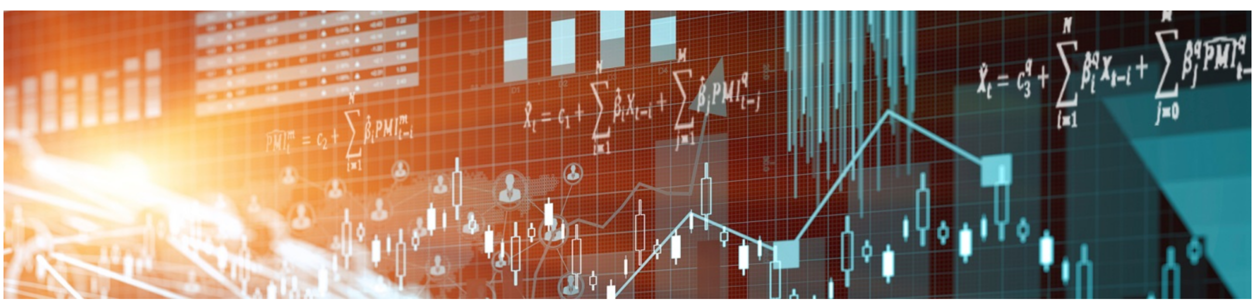
BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2018-47

Challenges in Implementing Worst-Case Analysis



by Jon Danielsson, Lerby M. Ergun and Casper G. de Vries

www.bank-banque-canada.ca

Bank of Canada Staff Working Paper 2018-47

September 2018

Challenges in Implementing Worst-Case Analysis

by

Jon Danielsson,[1] Lerby M. Ergun[2] and Casper G. de Vries[3]

[1] London School of Economics

[2] London School of Economics
and
Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
lergun@bankofcanada.ca

[3] Erasmus University Rotterdam

ISSN 1701-9397

Acknowledgements

This work was supported by the Economic and Social Research Council (ESRC) in funding the Systemic Risk Centre [grant number ES/K002309/1] and the Netherlands Organisation for Scientific Research Mozaiek grant [grant number: 017.005.108].

Abstract

Worst-case analysis is used among financial regulators in the wake of the recent financial crisis to gauge the tail risk. We provide insight into worst-case analysis and provide guidance on how to estimate it. We derive the bias for the non-parametric heavy-tailed order statistics and contrast it with the semi-parametric extreme value theory (EVT) approach. We find that if the return distribution has a heavy tail, the non-parametric worst-case analysis, i.e. the minimum of the sample, is always downwards biased and hence is overly conservative. Relying on semi-parametric EVT reduces the bias considerably in the case of relatively heavy tails. But for the less-heavy tails this relationship is reversed. Estimates for a large sample of US stock returns indicate that this pattern in the bias is indeed present in financial data. With respect to risk management, this induces an overly conservative capital allocation if the worst case is estimated incorrectly.

Bank topic: Financial stability
JEL codes: C01, C14, C58

Résumé

Depuis la récente crise financière, l’analyse du pire scénario est utilisée par les autorités de réglementation du secteur financier pour évaluer le risque extrême. Nous apportons de nouvelles perspectives sur cette méthode et sur l’estimation de la valeur extrême qui en découle. Nous calculons le biais des estimateurs d’ordre non paramétrique de la queue de distribution et le comparons au biais associé à la méthode semi-paramétrique de la théorie des valeurs extrêmes (TVE). Quand la distribution des rendements a une queue épaisse, nous trouvons que la valeur minimum de l’échantillon – c’est-à-dire l’estimateur issu de l’analyse du pire scénario – est très modérée dans la mesure où elle surestime toujours le risque. Dans le cas des distributions à queues relativement épaisses, le biais se réduit substantiellement grâce à l’estimateur semi-paramétrique résultant de la TVE. Pour les queues moins épaisses, la relation s’inverse. Les estimations tirées d’un large échantillon du rendement d’actions du marché américain révèlent en effet la présence d’un comportement semblable dans les données financières. Sur le plan de la gestion du risque, ce résultat se traduit par une allocation très prudente des capitaux si le pire scénario est incorrectement estimé.

Sujet : Stabilité financière
Codes JEL : C01, C14, C58

Non-technical Summary

Worst-case analysis studies the worst expected outcome over a predetermined time length, with a typical question: What is the worst daily market outcome in 10 years? This type of analysis is increasingly common since the recent financial crisis. Much of the bank stress testing scenario production is based on the worst observed historical event (BIS, 2017 and EIOPA, 2014). Others use it as a value-at-risk metric. In spite of its increasing importance, little is known about worst-case analysis and its proper estimation.

There are generally three main approaches for worst-case analysis. The simplest, and the most obvious, is to directly read the object of interest from the empirical distribution, in our case the historical minima. One can also assume a model only for the tail of the distribution and not model the center of the distribution, i.e. a semi-parametric approach. The third approach is based on specifying a parametric distribution for all outcomes and estimating its parameters. Of these three alternatives, the last is the only one that cannot be recommended due to an over-representation of observations from the center of the distribution.

In this paper we compare the historical minima and semi-parametric approaches. Both worst-case estimators are biased towards giving larger values. However, we find the method that produces the smallest bias depends on how heavy the tail of the distribution is. The semi-parametric approach produces the smallest bias for very heavy-tailed distributions, and the historical minima produces the smallest bias for the relatively lighter heavy-tailed distributions. We confirm this relationship for the individual securities traded on the US stock exchanges. Choosing the inappropriate estimator can lead to an overly conservative capital allocation.

1 Introduction

Worst-case analysis studies the worst expected outcome over a predetermined time length, with a typical question: What is the worst daily market outcome in 10 years or 2,500 days? This type of analysis is increasingly common due to the recent financial crises. Much of the bank stress testing scenario production is based on the worst observed historical event (BIS, 2017 and EIOPA, 2014). Others use it as a value-at-risk (VaR) metric.[1] In spite of its increasing importance, little is known about worst-case analysis and its proper estimation.

An inappropriate measurement of this risk metric can lead to a misallocation of capital. In this paper we compare the non-parametric and semi-parametric approaches. Both worst-case estimators are downwards biased for the left tail quantile. The method that produces the smallest bias depends on the heaviness of the tail. The semi-parametric approach is first-moment stochastic dominant for very heavy-tailed distributions, and the non-parametric approach is first-moment stochastic dominant for relatively lighter heavy-tailed distributions. Given the second-moment stochastic dominance of the semi-parametric approach, there is a strict preference for the semi-parametric estimator in the case of the more heavy-tailed distributions. We confirm this is the relevant case for the individual securities traded on the US stock exchanges.

There are generally three main approaches to worst-case analysis. The simplest is to directly read the object of interest from the empirical distribution, in our case the historical minima. This is the non-parametric approach (NP). One can also assume a model only for the tail of the distribution and not model the center of the distribution; this constitutes the semi-parametric approach (SP). The third approach is based on specifying a fully parametric distribution for all outcomes and estimating its parameters. Of these three alternatives, the last is the only one that is not recommended. The reason is that the estimates are dominated by the center of the distribution, so that the fit is optimal for a typical observation, but not the lowest. Therefore, such an approach would in most cases deliver less precise and more uncertain worst-case estimates than either the NP or the SP approach. We therefore focus on NP and the SP estimators and provide guidance on the appropriate use of either.

