period is from 1963 to 2018. The p-values for the overlapping Fama-MacBeth regression are computed using the Newey and West (1987) autocorrelation and heteroskedastic robust standard errors. The last row reports the average R^2 over the rolling windows of the second-stage regressions. *, **, *** are indicators for the significance level at 10%, 5% and 1% of the coefficients, respectively.

Table 13: Correlation matrix of factor loadings

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
β^m	1.00	0.20	-0.24	-0.05	-0.09	0.24	0.92	-0.19	0.14	-0.11	0.01	0.45	0.07	0.28
β^{SMB}	0.20	1.00	-0.02	-0.08	-0.11	-0.06	0.27	0.15	-0.11	0.03	0.01	0.48	-0.00	-0.37
β^{HML}	-0.24	-0.02	1.00	0.08	-0.03	-0.04	-0.24	-0.02	-0.05	0.06	-0.08	-0.16	-0.08	-0.02
β^{Mom}	-0.05	-0.08	0.08	1.00	-0.02	0.00	-0.04	0.04	-0.15	0.09	0.10	-0.08	-0.00	-0.05
β^{Liq}	-0.09	-0.11	-0.03	-0.02	1.00	-0.03	-0.06	0.09	-0.19	0.05	0.10	-0.03	-0.01	-0.03
δ	0.24	-0.06	-0.04	0.00	-0.03	1.00	0.21	-0.07	0.06	-0.03	-0.01	-0.03	0.01	0.26
β^-	0.92	0.27	-0.24	-0.04	-0.06	0.21	1.00	0.20	0.05	-0.05	0.02	0.53	0.07	0.14
$\beta^- - \beta^m$	-0.19	0.15	-0.02	0.04	0.09	-0.07	0.20	1.00	-0.21	0.15	0.01	0.23	0.01	-0.34
Coskewness	0.14	-0.11	-0.05	-0.15	-0.19	0.06	0.05	-0.21	1.00	-0.58	0.04	-0.01	0.01	0.17
Cokurtosis	-0.11	0.03	0.06	0.09	0.05	-0.03	-0.05	0.15	-0.58	1.00	-0.05	-0.00	-0.01	-0.11
$\beta^{\alpha_{CS}}$	0.01	0.01	-0.08	0.10	0.10	-0.01	0.02	0.01	0.04	-0.05	1.00	0.08	0.00	-0.03
σ_{R_i}	0.45	0.48	-0.16	-0.08	-0.03	-0.03	0.53	0.23	-0.01	-0.00	0.08	1.00	0.06	-0.41
Bk-Mkt	0.07	-0.00	-0.08	-0.00	-0.01	0.01	0.07	0.01	0.01	-0.01	0.00	0.06	1.00	0.07
log(size)	0.28	-0.37	-0.02	-0.05	-0.03	0.26	0.14	-0.34	0.17	-0.11	-0.03	-0.41	0.07	1.00

This table describes the pairwise correlation between factor loadings that are used in this paper. The correlations shown are the average correlations measured over a 5-year overlapping rolling window. The number of firms included in the analysis is between 1437 and 3889.