Taylor's law of fluctuation scaling for semivariances and higher moments of heavy-tailed data

Mark Brown[a,1,2], Joel E. Cohen[b,c,d,e,1,2], Chuan-Fa Tang[f,1,2], and Sheung Chi Phillip Yam[g,1,2]

[a]Department of Statistics, Columbia University, New York, NY 10027; [b]Laboratory of Populations, The Rockefeller University, New York, NY 10065-6399; [c]Earth Institute, Columbia University, New York, NY 10027; [d]Department of Statistics, Columbia University, New York, NY 10027; [e]Department of Statistics, University of Chicago, Chicago, IL 60637; [f]Department of Mathematical Sciences, University of Texas at Dallas, Richardson, TX 75080; and [g]Department of Statistics, Chinese University of Hong Kong, Hong Kong

Contributed by Joel E. Cohen, September 28, 2021 (sent for review April 28, 2021; reviewed by Svetlozar Rachev and Johannes Ruf)

We generalize Taylor's law for the variance of light-tailed distributions to many sample statistics of heavy-tailed distributions with tail index α in (0, 1), which have infinite mean. We show that, as the sample size increases, the sample upper and lower semivariances, the sample higher moments, the skewness, and the kurtosis of a random sample from such a law increase asymptotically in direct proportion to a power of the sample mean. Specifically, the lower sample semivariance asymptotically scales in proportion to the sample mean raised to the power 2, while the upper sample semivariance asymptotically scales in proportion to the sample mean raised to the power $(2-\alpha)/(1-\alpha) > 2$. The local upper sample semivariance (counting only observations that exceed the sample mean) asymptotically scales in proportion to the sample mean raised to the power $(2-\alpha^2)/(1-\alpha)$. These and additional scaling laws characterize the asymptotic behavior of commonly used measures of the risk-adjusted performance of investments, such as the Sortino ratio, the Sharpe ratio, the Omega index, the upside potential ratio, and the Farinelli–Tibiletti ratio, when returns follow a heavy-tailed nonnegative distribution. Such power-law scaling relationships are known in ecology as Taylor's law and in physics as fluctuation scaling. We find the asymptotic distribution and moments of the number of observations exceeding the sample mean. We propose estimators of α based on these scaling laws and the number of observations exceeding the sample mean and compare these estimators with some prior estimators of α.

stable law | semivariance | Pareto | Taylor's law | power law

Heavy-tailed nonnegative random variables with infinite moments, such as nonnegative stable laws with index α in (0,1), have theoretical and practical importance [e.g., Carmona (1), Feller (2), Resnick (3), and Samorodnitsky and Taqqu (4)]. Heavy-tailed nonnegative random variables with some or all infinite moments have been claimed to arise empirically in finance [operational risks in Nešlehová et al. (5)], economics [income distributions in Campolieti (6) and Schluter (7); returns to technological innovations in Scherer et al. (8) and Silverberg and Verspagen (9)], demography [city sizes in Cen (10)], linguistics [word frequencies in Bérubé et al. (11)], and insurance [economic losses from earthquakes in Embrechts et al. (12) and Ibragimov et al. (13)]. Partial reviews are in Carmona (1) and Ibragimov (14).

Brown et al. (15) (hereafter BCD) showed that when a random sample is drawn from a nonnegative stable law with index $\alpha \in (0,1)$, the sample variance is asymptotically (as the sample size n goes to ∞) proportional to the sample mean raised to a power that is an explicit function of α (Eqs. **11** and **13**). This relationship generalizes to stable laws with infinite moments a widely observed power-law relationship between the variance and the mean in families of distributions with finite population mean and finite population variance. This power-law relationship is commonly known as Taylor's law in ecology [Taylor (16, 17)] and as fluctuation scaling in physics [Eisler et al. (18)].

To the two ingredients combined by BCD (nonnegative stable laws with infinite moments and Taylor's law), this paper adds two more ingredients. We establish scaling relationships that generalize the usual Taylor's law, for light-tailed distributions, to many functions of the sample in addition to the variance, including all positive absolute and central moments, upper and lower semivariances, and several measures of risk-adjusted investment performance such as the Sortino, Sharpe, and Farinelli–Tibiletti ratios. In addition, based on these scaling relationships, we propose several estimators of the index α of a nonnegative stable law with infinite first moment.

Section 1 defines most of the sample functions studied here. Section 2 gives background on Taylor's law, semivariances, and nonnegative stable laws, including key prior results from BCD. Section 3 establishes that the lower sample semivariance, the upper sample semivariance, the local lower sample semivariance, and the local upper sample semivariance are asymptotically each a power of the sample mean with explicitly given exponents. These results are the core of the paper. When investment returns obey a nonnegative heavy-tailed law with index $\alpha \in (0,1)$, these results reveal the asymptotic behavior of the Sharpe ratio, the Sortino ratio, and the Farinelli–Tibiletti ratio. Section 4 extends these results to higher central and noncentral moments and various indices of volatility. Section 5 analyzes the number of observations from a stable law or an approximately stable (i.e., regularly varying) law that exceed the sample mean.