The NP quantile estimator is the maximum sample ordered observation, i.e.

[1]Danielsson (2011) gives a comprehensive overview of the different methodologies and issues regarding VaR estimation.

the most extreme order statistic. Through the use of extreme value theory (EVT) and under the assumption that the underlying distribution is heavy-tailed, Leadbetter, Lindgren and Rootzén (1983) derive the asymptotic distribution of the order statistics. Our focus is on the most extreme observation. Using the asymptotic distribution of the extreme order statistics we derive the bias and the variance of the extreme order statistics. The NP estimator is downwards biased and increases in the heaviness of the tail. The variance of the NP approach is large and does not exist for distributions with a tail index lower than or equal to 2.

The SP estimator for the class of heavy-tailed distribution is the Weissman (1978) estimator for the most extreme quantiles. The SP estimator is derived by inverting the first-order Taylor expansion at infinity of the CDF for heavy-tailed distributions. This inversion necessitates the estimation of the tail index. The tail index is estimated by means of the Hill (1975) estimator. The statistical properties of the tail index estimate dominate the properties of Weissman's quantile estimator. Goldie and Smith (1987) provide the asymptotic distribution of this estimator. The SP estimator is normally distributed with the bias and variance decreasing in the heaviness of the distribution. Additionally, the bias and the variance are dependent on the number of order statistics, t, utilized to estimate the tail index and scaling constant of the quantile estimator. The bias of the SP estimator is decreasing for $t > exp(2)$.

The literature on the estimation of the extreme quantiles has put forth various bias-reducing estimators. The bias of the SP estimator can be largely attributed to the bias in the tail index estimator. Gomes and Pestana (2007), for instance, proposed an adequate bias-corrected tail index estimator by estimating the parameters of the second order in the Taylor expansion of the tail distribution function. This paper tries to minimize the influence of this bias, utilizing a distance metric which minimizes the distance between the empirical and theorized distribution. This approach implicitly penalizes large deviations in the very extreme part of the tail. It therefore often selects only a very limited number of order statistics to estimate the Hill estimator and therefore reduces the influence of this bias.

Many of the empirical applications based on VaR focus on a probability closer to the center of the distribution. The academic literature has scarcely focused on the worst case as a risk measure. Ghaoui, Oks, and Oustry (2003) use the maximum VaR over a random space of probability distributions for robust portfolio optimization. Zhu and Fukushima (2009) extend their paper by including expected shortfall as the basis of their risk measure. Along this line, Kerkhof, Melenberg, and Schumacher (2010) use the worst case across classes of models to incorporate model risk in capital reserve requirements.

The contribution of this paper lies in the comparison of the biases of the two worst-case estimators. We show that the choice of estimator with the lowest bias hinges on the tail index. For relatively heavy-tailed distributions, the SP estimator has the smallest bias. This relationship reverses as the distribution becomes less heavy-tailed. For example, for the Student-t family of distributions, the point where the SP bias becomes larger than the NP bias occurs around the Student-t distribution with 6 degrees of freedom. Given that the variance of the NP approach is strictly larger than that of the SP approach, for very heavy-tailed distributed variables the SP approach is the strictly preferred estimator. Beyond the point where the relative size of their biases switches, one needs to consider the bias-variance trade-off of the estimators.

The comparison of the bias puts forth two predictions. First, the difference between SP and NP, i.e. SP-NP, is an increasing function in the tail index. Second, this difference is decreasing in t for $t > exp(2)$. To investigate these predictions, we use the securities return data by the Center for Research in Security Price (CRSP) to apply the two worst-case estimators. We estimate for each individual stock the NP and SP estimator. We evaluate the relationship between the difference of the two estimates and the tail estimate and t. The results from the empirical analysis reveal that for stocks with a heavy-tailed distribution, the SP estimate is smaller than the NP estimate. This changes for stocks with a larger estimated tail index, i.e. less heavy tail. The switching of the relative size of the bias occurs for stocks with a tail index above 3. We also find that t is negatively related to SP-NP. This shows that the predicted relationships, with respect to t and $\hat{\alpha}$, in the relative bias of the worst-case estimates can also be found in financial return data. Therefore, the guidelines provided by the comparison of the bias and variance should be taken into account when choosing the estimator for worst-case analysis.

In the next section we introduce the two quantile estimators and analyze their bias and variance. In the subsequent section we explore the extent of the bias in US securities data. The last section concludes.

2 Worst-case Estimators

This paper defines the worst case as the worst potential daily loss over n number of days. Under the i.i.d. assumption this equates to the daily VaR at probability level $1/n$. Given this approach, we rely on EVT to further derive the properties of the SP and NP estimator.

2.1 The non-parametric approach

To derive the bias of the worst observation as a worst-case estimator, we start with a relatively general approach. We begin by deriving the distribution of observations in an ordered sample. Suppose one observes some i.i.d. heavy-tailed random variable $Y_1, ..., Y_n$ with distribution F, where

$$\lim_{s \to \infty} \frac{1 - F(sx)}{1 - F(s)} = x^{\alpha}, \quad \alpha > 0. \tag{1}$$

The class of distribution functions with heavy tails, like the Student-t, Pareto, stable distribution or the unconditional distribution of the stationary solution to a GARCH(1,1) process, is precisely defined in terms of the regular variation property. The sorted sample, i.e. order statistics, can be represented as

$$\max(Y_1, ..., Y_n) = X^{(1,n)} \geq X^{(2,n)} \geq \cdots \geq X^{(n,n)} = \min(Y_1, ..., Y_n).$$

The distribution of the order statistics can be studied through the number of exceedances. These follow a binomial distribution:

$$G^{(k,n)}(x) = \sum_{r=0}^{k-1} \binom{n}{r} [1 - \mathrm{F}(x)]^r [\mathrm{F}(x)]^{n-r}. \tag{2}$$

Suppose one is interested in the distribution of the maximum realization:

$$\Pr(\max(Y_1, ..., Y_n) < x) = G^{(1,n)}(x) = [\mathrm{F}(x)]^n. \tag{3}$$

Similar to the standard central limit theorem for the asymptotic distribution of the arithmetic mean, Fisher and Tippett (1928) and Gnedenko (1943) provide a limit theorem for the asymptotic distribution of the maximum, i.e. EVT.

EVT gives the conditions under which there exist sequences b_n and a_n such that

$$\lim_{n \to \infty} [\mathrm{F}(a_n x + b_n)]^n \to \mathrm{G}^{(1,n)}(x),$$

where $\mathrm{G}^{(1,n)}(x)$ is the Fréchet distribution for heavy-tailed distributions.

Theorem 2.2.2 in Leadbetter, Lindgren, and Rootzén (1983) extends the EVT for the maximum to lower order statistics by means of the Poisson property of the lower order statistics. In particular, the asymptotic distribution of the k^{th} largest order statistic is

$$\mathrm{G}^{(k,n)}(x) \to \mathrm{G}^{(1,n)}(x) \sum_{s=0}^{k-1} \frac{\left(-\log\left[\mathrm{G}^{(1,n)}(x)\right]\right)^s}{s!}. \tag{4}$$

From (4) we determine the expectation of the order statistics[2]

$$\mathrm{E}\left[X^{(k,n)}\right] = \frac{a_n}{[k-1]!}\Gamma\left[k - \frac{1}{\alpha}\right] \tag{5}$$

and the variance

$$\mathrm{var}\left[X^{(k,n)}\right] = \frac{a_n^2}{[k-1]!}\left[\Gamma\left[k-\frac{2}{\alpha}\right] - \frac{1}{[k-1]!}\Gamma\left[k-\frac{1}{\alpha}\right]^2\right]. \tag{6}$$

Here $\Gamma\left(\right)$ refers to the gamma function. To determine this expectation for a specific heavy-tailed distribution, a_n needs to be chosen appropriately.[3] To find a good approximation, we use the first-order term of the Hall expansion (Hall and Welsh, 1985):

$$\Pr\left(Y \leq -y\right) = \mathrm{F}(-y) = Ay^{-\alpha}[1 + By^{-\beta} + o(y^{-\beta})]. \tag{7}$$

For the Pareto distribution, $\mathrm{F}\left(-y\right) = Ay^{-\alpha}$, we observe that the Hall expansion perfectly fits the first-order term.[4] For the Pareto distribution the scaling constant a_n is $(An)^{\frac{1}{\alpha}}$, where A is fixed the scale parameter. Therefore, we can, through the Hall expansion, extract a good approximation of the expectation of the order statistics of heavy-tailed distributions.

2.2 The semi-parametric approach

To contrast the expectation of the maximum observation, we compare it to an SP estimator of the worst case. By inverting the first-order expansion

[2]See Appendix for the derivations.

[3]For the heavy-tailed distributions, $b_n = 0$.

[4]All of the standard heavy-tailed distributions satisfy the Hall expansion. This also applies to the GARCH(1,1) unconditional distribution, but the class by (7) is a bit narrower than the class defined by (1).

in (7), using the empirical counterpart of $A = \frac{t/n}{X^{(t,n)^{-\alpha_t}}}$ measured at some threshold t and $F(-y) = {}^k\!/n$, one obtains the SP tail quantile estimator by Weissman (1978)

$$\hat{x}_{SP}^{(n-k/n)}(t) = X^{(t,n)} \left(\frac{t}{k}\right)^{\frac{1}{\alpha_t}}.$$

For distributions where (7) applies, Goldie and Smith (1987) derive the distribution of the SP quantile estimator

$$\frac{\sqrt{t}}{\log(t/np)} \left(\frac{\hat{x}_{SP}^{(n-k/n)}(t)}{x^{(p)}} - 1\right) \sim N\left(-\frac{\operatorname{sign}(B)}{\sqrt{2\beta\alpha}}, \frac{1}{\alpha^2}\right), \tag{8}$$

where B and β are the second-order scale and shape parameters from (7).

2.3 Comparing the non-parametric and semi-parametric quantile estimators

The two approaches, the NP and the SP, each have their own advantages and disadvantages. While the NP is much simpler to implement, the SP might be more accurate because it uses more tail observations in the estimation, and therefore might result in an estimator with a smaller variance. However, the SP is dependent on correctly specifying the SP distribution and identifying a threshold $X^{(t,n)}$.

To shed more light on the use of these two estimators, we compare their bias at $p = {}^1\!/n$. From (5) and (8) the bias of the two approaches is as follows:

$$\left(\frac{\hat{x}_{SP}^{(n-1/n)}(t)}{x^{(p)}} - 1\right) \sim -\frac{\operatorname{sign}(B)}{\sqrt{2\beta\alpha}} \frac{\log(t)}{\sqrt{t}} \qquad \text{SP} \tag{9}$$

$$\left(\frac{\hat{x}_{NP}^{(n-1/n)}}{x^{(p)}} - 1\right) \sim \Gamma\left[1 - \frac{1}{\alpha}\right] - 1. \qquad \text{NP} \tag{10}$$

Expressions (9) and (10) indicate that neither approach is first-moment stochastic dominant in all circumstances. For the NP estimator, the asymptotic bias approaches infinity as α approaches 1. However, as α increases, the $\Gamma()$ function decreases rapidly. As α approaches 1, the bias in the SP estimator is relatively small for moderate values of β and t. This leads to a crossing point in the bias of the two estimators with respect to α.

Given values of t and β, we define switching point α^*. For $\alpha < \alpha^*$, the absolute bias of the SP estimator is smaller than that of the NP estimator.

When $\alpha > \alpha^*$, the relationship is reversed. For a fixed t, this relationship is depicted in Figure 1.[5] This figure portrays at which combination of α and β the bias of the NP worst-case estimator becomes smaller than the SP approach.

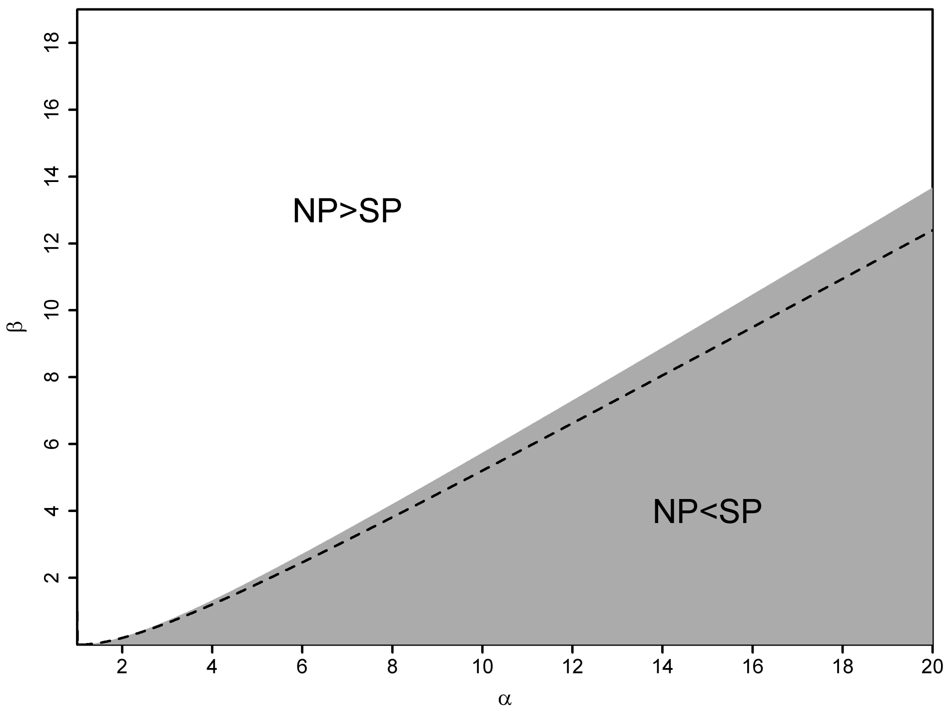


Figure 1: This figure depicts the area where the absolute bias of the semi-parametric estimator becomes larger than the bias of the order statistic (gray area). The biases of the estimators are at $p = 1/n$ as in Equation (9) and (10). For this figure we fix t at $exp(2)$. To the right of the lines, the combination of α and β produces a larger bias for the semi-parametric approach. The dotted line shows where the boundary shifts to when the threshold t is doubled to $2exp(2)$.