APPLIED MATHEMATICS

ECONOMIC SCIENCES

Significance

Many quantities are extremely large extremely rarely. Examples include income, wealth, financial returns, insurance losses, firm size, and city population size; earthquake magnitude, hurricane energy, tornado outbreaks, precipitation, and flooding; and pest outbreaks, infectious epidemics, and forest fires. When such a quantity is modeled as a nonnegative random variable with a heavy upper tail, the probability of an observation larger than some threshold falls as a small power (the "tail index") of the threshold. When the tail index is small enough, the mean and all higher moments of the random quantity are infinite. Surprisingly, the sample mean and the sample higher moments obey orderly scaling laws, which we prove and apply to estimating the tail index.

Author contributions: M.B., J.E.C., C.-F.T., and S.C.P.Y. designed research, performed research, contributed new reagents/analytic tools, analyzed data, and wrote the paper.

Reviewers: S.R., Texas Tech University; and J.R., London School of Economics.

The authors declare no competing interest.

See online for related content such as Commentaries.

[1]M.B., J.E.C., C.-F.T., and S.C.P.Y. contributed equally to this work.

[2]To whom correspondence may be addressed. Email: mb2484@columbia.edu, cohen@rockefeller.edu, chuan-fa.tang@utdallas.edu, or scpyam@sta.cuhk.edu.hk.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2108031118/-/DCSupplemental.

Published November 12, 2021.

Section 6 proposes and compares estimators of α by simulation. *SI Appendix* gives all proofs of results stated in the text and additional numerical simulations.

1. Preliminary

Let $\xrightarrow{d}$ mean "converges in distribution to." Let $\xrightarrow{p}$ mean "converges in probability to." Let $\xrightarrow{\text{a.s.}}$ mean "converges almost surely to."

Let X be a real-valued nonnegative random variable. Let n be a positive integer and assume that $n > 1$. For $i = 1, \ldots, n$, let X_i be independent and identically distributed as X. For any real $h \geq 0$, the hth (raw) sample moment is defined as

$$M'_h := \frac{1}{n}\sum_{i=1}^{n} X_i^h. \quad \textbf{[1]}$$

Thus, M'_1 is the sample mean. For any nonnegative integer h, the hth sample central moment is defined as

$$M_h := \frac{1}{n}\sum_{i=1}^{n}(X_i - M'_1)^h. \quad \textbf{[2]}$$

Clearly, $M_1 = 0$, and M_2 is the sample variance normalized by n. The sample variance normalized by $n-1$ is defined as

$$v_n := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - M'_1)^2. \quad \textbf{[3]}$$

Obviously, $v_n = M_2 n/(n-1)$ and $v_n/M_2 \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$.

The lower sample semivariance and the upper sample semivariance are defined as

$$v_n^- := \frac{1}{n-1}\sum_{i:X_i \leq M'_1}(X_i - M'_1)^2,$$
$$v_n^+ := \frac{1}{n-1}\sum_{i:X_i > M'_1}(X_i - M'_1)^2, \quad \textbf{[4]}$$

so that $v_n = v_n^- + v_n^+$. Define N_n^- as the number of values of X_i that do not exceed the sample mean and N_n^+ as the number of values of X_i that (strictly) exceed the sample mean:

$$N_n^- := \#\{i : X_i \leq M'_1\}, \quad N_n^+ := \#\{i : X_i > M'_1\}. \quad \textbf{[5]}$$

Then, $N_n^- N_n^+ > 0$ unless $X_i = M'_1$ for all $i = 1, \ldots, n$. The local lower sample semivariance and the local upper sample semivariance are defined only when $N_n^- > 0$ and $N_n^+ > 0$, respectively, as

$$v_n^{-*} := \frac{1}{N_n^-}\sum_{i:X_i \leq M'_1}(X_i - M'_1)^2,$$
$$v_n^{+*} := \frac{1}{N_n^+}\sum_{i:X_i > M'_1}(X_i - M'_1)^2. \quad \textbf{[6]}$$

The local upper sample semivariance v_n^{+*} is the more mathematically challenging sequence to analyze because it depends on the asymptotic behavior of the number of observations that exceed the sample mean. Our result, *Theorem 9*, may be of independent interest in the study of heavy-tailed distributions.

For the remainder of this article, we assume two restrictions on X without further restatement. First, we assume that X takes only nonnegative values. Second, to assure that $\mathbb{P}(N_n^+ = 0) = 0$, we assume that X is not atomic [i.e., for all real a, we assume that $\mathbb{P}(X = a) = 0$]. Then, $\mathbb{P}(N_n^+ = 0) = 0$ and conversely; for otherwise, if $\mathbb{P}(X = a) > 0$ for some a, then $\mathbb{P}(N_n^+ = 0) \geq \{\mathbb{P}(X = a)\}^n > 0$. Under the assumption that X is not atomic, $\mathbb{P}(N_n^- N_n^+ > 0) = 1$, and v_n^{-*} and v_n^{+*} are well defined almost surely (a.s.); also, $v_n^- = N_n^- v_n^{-*}/(n-1)$, $v_n^+ = N_n^+ v_n^{+*}/(n-1)$, and $v_n = (N_n^- v_n^{-*} + N_n^+ v_n^{+*})/(n-1)$ a.s. The assumption that X is not atomic also plays an important role in *Theorems 5* and *8(3), Remark 2,* and *Corollaries 6(3)* and *8*.

Alternatively, we could assume that X is not constant (i.e., not a degenerate random variable with all probability mass concentrated at a single value). If X is atomic but not a constant, then $\mathbb{P}(N_n^- N_n^+ > 0) \to 1$ as $n \to \infty$, but $\mathbb{P}(N_n^- N_n^+ > 0) \neq 1$. Nevertheless, similar asymptotic results could still be proved.