In the case of the family of Student-t distributions, $\beta = 2$ and α equals the degrees of freedom for the specific Student-t distribution. From Figure 1, we read that in the case of the family of Student-t distributions, the switching of the biases occurs around $\alpha^* \approx 5$. For higher and lower values of t, the α^* increases. For the family of symmetric stable distributions, the bias is always smaller for the SP estimator, as $\beta = \alpha$ and $\alpha < 2$.

[5]The bias of SP reaches its maximum at $t = exp(2)$.

Given the above expectation, determining the variance of the order statistics is a trivial matter.

$$var\left[\hat{x}_{NP}^{(n-1/n)}\right] = \frac{a_n^2}{[k-1]!}\Gamma\left[k-\frac{2}{\alpha}\right] - \left[\frac{a_n}{[k-1]!}\Gamma\left[k-\frac{1}{\alpha}\right]\right]^2 \tag{11}$$

$$var\left[\hat{x}_{SP}^{(n-1/n)}\right] = a_n^2\frac{1}{\alpha^2}\frac{\log\left(t/np\right)}{t}. \tag{12}$$

The variance of the NP estimators is strictly larger than that of the SP estimator. For $\alpha \leq 2$, the variance of the NP does not exist. Figure 2 shows that for the relatively heavy-tailed distributions, the NP estimator has a disproportionally larger variance than the SP estimator. This result, combined with the comparison of the biases, indicates a strong preference for the use of the SP approach over the NP approach for very heavy-tailed distributed variables.[6]

The comparison of the biases leads to two empirical predictions. First, the difference between the SP and NP estimator is an increasing function in α. Second, given that t is above $exp(2)$, the difference is decreasing in the number of order statistics used in the SP approach. This dictates a negative relationship between the difference in the biases and t. In the next section we test these predictions for US stock market data.

3 Empirical Application

Financial institutions use VaR-based risk metrics to assess the risk of assets they have on their books. The CRSP dataset contains a large cross-section of daily stock prices for US stocks, the kind of assets financial institutions typically hold. The large cross-section of stocks allows us to compare a large number of worst-case estimates for the NP and SP estimators. Therefore, we are able to determine whether the predicted differences in the bias and variance of the estimators are also present in financial assets.

3.1 Data

The CRSP database contains individual stock data from 1925-12-31 to 2015-12-31 for NYSE, AMEX, NASDAQ, and NYSE Arca. In the main analysis,

[6]Figure 3 in the Appendix depicts the ratio of the MSE of the two worst-case estimators. The MSE of the two estimators also shows that for the more heavy-tailed distributions the SP estimator is strictly preferred over the NP estimator.

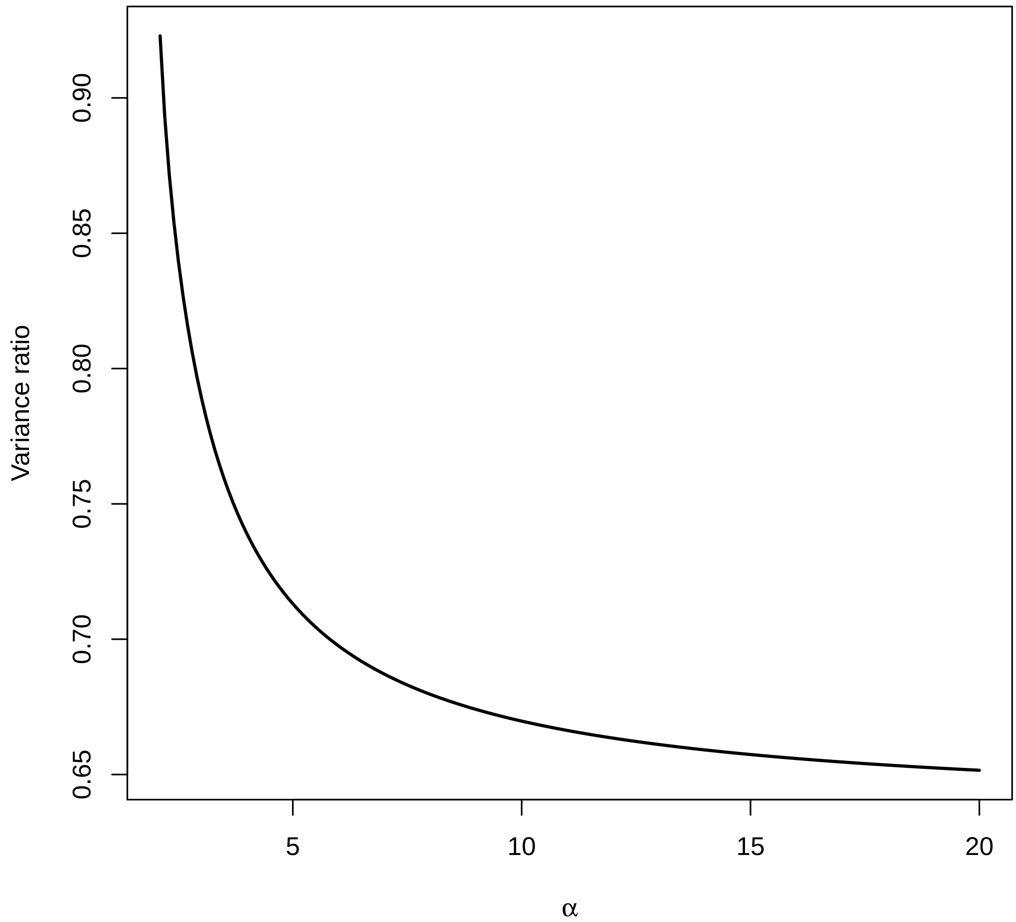


Figure 2: This figure displays the variance ratio of the semi-parametric and the non-parametric worst-case estimator as a function of α. The variance ratio is given by $\sigma^2_{NP}/(\sigma^2_{SP}+\sigma^2_{NP})$. For the variance of the semi-parametric estimator, we choose $t = exp(2)$.

$n = 1,986$ stocks are used. For every stock that is included in the analysis, we require that it be traded on one of the four exchanges during the whole measurement period, which is between 01-01-1995 and 01-01-2011.[7] The fixed time period is to ensure that the sample size is large enough for the

[7] In the CRSP database, exchange code -2, -1, 0 indicates that a stock was not traded on one of the four exchanges and thus no price data is recorded for these days. Stocks that contain exchange code -2, -1, 0 are not included in the analysis. We only use stocks with share code 10 and 11.

EVT estimation.[8] Furthermore, this ensures that the empirical probability at the largest order statistic is the same across different securities. The choice of the specific sample period is to ensure a large cross-sectional sample. We also require the average price of the stock to be above 5 dollars over the measurement period.