The infinite sequences of random variables defined in Eqs. **1** to **6** (one random variable for each $n = 1, 2, \ldots$) exist a.s., whether or not X has any finite moments. Our goal here is to show that, if X is a stable distribution (or an approximately stable distribution under *Definition 1*) with support $(0, \infty)$ and index $\alpha \in (0, 1)$, then as $n \to \infty$, the quantities in Eqs. **1** to **6** and other related quantities defined in section 3, when divided by some power b of the sample mean M'_1, converge in distribution, in probability or almost surely, depending on the case. Here, b may depend on α and on which quantity is being examined.

2. Background and Prior Results

Taylor's law [Taylor (16)] says that the sample variance v_n scales approximately in direct proportion to a nonzero power b (positive or negative) of the sample mean M'_1. Taylor's law is a widely confirmed empirical pattern in ecology and other sciences [Taylor (17)], nearly always with $b > 0$ and often with $b \in (1, 2)$. Taylor's law holds also for the mean and variance of some single-parameter probability distributions, in addition to holding for the sample mean and sample variance. For example, for varying values of the population mean μ, the population variance σ^2 varies according to Taylor's law $\sigma^2 = a\mu^b$ with $a = 1$, $b = 1$ for the Poisson distribution and $a = 1$, $b = 2$ for the exponential distribution.

The semivariances, especially the lower, have important applications in agricultural and financial economics [Berck and Hihn (19), Bond and Satchell (20), Hogan and Warren (21), Jin et al. (22), Liagkouras and Metaxiotis (23), Nantell and Price (24), Porter (25), Turvey and Nayak (26), and van de Beek et al. (27)]. We know no prior proofs that the sample semivariances of a nonnegative stable law satisfy Taylor's law.

Higher moments include skewness and kurtosis in statistics and the Farinelli–Tibiletti ratio in finance. Power-law scaling relationships for moments other than the sample variance are generalized Taylor's laws [Giometto et al. (28)]. Generalized Taylor's laws are less widely studied empirically or theoretically.

Every stable random variable X with support $(0, \infty)$ has Laplace transform [Feller (2), pp. 448–449]

$$\mathscr{L}(s) := \mathbb{E}(e^{-sX}) = e^{-(cs)^\alpha}, \quad \textbf{[7]}$$

for $s \geq 0$, $0 < \alpha < 1$, and $c > 0$. We say that $X \overset{d}{=} F(c, \alpha)$ when the distribution of X has Laplace transform Eq. **7**, and then we say that X has index α. We have $X \overset{d}{=} F(c, \alpha) \overset{d}{=} cF(1, \alpha)$. Such a heavy-tailed distribution has an infinite mean. Consequently, the sample mean, sample variance, sample semivariances, and sample higher moments are not estimators of population moments, and the normal central limit theorem does not apply.

If $X \overset{d}{=} F(c, \alpha)$ for some $0 < \alpha < 1$, $c > 0$, the survival function of X evaluated at $t \in (0, \infty)$ is defined as $\overline{F}(c, \alpha)(t) := 1 - F(c, \alpha)(t)$. By Feller (2, p. 448), if $0 < \alpha < 1$ and $c > 0$, then as $t \to \infty$,

$$\overline{F}(c, \alpha)(t) \Big/ \frac{c^\alpha t^{-\alpha}}{\Gamma(1-\alpha)} \to 1. \quad \textbf{[8]}$$

Many distributions on $(0, \infty)$ satisfy Eq. **8** but are not of the special form $F(c, \alpha)$ in Eq. **7**.

Definition 1. $X \overset{d}{\approx} F(c,\alpha)$ and $F_X \overset{d}{\approx} F(c,\alpha)$ both mean that a nonnegative random variable X has a distribution function F_X that satisfies Eq. **8**: that is, as $t \to \infty$,

$$\{1 - F_X(t)\}\Big/\frac{c^\alpha t^{-\alpha}}{\Gamma(1-\alpha)} \to 1. \qquad \textbf{[9]}$$

When Eq. **9** holds, we say that X is approximately stable.

For $\alpha \in (0,1)$ and real $g > \alpha$, $h > \alpha$, define

$$\alpha(g,h) := \frac{g-\alpha}{h-\alpha}, \qquad \alpha^* := \alpha(2,1) = \frac{2-\alpha}{1-\alpha}. \qquad \textbf{[10]}$$

If $g > h$, then $\alpha(g,h) > g/h$. Consequently, $\alpha^* > 2$. If $g < h$, then $\alpha(g,h) < g/h < 1$. Thus if, as we shall prove below, $\alpha(g,h)$ is the exponent b in Taylor's law for a stable nonnegative law with index $\alpha \in (0,1)$ and if $g \geq 2h$ or $g < h$, then the exponent b must fall outside the interval $(1,2)$ that is commonly (although not universally) observed in many ecological applications [Cohen et al. (29, 30)].