3.2 Empirical analysis

The SP estimator requires the estimation of α and t. For this empirical application we use the Hill estimator to estimate the tail exponent α. This estimator depends on a selection of a high order statistic as a threshold, i.e. $X^{(t,n)}$. This nuisance statistic is obtained by the KS-distance metric developed in Danielsson et al. (2016).[9] The KS-distance metric is focused on picking t to fit the quantile of distribution. Danielsson et al. (2016) show that alternative approaches, e.g. Danielsson et al. (2001) and Drees and Kaufmann (1998), underperform significantly, especially when it comes to the quantiles deep in the tail of the distribution.[10]

Given the estimate of the tail index and nuisance statistic, the quantile can be estimated semi-parametrically for every individual stock. We compare the difference between the previously introduced worst-case estimators for each stock at the $1/n$ quantile. The difference,

$$SP_i - NP_i = X_i^{(t,n)} \left(t_i\right)^{\frac{1}{\hat{\alpha}_t^i}} - X_i^{(1,n)}, \tag{13}$$

for stock i has an estimate of the tail index in the SP quantile estimator. To look at the initial relationship between the bias in the two estimators, we sort the individual stocks by their estimated α^i. Based on $\hat{\alpha}^i$, the stocks are assigned to five different baskets with a range of $\{\hat{\alpha}^i < 2, 2 < \hat{\alpha}^i \leq 3, 3 < \hat{\alpha}^i \leq 4, 4 < \hat{\alpha}^i \leq 5, 5 \leq \hat{\alpha}^i\}$. Table 1 reports the aggregate statistics of the difference in the quantile estimators, $SP_i - NP_i$, for each basket.

The theoretical results stipulate that the relative size of the bias changes as a function of $\hat{\alpha}$. Table 1 portrays this pattern for the left tail of the securities in

[8] The size of the time series for each individual firm is 4,030 days.

[9] The KS-distance metric chooses the threshold which minimizes the maximum quantile distance between the empirical and Pareto distribution. This approach is further explained in the Appendix.

[10] We also use a fixed number of order statistics, t, for each stock and obtain similar results. See Table 5 in the Appendix.

Table 1: $SP_i - NP_i$ sorted by $\hat{\alpha}^i$

	All	$\hat{\alpha}^i < 2$	$2 \leq \hat{\alpha}^i < 3$	$3 \leq \hat{\alpha}^i < 4$	$4 \leq \hat{\alpha}^i < 5$	$\hat{\alpha}^i \geq 5$
Mean	0.041	-4.234	-0.909	0.369	1.197	1.861
Median	0.701	-5.749	-0.754	0.701	1.337	1.865
St. Dev.	2.235	5.664	2.926	1.303	1.154	0.604
Q0.01	-7.708	-9.186	-10.472	-3.087	-1.768	0.491
Q0.99	3.745	3.567	4.650	2.520	3.480	2.921
Rank Sum test	0.000	0.250	0.397	0.000	0.000	0.000
N	888	4	329	392	144	19

This table reports summary statistics for the difference between the semi-parametric quantile estimator and the largest order statistic, $SP_i - NP_i$, for the **left tail** of US stocks. For the SP_i estimator, α^i is estimated with the Hill estimator. To determine the number of order statistics for the Hill estimator, we use the KS-distance metric described in Danielsson et al. (2016). Column 1 reports the summary statistics of SP-NP for all stocks. The second column reports the summary statistics of the difference for the stock with $\hat{\alpha}^i \leq 2$. Columns 3 through 6 report the summary statistics for the stocks with the corresponding $\hat{\alpha}^i$. The first three rows report the mean, median and standard deviation of the corresponding baskets. Q0.01 and Q0.99 report the 1% and 99% quantile for the distribution of $SP_i - NP_i$ for the different basket of stocks. The next row reports the Wilcoxon signed-rank test p-value, testing non-parametrically for a difference in mean rank. N is the number of stocks in each basket. The individual stock data is from the CRSP dataset. The securities need to be traded on NYSE, AMEX, NASDAQ, and NYSE Arca exchanges over the period from 01-01-1995 to 01-01-2011. To be included, the average stock price over the sample needs to be above 5 dollars.

the CRSP database.[11] For these stocks the switch point is around $\hat{\alpha}^* = 3$. It is difficult to determine the exact switch point for real data. This is because the second order parameter β, in the bias of the SP quantile estimator, is hard to estimate precisely. In addition, the Hill estimator is generally biased (Hall, 1982). This makes it difficult to determine the exact switch point. It is encouraging that we see a monotonic decline in the average difference as $\hat{\alpha}$ increases. This is supportive of the result that the bias of the EVT-based worst-case estimator overtakes the bias of the NP quantile estimator. The results for the difference in the median of each basket convey the same story.

The 1% and 99% quantiles of the buckets show that although the mean and median showcase a switch between the severity of the bias of the quantile estimators, this might be statistically insignificant. Therefore, we employ the Wilcoxon signed-rank sum test to test for the difference in size of SP_i and NP_i estimates. We find that for the stocks with a lighter heavy-tailed return distribution the estimates are significantly different from one another.

[11]See Table 3 in the Appendix for the results for the right tail of the distribution for the individual stocks.

The empirical distribution of the SP quantile estimates tends to have larger values than the distribution of the NP quantile estimates. This is significant for $\hat{\alpha} \leq 2$ at a 5% significance level. This is due to the small sample and the fact that the variance of both estimators is relatively large for low values of α. For stocks with $\alpha \approx \alpha^*$, the test shows that there is no significant difference in the size of the estimates.

Table 2: CRSP data

	SP-NP					
	Left tail			Right tail		
$\hat{\alpha}^i$	1.261***		0.724***	1.808***		1.032***
	(0.116)		(0.106)	(0.178)		(0.143)
$t_i/n * 100$		−0.336***	−0.222***		−0.417***	−0.263***
		(0.040)	(0.044)		(0.051)	(0.052)
Constant	−4.135***	0.911***	−1.777***	−5.929***	1.156***	−2.651***
	(0.433)	(0.092)	(0.436)	(0.640)	(0.121)	(0.552)
Observations	864	864	864	867	867	867
R^2	0.174	0.190	0.226	0.216	0.227	0.266

This table reports the regression results for the difference between the semi-parametric quantile estimator and the largest order statistic, $SP_i - NP_i$, for US stocks. For the SP estimator, α^i is estimated with the Hill estimator. To determine the number of order statistics for the Hill estimator we use the KS-distance metric described in Danielsson et al. (2016). Here $t_i/n * 100$ is the percentage of order statistics from the total sample to estimate the Hill estimate. We include only stocks with $t_i > exp(2)$. The individual stock data is from the CRSP dataset. The securities need to be traded on NYSE, AMEX, NASDAQ, and NYSE Arca exchanges over the period from 01-01-1995 to 01-01-2011. To be included, the average stock price over the sample needs to be above 5 dollars.

Table 2 reports the results of regressing $SP_i - NP_i$ on their respective tail index and nuisance parameter t. The signs of the parameter estimates are as prescribed by the comparisons of the biases in Section 2.3. The coefficient of $\hat{\alpha}$ in the first column shows that an increase of the tail index by 1 increases the difference in the worst-case return estimates by 1.261 percentage points. The difference switches from negative to positive around a tail index of 3.3. When including the fraction of order statistics utilized in the estimation of the SP approach, the coefficient is as predicted. An increase in the number of order statistics past $t = exp(2)$ decreases the bias in the SP approach and therefore decreases the difference in the worst-case estimates. Both $\hat{\alpha}$ and $\hat{t}$

have a significant effect on the difference in estimates. This holds for both the left and right tails of the distribution.[12]

For the regressions presented in Table 2 we use $\hat{\alpha}_i$ instead of the true tail index. The measurement error in $\hat{\alpha}_i$ could be correlated with $SP_i - NP_i$. To address this issue we use an instrumental variable approach. In a two-stage least-square regression, we use kurtosis, skewness and the standard deviation of the empirical distribution as instruments for the tail index.[13] Table 4, in the Appendix, shows that the higher moments of the return distribution explain a large portion of the variation in $\hat{\alpha}$. The second-stage regression shows that the relationship between the bias and the tail index is not driven by the measurement error in $\hat{\alpha}_i$.