Among other results, BCD (ref. 15, p. 663, proposition 2) showed that if $X \overset{d}{=} F(1,\alpha)$, then as $n \to \infty$,

$$W_n := \frac{v_n}{M_1'^{\alpha^*}} \overset{d}{\to} W, \qquad \textbf{[11]}$$

where $\mathbb{E}(W_n) = 1-\alpha$, $Var(W_n) = \{\mathbb{E}(W_n)\}^2\{1 + 2\alpha/(n-1)\}$, and the limiting random variable W has $\mathbb{P}(0 < W < \infty) = 1$. W has a finite mean and a finite SD, both of which equal $1-\alpha$. Moreover, for all $h = 1, 2, \ldots, \mathbb{E}(W_n^h) \to \mathbb{E}(W^h)$. The second and third moments of W are

$$\mathbb{E}(W^2) = 2\{\mathbb{E}(W)\}^2, \quad \mathbb{E}(W^3) = \left(6 - \frac{\alpha}{(5-2\alpha)}\right)\{\mathbb{E}(W)\}^3, \qquad \textbf{[12]}$$

while for an exponentially distributed random variable Y, $\mathbb{E}(Y^3) = 6\{\mathbb{E}(Y)\}^3$ (ref. 15, p. 666).

For general $c > 0$ in Eq. **7**, BCD showed that $v_n/M_1'^{\alpha^*} \overset{d}{\to} c^{-\frac{\alpha}{1-\alpha}} W$, where W is the limiting random variable in Eq. **11**. Consequently, for any $c > 0$, BCD showed that as $n \to \infty$,

$$\frac{\log v_n}{\log M_1'} \overset{p}{\to} \alpha^*. \qquad \textbf{[13]}$$

Thus, for large n, with arbitrarily high probability, $(\log v_n)/(\log M_1')$ will be close to α^*, regardless of $c > 0$. This scaling relationship is an asymptotic form of Taylor's law with exponent $b = \alpha^* > 2$. BCD further argued without detailed proofs that $X \overset{d}{\approx} F(c,\alpha)$ satisfies Eq. **13**.

A common sample statistic used to compare the effectiveness of investments is the well-known Sharpe ratio [Sharpe (31)] $(M_1' - r_f)/v_n^{1/2}$ for the period rates of return of a security, where r_f is a zero-risk reference: for example, the London interbank offered rate. In signal processing, the Sharpe ratio (with $r_f = 0$) is a useful but biased estimator of the signal-to-noise ratio [Miller and Gehr (32)]. In statistics, the reciprocal of the Sharpe ratio (with $r_f = 0$) is called the coefficient of variation. If the period rate of return has a distribution $X \overset{d}{\approx} F(c,\alpha)$, where $0 < c < \infty$ and $0 < \alpha < 1$, then the Sharpe ratio converges in probability to zero as $n \to \infty$. Why? Eq. **11** implies that, as $n \to \infty$, $M_1'^{\alpha^*}/v_n \overset{d}{\to} 1/W$, so $M_1'^{\alpha^*/2}/v_n^{1/2} \overset{d}{\to} 1/W^{1/2}$. However, $M_1'^{\alpha^*/2} = M_1' \times M_1'^{(\alpha^*/2)-1}$, and because $\alpha^* > 2$ (as noted just after Eq. **10**), the second factor $M_1'^{(\alpha^*/2)-1}$ goes a.s. to ∞. Therefore, the Sharpe ratio $(M_1' - r_f)/v_n^{1/2}$ must converge in probability to zero. Asymptotically, for large n, the Sharpe ratio reveals no information about the distribution.

Inspired by Taylor's law in Eq. **13**, one may consider $\log(M_1' - r_f)/\log v_n$ as a modified financial ratio, which converges to $1/\alpha^* = (1-\alpha)/(2-\alpha)$ in probability. Because $(1-\alpha)/(2-\alpha)$ is decreasing in α over $(0,1)$, the smaller α is, the heavier the distribution, so the larger the risk. The original Sharpe ratio is quasiconcave, scale invariant, and distribution based [Eling et al. (33)]. The modified ratio is also distribution based and reveals the tail index α for large-enough n. Because of the logarithmic transformation, the modified ratio is not scale invariant. However, both numerator and denominator diverge to infinity. The effect of finite scaling becomes negligible for large sample sizes, and hence, the ratio is F_α-asymptotically scale invariant.* In other words, when $X \overset{d}{\approx} F(c,\alpha)$, the modified ratio is asymptotically invariant with respect to c. The modified Sharpe ratio is F_α-asymptotically quasiconcave.† The proof is in *SI Appendix*. Thus, asymptotically with large sample size n, the modified Sharpe ratio inherits all the properties of the original Sharpe ratio. We discuss this using semivariances and partial moments for the financial ratios in the following sections.

3. Taylor's Laws for Semivariances

A. Lower Semivariances and Sortino Ratio. The lower semivariance of any nonnegative random variable with infinite expectation is almost surely asymptotic to the square of the sample mean.

Theorem 1 (Taylor's law for the lower semivariance). *Let X be a nonnegative random variable with $\mathbb{E}(X) = \infty$. Then, as $n \to \infty$,*

$$\frac{v_n^-}{M_1'^2} \overset{a.s.}{\longrightarrow} 1. \qquad \textbf{[14]}$$

This theorem does not assume X is stable or approximately stable.