[12]The restricted sample period in Table 2 is chosen to maximize the cross-section of returns. To show that these results are not sample-specific, Figure 4 in the Appendix depicts the coefficients of the third and sixth regression model for a 10-year sample period each year between 1975 to 2015.

[13]We have excluded the top and bottom 5% of the sample to prevent the tail observations from influencing the instruments. Results where the instruments are based on the full sample are quantitatively equivalent to the censored sample.

4 Conclusion

With worst-case analysis becoming increasingly common in both policymaking and practice, it is of interest to evaluate the qualities of common methods for such applications. The simplest and perhaps the most common way is to estimate the worst case by taking the most negative outcome in the historical sample. Alternatively, one could estimate the lower tail of the distribution by semi-parametric methods and use that to calculate the worst case.

Our overall conclusion is that either method is best, depending on how heavy the tails are and their specific shape. Generally, for the heaviest, the semi-parametric approach is best, and as it thins, the historical minima eventually becomes better. This is further reinforced by the strictly higher variance of the non-parametric estimator compared to the semi-parametric estimator. These results are further confirmed in US stock market data. Individual stocks with a relatively heavy tail have on average a lower semi-parametric worst-case estimate. This relationship is reversed for stocks with a thinner heavy tail.

References

BIS, 2017. Supervisory and bank stress testing: Range of practices. Unpublished working paper. Bank for International Settlements.

Danielsson, J., 2011. Financial risk forecasting: The theory and practice of forecasting market risk with implementation in R and Matlab vol. 587. John Wiley & Sons, New Jersey.

Danielsson, J., L. de Haan, L. Peng, C. de Vries, 2001. Using a bootstrap method to choose the sample fraction in tail index estimation. Journal of Multivariate Analysis 76(2), 226–248.

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

Drees, H., E. Kaufmann, 1998. Selecting the optimal sample fraction in univariate extreme value estimation. Stochastic Processes and Their Applications 75(2), 149–172.

EIOPA, 2014. EIOPA insurance stress test. Unpublished working paper. Consultative Document EIOPA.

Fisher, R. A., L. Tippett, 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. In *Proceedings of the Cambridge Philosophical Society* vol. 24 pp. 180–190.

Ghaoui, L. E., M. Oks, F. Oustry, 2003. Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. Operations Research 51(4), 543–556.

Gnedenko, B., 1943. Sur la distribution limite du terme maximum d'une série aléatoire. Annals of Mathematics 44(1), 423–453.

Goldie, C. M., R. L. Smith, 1987. Slow variation with remainder: Theory and applications. The Quarterly Journal of Mathematics 38(1), 45–71.

Gomes, M. I., D. Pestana, 2007. A sturdy reduced-bias extreme quantile (VaR) estimator. Journal of the American Statistical Association 102(477), 280–292.

Hall, P., 1982. On some simple estimates of an exponent of regular variation. Journal of the Royal Statistical Society. Series B (Methodological) pp. 37–42.

Hall, P., A. Welsh, 1985. Adaptive estimates of parameters of regular variation. The Annals of Statistics 13(1), 331–341.

Hill, B. M., 1975. A simple general approach to inference about the tail of a distribution. The Annals of Statistics 3(5), 1163–1174.

Kerkhof, J., B. Melenberg, H. Schumacher, 2010. Model risk and capital reserves. Journal of Banking & Finance 34(1), 267–279.

Leadbetter, M. R., G. Lindgren, H. Rootzén, 1983. Extremes and related properties of random sequences and processes. Springer Series in Statistics.

Weissman, I., 1978. Estimation of parameters and large quantiles based on the k largest observations. Journal of the American Statistical Association 73(364), 812–815.

Zhu, S., M. Fukushima, 2009. Worst-case conditional value-at-risk with application to robust portfolio management. Operations Research 57(5), 1155–1168.

A Expectation and Variance

Given the CDF for the lower order statistics,

$$\mathrm{G}^{(k,n)}(x) \to \mathrm{G}^{(1,n)}(x) \sum_{s=0}^{k-1} \frac{\left(-\log\left[\mathrm{G}^{(1,n)}(x)\right]\right)^{s}}{s!}, \tag{14}$$

where $\mathrm{G}^{(k,n)}(x)$ is the CDF of the k^{th} order statistic. For the domain of attraction of the heavy-tailed distributions, $\mathrm{G}^{(1,n)}(x)$ is the CDF of the Frechet distribution. Therefore, we have

$$\mathrm{G}^{(k,n)}(x) = e^{-a_n^{\alpha}x^{-\alpha}} \sum_{s=0}^{k-1} \frac{\left(a_n^{\alpha}x^{-\alpha}\right)^{s}}{s!}.$$

For the density we find

$$g^{(k,n)}(x) = \alpha a_n^{\alpha} x^{-\alpha-1} e^{-a_n^{\alpha}x^{-\alpha}} \left[\frac{\left(a_n^{\alpha}x^{-\alpha}\right)^{k-1}}{[k-1]!}\right].$$

Given the density, determining the expectation of the k^{th} order statistic is straightforward:

$$E\left[X_{n-k+1,n}\right] = \int_0^{\infty} x\alpha a_n^{\alpha} x^{-\alpha-1} e^{-a_n^{\alpha}x^{-\alpha}} \left[\frac{\left(a_n^{\alpha}x^{-\alpha}\right)^{k-1}}{[k-1]!}\right] dx.$$

Applying a change of variable $y = a_n^{\alpha}x^{-\alpha}$ we get

$$\begin{aligned} E\left[X_{n-k+1,n}\right] &= \frac{a_n}{k-1} \int_0^{\infty} y^{\frac{1}{\alpha}} y^{k-1} e^{-y} dy \\ &= \frac{a_n}{[k-1]!} \Gamma\left[k - \frac{1}{\alpha}\right]. \end{aligned}$$

Given the above expectation, determining the variance of the order statistics is a trivial matter:

$$\begin{aligned} var\left[X_{n-k+1,n}\right] &= E\left[X_{n-k+1,n}^2\right] - E\left[X_{n-k+1,n}\right]^2 \\ &= \frac{a_n^2}{[k-1]!}\Gamma\left[k-\frac{2}{\alpha}\right] - \left[\frac{a_n}{[k-1]!}\Gamma\left[k-\frac{1}{\alpha}\right]\right]^2 \end{aligned} \tag{15}$$

A.1 KS-distance metric

The purpose of the KS-distance metric is to find the optimal number of order statistics to estimate the tail index with the Hill estimator. This method achieves this by minimizing the distance between the empirical distribution and Pareto distribution over the quantile dimension. The starting point for locating t^* is the first-order term of the power expansion:

$$\Pr(X \leq x) = \mathrm{F}(x) = 1 - Ax^{-\alpha}[1 + o(1)]. \tag{16}$$

This function is identical to a Pareto distribution if the higher-order terms are ignored. By inverting (16), we get the quantile function

$$x = \left(\frac{\Pr(X \geq x)}{A}\right)^{\frac{1}{-\alpha}}. \tag{17}$$

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

$$q(k,t) = \left(\frac{\Pr(X > x)}{A}\right)^{\frac{1}{-\alpha}} = \left[\frac{t}{k}(x_{n-t+1,n})^{\widehat{\alpha}_t}\right]^{\frac{1}{\widehat{\alpha}_t}}. \tag{18}$$

Here k is the $(n-k)^{th}$ order statistic $X_{1,n} \leq X_{2,n} \leq ... \leq X_{n-k,n} \leq ... \leq X_{n,n}$ such that k/n comes closest to the probability level $\Pr(X > x)$.