The Sortino ratio [Sortino and Price (34)] is another sample statistic used to compare the risks and rewards in some period of a set of investments such as individual equities, mutual funds, trading systems, or investment managers. It is defined as $(M_1' - r_f)/s_d$, where M_1' is the sample mean of the period rate of return X, r_f is a threshold or reference point or target return, the zero-risk rate of return or minimal acceptable return, which we take to be zero, and $s_d := (v_n^-)^{1/2}$ is the downside risk, equal to the square root of the lower sample semivariance v_n^- of the period rate of return [e.g., Sortino and Price (34) and Rollinger and Hoffman (35)]. Under our assumption that $\mathbb{P}(0 < X < \infty) = 1$, one might interpret X as the ratio of final price to initial price, so that $0 < X < 1$ would represent a loss, while $X > 1$ would represent a gain. The possible use of n instead of $n-1$ in the denominator of Eq. **4** is immaterial for large samples. Eq. **14** shows that if the period rate of return X is a nonnegative random variable with an infinite mean, then the Sortino ratio converges a.s. to one as $n \to \infty$. When the mean is infinite, asymptotically, for large n, the Sortino ratio reveals no information about the distribution.

Similar to our modified Sharpe ratio for heavy-tailed distributions, for the Sortino ratio, we consider the ratio between the logarithm of the sample mean minus r_f and the logarithm of the sample lower semivariance, namely $\log(M_1' - r_f)/\log v_n^-$. *Theorem 1* and Slutsky's theorem imply that a power law with exponent 2 relates the lower semivariance to the sample mean. So Taylor's law holds between the sample mean and the lower semivariance.

*F_α-asymptotic scale invariance is defined in *SI Appendix*, section D.

†F_α-asymptotic quasiconcavity is defined in *SI Appendix*, section D.

APPLIED MATHEMATICS

ECONOMIC SCIENCES

Corollary 1. *Let X be a nonnegative random variable with $\mathbb{E}(X) = \infty$. As $n \to \infty$,*

$$\frac{\log v_n^-}{\log M_1'} \xrightarrow{a.s.} 2. \qquad [15]$$

The modified Sortino ratio is F_α-asymptotically quasiconcave and F_α-asymptotically scale invariant, like the original Sortino ratio; proofs are in *SI Appendix*. However, from *Corollary 1*, the limiting value of the modified Sortino ratio is independent of the tail index α.

We now extend Taylor's law to the local lower semivariance v_n^{-*}. The local lower semivariance differs from the lower semivariance by a factor equal to the ratio N_n^-/n. We show that $N_n^-/n \to 1$ almost surely if $\mathbb{E}(X) = \infty$.

Lemma 1. *Let X be a nonnegative random variable with $\mathbb{E}(X) = \infty$. Then, with N_n^- defined in Eq.* **5**, *as $n \to \infty$,*

$$\frac{N_n^-}{n} \xrightarrow{a.s.} 1. \qquad [16]$$

Corollary 1 and *Lemma 1* imply that a power law with exponent 2 relates the local lower semivariance to the sample mean.

Corollary 2. *Let X be a nonnegative random variable with $\mathbb{E}(X) = \infty$. Then, as $n \to \infty$,*

$$\frac{\log v_n^{-*}}{\log M_1'} \xrightarrow{a.s.} 2. \qquad [17]$$

If X is approximately stable with infinite expectation, then *Lemma 1* and *Corollaries 1* and *2* imply further results that will be useful later for studying the local upper semivariance and upper semivariance.

Corollary 3. *Let $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$. Let $\alpha^* := (2 - \alpha)/(1 - \alpha)$ as defined in Eq.* **10**. *Then, as $n \to \infty$,*

$$\frac{v_n^-}{M_1'^{\alpha^*}} \xrightarrow{a.s.} 0 \quad \text{and} \quad \frac{v_n^{-*}}{M_1'^{\alpha^*}} \xrightarrow{a.s.} 0. \qquad [18]$$

B. Upper Semivariances. Although the asymptotic values of the ratios in Eqs. **15** and **17** are both two, which is independent of α, if one replaces the lower or local lower semivariances by the upper or local upper semivariances, respectively, Taylor's law continues to hold, and it depends on α.

Theorem 2. *Let $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$. Then, as $n \to \infty$,*

$$\frac{\log v_n^+}{\log M_1'} \xrightarrow{p} \alpha^* \quad \text{and} \quad \frac{\log v_n^{+*}}{\log M_1'} \xrightarrow{p} \alpha^* + \alpha = \frac{2 - \alpha^2}{1 - \alpha}. \qquad [19]$$

Inspired by Taylor's law in Eq. **19**, one may consider ratios between the logarithm of the sample mean minus r_f and the logarithm of either the sample upper or local upper semivariances, namely $\log(M_1' - r_f)/\log v_n^+$ and $\log(M_1' - r_f)/\log v_n^{+*}$, respectively, which converge in probability to $1/\alpha^* = (1 - \alpha)/(2 - \alpha)$ and $(1 - \alpha)/(2 - \alpha^2)$, respectively. Because $(1 - \alpha)/(2 - \alpha)$ and $(1 - \alpha)/(2 - \alpha^2)$ are both decreasing in α, the smaller α is, the heavier the distribution is, and the larger these ratios are asymptotically. The asymptotic properties and proofs are in *SI Appendix, Proposition* D.3.

4. Fluctuation Scaling for Higher Moments

In this section, we show that the sample higher moments are proportional to a power of the sample mean. These relations imply power-law relations between sample higher moments used in financial ratios such as the Farinelli–Tibiletti ratio (36).