Given the quantile estimator, the empirical quantile and the penalty function, we get

$$t^* = \arg\inf_t \left[\sup_k |x_{n-k,n} - q(k,t)|\right], \qquad for\ k = 1, ..., T, \tag{19}$$

where $T > t$ is the region over which the KS-distance metric is measured. Here $x_{n-k,n}$ is the empirical quantile and $q(k,t)$ is the estimated quantile from (18). This is done for different levels of t. The t, which produces the smallest maximum horizontal deviation along all the tail observations until T, is the t^* for the Hill estimator.

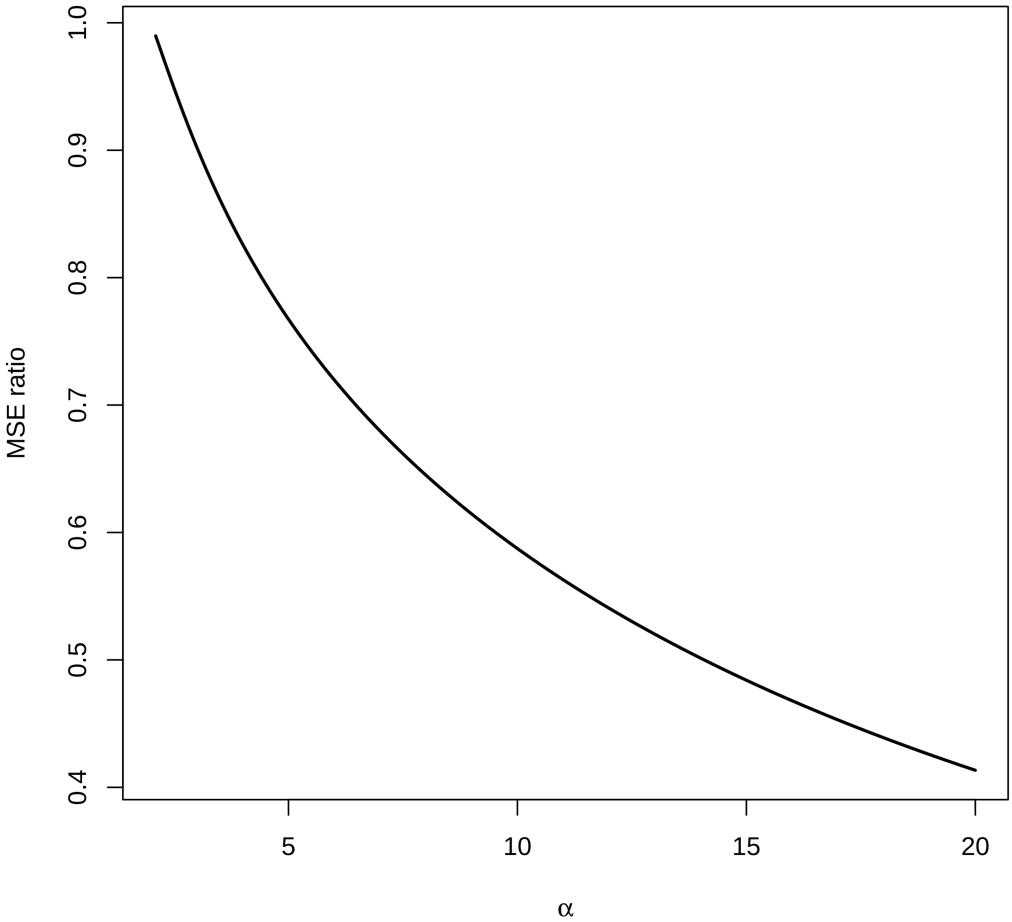


Figure 3: This figure displays the MSE ratio of the semi-parametric and the non-parametric worst-case estimator as a function of α. From (8), (10), (15) and (12) to construct the $MSE = Variance + Bais^2$. The MSE ratio is given by $^{MSE_{NP}}/_{(MSE_{SP}+MSE_{NP})}$. For the MSE of the semi-parametric estimator we choose $t = exp(2)$ and $\beta = 2$.

Table 3: CRSP data

	All	$\hat{\alpha}^i < 2$	$2 \leq \hat{\alpha}^i < 3$	$3 \leq \hat{\alpha}^i < 4$	$4 \leq \hat{\alpha}^i < 5$	$\hat{\alpha}^i \geq 5$
Mean	-0.028	-6.597	-1.224	0.648	1.518	2.203
Median	0.832	-4.382	-1.125	1.006	1.464	2.106
St. Dev.	2.798	6.893	2.902	1.567	1.458	0.995
Q0.01	-7.973	-27.236	-8.216	-4.072	-1.689	0.334
Q0.99	4.000	-0.051	4.311	3.321	4.257	4.097
Rank Sum test	0.001	0.000	0.115	0.000	0.000	0.000
N	884	22	315	383	151	13

This table reports summary statistics for the difference between the semi-parametric quantile estimator and the largest order statistic, $SP_i - NP_i$, for the **right tail** of US stocks. For the SP estimator, α is estimated with the Hill estimator. To determine the number of order statistics for the Hill estimator we use the KS-distance metric described in Danielsson et al. (2016). Column 1 reports the summary statistics of SP-NP for all stocks. The second column reports the summary statistics of the difference for the stock with $\hat{\alpha} \leq 2$. Columns 3 through 6 report the summary statistics for the stocks with the corresponding $\hat{\alpha}$. The first three rows report the mean, median and standard deviation of the corresponding basket. Q0.01 and Q0.99 report the 1% and 99% quantile for the distribution of $SP_i - NP_i$ for the different baskets of stocks. The next row reports the Wilcoxon signed-rank test p-value, testing non-parametrically for a difference in mean rank. N is the number of stocks in each basket. The individual stock data is from the CRSP dataset. The securities need to be traded on NYSE, AMEX, NASDAQ, and NYSE Arca exchanges over the period from 01-01-1995 to 01-01-2011. To be included, the average stock price over the sample needs to be above 5 dollars.