A. Higher Sample Moments, Skewness, and Kurtosis.

Theorem 3. *If $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, and $h > \alpha$, then, as $n \to \infty$,*

$$\frac{M_h'}{(M_1')^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1 - \alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}},$$

where the random vector (U_h, V) has the joint Laplace transform

$$\mathbb{E}(e^{-sU_h - tV}) = \exp\left\{-\int_0^\infty \{r_h(y, s, t)\}^{-\alpha} e^{-y}\, dy\right\},$$

for $s, t, y > 0$, and $r_h(y, s, t)$ is the unique positive root of the equation $sx^h + tx - y = 0$.

The ratio in *Theorem 3* may not be a practically useful financial ratio since α is usually unknown. However, the following *Theorem 4* and its corollaries heavily depend on it. The following remark uses the joint moment-generating function to give the marginal distributions of U_h and V.

Remark 1. In the joint Laplace transform defined in *Theorem 3*, if we set $t = 0$, then $r_h(y, s, 0) = (y/s)^{1/h}$ and

$$\mathbb{E}(e^{-sU_h}) = \exp\left\{-\int_0^\infty \left\{(y/s)^{1/h}\right\}^{-\alpha} e^{-y}\, dy\right\}.$$

Hence, U_h follows the distribution $F(\{\Gamma(1 - \alpha/h)\}^{h/\alpha}, \alpha/h)$. On the other hand, if we set $s = 0$, then $r_h(y, 0, t) = y/t$ and

$$\mathbb{E}(e^{-tV}) = \exp\left\{-\int_0^\infty \{(y/t)\}^{-\alpha} e^{-y}\, dy\right\}.$$

Hence, V follows the distribution $F(\{\Gamma(1 - \alpha)\}^{1/\alpha}, \alpha)$.

These results follow Albrecher et al. (ref. 37, remark 2.1) by the arguments in their proof. The following theorem shows that Taylor's law holds for raw moments.

Theorem 4. *If $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, $h_1 > \alpha$, and $h_2 > \alpha$, then as $n \to \infty$,*

$$\frac{\log M_{h_2}'}{\log M_{h_1}'} \xrightarrow{p} \alpha(h_2, h_1).$$

In particular, for $h > \alpha$, as $n \to \infty$,

$$\frac{\log M_h'}{\log M_1'} \xrightarrow{p} \alpha(h, 1).$$

For a positive integer $h > 1$, the ratio between the central moment M_h and the $\alpha(h, 1)$ power of the sample mean M_1' converges to a distribution given in *Corollary 4*.

Corollary 4. *If $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, and $h > 1$ is a positive integer, then as $n \to \infty$,*

$$\frac{M_h}{(M_1')^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1 - \alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}},$$

where the random vector (U_h, V) is specified in Theorem 3.

Theorem 5. *If $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, and $h > 1$ is a positive integer, then as $n \to \infty$,*

$$\frac{\log |M_h|}{\log M_1'} \xrightarrow{p} \alpha(h, 1).$$

For any positive integers $h_1 > 1$ and $h_2 > 1$, as $n \to \infty$,

$$\frac{\log |M_{h_2}|}{\log |M_{h_1}|} \xrightarrow{p} \alpha(h_2, h_1).$$

For the raw moments, we have generalized *Theorem 3* for the ratio of two raw moments with orders both larger than α.

Theorem 6. *If* $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, *and both* $h_1, h_2 > \alpha$, *then as* $n \to \infty$,

$$\frac{M'_{h_2}}{(M'_{h_1})^{\alpha(h_2, h_1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h_2-h_1}{h_1-\alpha}} \frac{U_{h_2}}{(U_{h_1})^{\alpha(h_2,h_1)}},$$

where (U_{h_1}, U_{h_2}) *has the joint Laplace transform*

$$\mathbb{E}(e^{-sU_{h_2} - tU_{h_1}}) = \exp\left\{-\int_0^\infty \{r_{h_2,h_1}(y, s, t)\}^{-\alpha} e^{-y}\, dy\right\},$$

with $y > 0$, $s > 0$, $t > 0$, *and* $r_{h_2,h_1}(y, s, t)$ *is the unique positive root x of* $sx^{h_2} + tx^{h_1} - y = 0$. *Moreover, as* $n \to \infty$,

$$\frac{\log M'_{h_2}}{\log M'_{h_1}} \xrightarrow{p} \alpha(h_2, h_1).$$

Corollary 5. *If* $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, *and* $h_2 \geq h_1 > 1$ *are positive integers, then as* $n \to \infty$,

$$n^{\frac{h_1-h_2}{h_1}} \frac{M_{h_2}}{(M_{h_1})^{h_2/h_1}} \xrightarrow{d} \frac{U_{h_2}}{(U_{h_1})^{h_2/h_1}},$$

where (U_{h_1}, U_{h_2}) *is defined in Theorem 6.*

Remark 2. From *Corollary 5*, it is clear that the skewness $M_3/(v_n)^{3/2}$ and the kurtosis $M_4/(v_n)^2$ diverge to infinity, yet the scaled skewness and the scaled kurtosis have distributions, asymptotically as $n \to \infty$,

$$\frac{M_3}{n^{1/2}(v_n)^{3/2}} \xrightarrow{d} \frac{U_3}{(U_2)^{3/2}} \quad \text{and} \quad \frac{M_4}{n(v_n)^2} \xrightarrow{d} \frac{U_4}{(U_2)^2},$$

where the joint distributions of (U_2, U_3) and (U_2, U_4) are defined in *Theorem 6*. The limiting distribution of $M_4/\{n(v_n)^2\}$ matches the result derived in Cohen et al. (ref. 38, equation 3.9). Moreover, by Slutsky's theorem, as $n \to \infty$,

$$\frac{\log|M_3|}{\log[(v_n)^{3/2}]} \xrightarrow{p} \frac{2}{3}\alpha(3, 2) \quad \text{and} \quad \frac{\log M_4}{\log[(v_n)^2]} \xrightarrow{p} \frac{1}{2}\alpha(4, 2).$$