Table 4: IV Regression

	SP-NP			
	Left tail		Right tail	
	Stage 2	Stage 1	Stage 2	Stage 1
$\hat{\alpha}^i_{fitted}$	1.566***		0.824**	
	(0.269)		(0.401)	
$t_i/n * 100$	−0.091*	−0.092***	−0.294***	−0.102***
	(0.048)	(0.008)	(0.065)	(0.007)
Kurtosis		−0.014***		−0.004***
		(0.001)		(0.001)
Skewness		0.291***		−0.104***
		(0.031)		(0.026)
St. Dev.		11.502***		25.239***
		(2.640)		(2.823)
Constant	−4.902***	3.475***	−1.882	3.147***
	(1.004)	(0.064)	(1.482)	(0.071)
Observations	864	864	867	867
R^2	0.221	0.492	0.231	0.526
F Statistic	122.182***	208.147***	129.524***	239.171***

This table reports the regression results of the two-stage least-square estimation to instrument the estimated tail index, for US stocks. In the **first stage** we estimate $\hat{\alpha}^i = b_0 + b_1 * Kurtosis^i + b_2 * Skewness^i + b_3 * StDev^i + b_4 * (t_i/n * 100) + \varepsilon_i$. Here kurtosis, skewness and standard deviation are the moments of the return distribution of stock i. We exclude the top and bottom 5% of the observations in the measurement of the higher moments. In the **second stage** we estimate $SP_i - NP_i = c_0 + c_1 * \hat{\alpha}^i_{fitted} + c_2 * (t_i/n * 100) + \nu_i$. For the SP estimator, α^i is estimated with the Hill estimator. To determine the number of order statistics for the Hill estimator we use the KS-distance metric described in Danielsson et al. (2016). Here $t_i/n * 100$ is the percentage of order statistics from the total sample to estimate the Hill estimate. We include only stocks with $t_i > e^2$. The individual stock data is from the CRSP dataset. The securities need to be traded on NYSE, AMEX, NASDAQ, and NYSE Arca exchanges over the period from 01-01-1995 to 01-01-2011. To be included, the average stock price over the sample needs to be above 5 dollars.

Table 5: CRSP data: Fixed threshold

	SP-NP	
	Left tail	Right tail
$\hat{\alpha}^i$	0.481^{***}	0.849^{***}
	(0.092)	(0.138)
Constant	-4.352^{***}	-6.271^{***}
	(0.405)	(0.594)
Observations	889	889
R^2	0.030	0.041

This table reports the regression results for the difference between the semi-parametric quantile estimator and the largest order statistic, $SP_i - NP_i$, for US stocks. For the SP estimator, α^i is estimated with the Hill estimator. The number of order statistics is fixed at 0.25% of the total sample. The individual stock data is from the CRSP dataset. The securities need to be traded on NYSE, AMEX, NASDAQ, and NYSE Arca exchanges over the period from 01-01-1995 to 01-01-2011. To be included, the average stock price over the sample needs to be above 5 dollars.

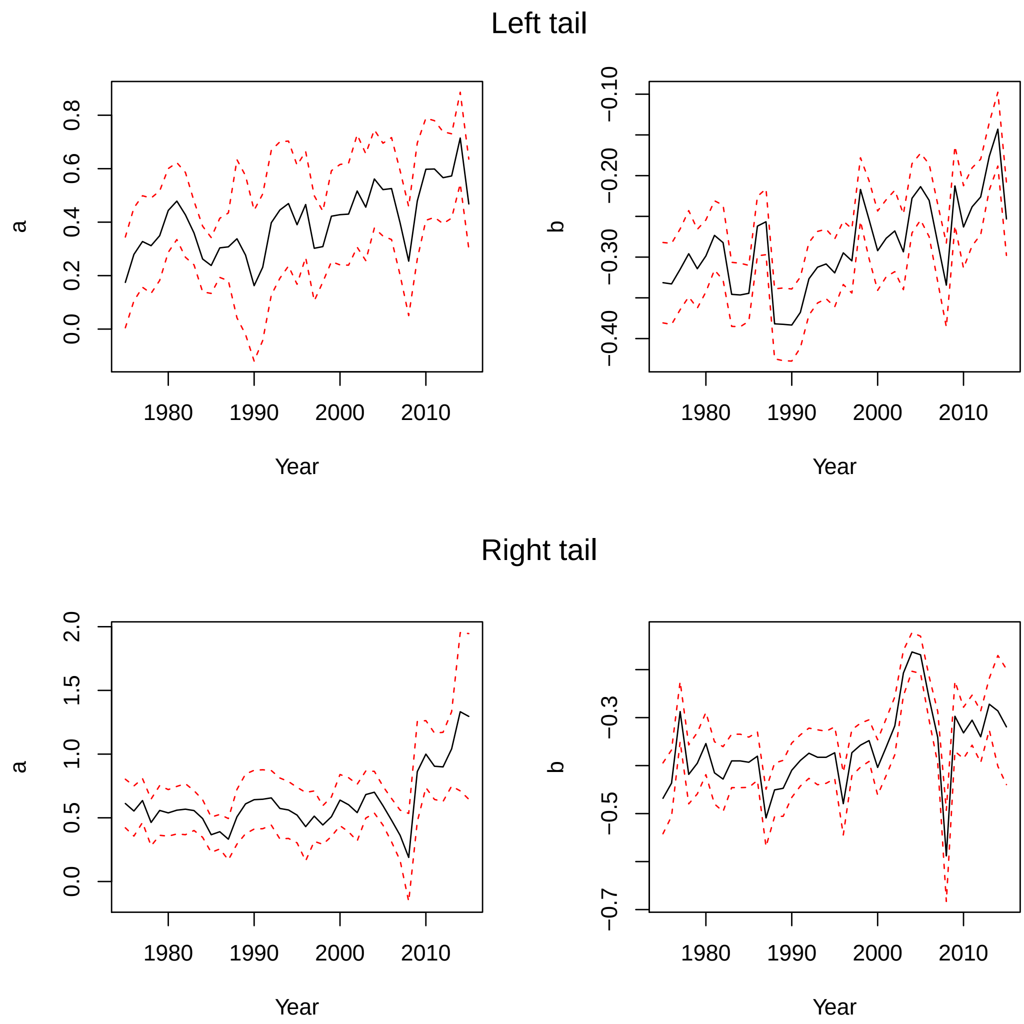


Figure 4: These figures depict the stability of the parameter estimates of Table 2. The solid lines are the parameter estimates over time and the dotted lines are their respective 95% error bounds. The two top and two bottom panels show the results for the left tail and right tail of the distribution, respectively. The left figures depict the results for the coefficient estimates $\hat{\alpha}^i$ and the right figures show the coefficient estimates ${}^{t_i}/_{n} * 100$. The regression equation, $SP_i - NP_i = c + a\ \hat{\alpha}^i + b\ {}^{t_i}/_{n} * 100 + e_i$, is re-estimated each year. In the re-estimation, the data from the preceding 10 years are used to proxy $SP_i - NP_i$, $\hat{\alpha}^i$, and ${}^{t_i}/_{n}$. We include only stocks with $t_i > exp(2)$. The individual stock data is from the CRSP dataset. The securities need to be traded on NYSE, AMEX, NASDAQ, and NYSE Arca exchanges over the period from 01-01-1965 to 01-01-2015. To be included, the average stock price over the sample needs to be above 5 dollars.