B. Central Lower and Local Lower Partial Moments.

Definition 2. Define $c_+ := \max\{0, c\}$ for $c \in \mathbb{R}$. For $h > 0$, define

$$M_h^- := \frac{1}{n}\sum_{i=1}^n [(M'_1 - X_i)_+]^h, \qquad M_h^{-*} := \frac{nM_h^-}{N_n^-}.$$

Theorem 7. *Let* X *be a nonnegative random variable with* $\mathbb{E}(X) = \infty$, *and let* $h > 0$. *Then, as* $n \to \infty$,

$$M_h^-/(M'_1)^h \xrightarrow{a.s.} 1 \quad \textit{and} \quad \log M_h^- - h \log M'_1 \xrightarrow{a.s.} 0.$$

Corollary 6. *Let* X *be a nonnegative random variable with* $\mathbb{E}(X) = \infty$. *Then, as* $n \to \infty$,

1) $M_1^-/M'_1 \xrightarrow{a.s.} 1$;
2) for $h > 1$, $M_h^-/(M'_1)^{\alpha(h,1)} \xrightarrow{a.s.} 0$;
3) for $h > 0$,

$$\frac{\log M_h^-}{\log M'_1} \xrightarrow{a.s.} h \quad \textit{and} \quad \frac{\log M_h^{-*}}{\log M'_1} \xrightarrow{a.s.} h.$$

C. Central Upper Moments and Local Upper Moments.

Definition 3. For $h > 0$, define the hth central upper moments and central local upper moments:

$$M_h^+ := \frac{1}{n}\sum_{i=1}^n [(X_i - M'_1)_+]^h, \qquad M_h^{+*} := \frac{nM_h^+}{N_n^+}.$$

Theorem 8 (central upper moments). *Let* $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$. *Then, as* $n \to \infty$,

1) for $0 < h < 1$, $M_h^+/(M'_1)^h \xrightarrow{p} 0$;
2) for $h \geq 1$,

$$\frac{M_h^+}{(M'_1)^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}},$$

where the random vector (U_h, V) *has the joint Laplace transform defined in Theorem 3;*
3) for $h \geq 1$,

$$\frac{\log M_h^+}{\log M'_1} \xrightarrow{p} \alpha(h, 1) \quad \textit{and} \quad \frac{\log M_h^{+*}}{\log M'_1} \xrightarrow{p} \frac{h-\alpha^2}{1-\alpha}. \qquad [20]$$

D. Omega Index, Upside Potential Ratio, and Farinelli–Tibiletti Ratio. Farinelli–Tibiletti (36) extended the Sharpe ratio to an index including asymmetrical information on the volatilities above and below the benchmark $r_f \in \mathbb{R}$. Their index Φ_{FT} is defined by

$$\Phi_{\mathrm{FT}}(r_f, p, q) := \frac{[\mathbb{E}[(X - r_f)_+]^p]^{1/p}}{[\mathbb{E}[(r_f - X)_+]^q]^{1/q}}.$$

The Omega index, introduced by Cascon et al. (39), is $\Phi_{\mathrm{FT}}(r_f, 1, 1)$ with $p = q = 1$. The upside potential index, introduced by Sortino et al. (40), is $\Phi_{\mathrm{FT}}(r_f, 1, 2)$ with $p = 1$ and $q = 2$. The ratio $\Phi_{\mathrm{FT}}(r_f, p, q)$ may not be well defined since the expectations may not exist for the heavy-tailed distributions. However, one can define an empirical version of the Farinelli–Tibiletti ratio by

$$\Phi_{\mathrm{FT}}^n(r_f, p, q) := \frac{[\frac{1}{n}\sum_{i=1}^n [(X_i - r_f)_+]^p]^{1/p}}{[\frac{1}{n}\sum_{i=1}^n [(r_f - X_i)_+]^q]^{1/q}}.$$

The following corollary shows that both $\Phi_{\mathrm{FT}}^n(r_f, p, q)$ and $\Phi_{\mathrm{FT}}^n(M'_1, p, q)$ converge to ∞ in probability.

Corollary 7. *If* $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, $r_f > 0$, $p > 1$, *and* $q > 1$, *then as* $n \to \infty$, $\Phi_{\mathrm{FT}}^n(r_f, p, q) \xrightarrow{p} \infty$ *and* $\Phi_{\mathrm{FT}}^n(M'_1, p, q) \xrightarrow{p} \infty$.

A modification of the usual Farinelli–Tibiletti ratio might have the ratio of the logarithm of the numerator to the logarithm of the denominator in $\Phi_{\mathrm{FT}}(r_f, p, q)$. However, for a fixed $r_f > 0$, the numerator converges to infinity in probability, while the denominator is bounded above with probability one. Therefore, this ratio diverges to infinity.

We propose as an alternative to the Farinelli–Tibiletti ratio:

$$\Phi_{\mathrm{FTlog}}(p, q) := p \log M_q^-/(q \log M_p^+),$$

which is the ratio of the logarithm of the numerator to that of the denominator in $\Phi_{\mathrm{FT}}(M'_1, p, q)$. The following corollary describes generalized Taylor's laws for the ratio of the logarithm of the upper central moment to the logarithm of the lower central moment.

Corollary 8. *If* $X \overset{d}{\approx} F(1, \alpha)$, $0 < \alpha < 1$, $p \geq 1$, *and* $q \geq 1$, *then as* $n \to \infty$,

$$\frac{\log M_p^+}{\log M_q^-} \xrightarrow{p} \frac{p-\alpha}{q(1-\alpha)}.$$

Corollary 8 implies that

$$\Phi_{\mathrm{FTlog}}(p,q) \xrightarrow{p} p(1-\alpha)/(p-\alpha),$$

which is decreasing in α for $p \geq 1, q \geq 1$. Therefore, the smaller α is, the heavier the distribution is, and the larger the risk is. Our modified Farinelli–Tibiletti ratio $\Phi_{\mathrm{FTlog}}(p,q)$ is asymptotically scale invariant and distribution based, like the original Farinelli–Tibiletti ratio, and satisfies F_α-asymptotic quasiconcavity (*SI Appendix*).

5. Number of Observations Exceeding Sample Mean of Stable Law

A. Asymptotic Distributions and Moments of N_n^+/n^α. In a sample of size *n* from an approximately stable law with index $\alpha \in (0,1)$, asymptotically the number of observations above the sample mean scales as n^α and has a distribution given by *Theorem 9*. To prove this result, we use Einmahl (ref. 41, corollary 2.1) together with *SI Appendix, Lemma C.1*.

Theorem 9. *If* $X \overset{d}{\approx} F(1,\alpha)$, $0<\alpha<1$, *and* $U \overset{d}{=} F(1,\alpha)$, *then as* $n \to \infty$,

$$\frac{N_n^+}{n^\alpha} \xrightarrow{d} V := \frac{U^{-\alpha}}{\Gamma(1-\alpha)}.$$

The asymptotic moments of N_n^+/n^α are the moments of V defined in *Theorems 9* and *10*.

Theorem 10. *Let* $U \overset{d}{=} F(1,\alpha)$, $0<\alpha<1$, $V := U^{-\alpha}/\Gamma(1-\alpha)$, *and* $\varepsilon \overset{d}{=} \mathrm{Exp}(1)$ *(an exponential random variable with mean and parameter 1), where* ε *is independent of U.*

1) $U^{-\alpha}\varepsilon^{\alpha} \overset{d}{=} \mathrm{Exp}(1)$.
2) *For integer* $K>0$,

$$\mathbb{E}[U^{-K\alpha}] = \frac{K!}{\Gamma(1+K\alpha)},$$

$$\mathbb{E}[V^K] = \frac{K!}{\Gamma(1+K\alpha)\{\Gamma(1-\alpha)\}^K}.$$

Specifically, when $K=1$, *then* $\mathbb{E}[U^{-\alpha}] = \{\Gamma(1+\alpha)\}^{-1}$ *and* $\mathbb{E}[V] = \{\Gamma(1+\alpha)\Gamma(1-\alpha)\}^{-1}$; *when* $K=2$, *then* $\mathbb{E}[U^{-2\alpha}] = 2\{\Gamma(1+2\alpha)\}^{-1}$, $\mathbb{E}[V^2] = 2\{\Gamma(1+2\alpha)\{\Gamma(1-\alpha)\}^2\}^{-1}$. *Hence*

$$\mathrm{Var}(U^{-\alpha}) = \frac{2}{\Gamma(1+2\alpha)} - \frac{1}{\{\Gamma(1+\alpha)\}^2},$$

$$\mathrm{Var}(V) = \frac{1}{\{\Gamma(1-\alpha)\}^2}\mathrm{Var}(U^{-\alpha}).$$

3) $\mathrm{SD}(V) < \mathbb{E}[V]$. *For example, when* $\alpha = 1/2$, $\mathbb{E}[V^2] = 2/\pi$, $\mathbb{E}[V] = 2/\pi$, $\mathrm{Var}(V) = \frac{2}{\pi}(1-\frac{2}{\pi})$. *Numerically,* $\mathrm{SD}(V) \approx 0.48097$, $\mathbb{E}[V] \approx 0.63662$, *where here "≈ " means the numerical approximation is inexact.*
4) *For* $K \geq 2$, $\mathbb{E}[V^K] < K!(\mathbb{E}[V])^K$.
5) $V \leq_{st} \varepsilon$ *[i.e., by the definition of the stochastic ordering* $\leq_{st}$, $\mathbb{P}(V>t) \leq \mathbb{P}(\varepsilon > t)$ *for all* $t \in \mathbb{R}$*].*

Part 1 of *Theorem 10* is not well known. The moment results in part 2 of *Theorem 10* are derived using fractional calculus by Wolfe (42). Because the logarithm of the moment-generating function of a nonnegative random variable is a convex function of the moment (by Artin's theorem) [Marshall and Olkin (ref. 43, theorem B.8)], it follows that $\log\mathbb{E}(U^{-x\alpha}) = \log\Gamma(1+x) - \log\mathbb{E}(W^x)$ is concave in $x \in [1,\infty)$.

The distribution of $U^{-\alpha}$ approximates the standard exponential distribution $\mathrm{Exp}(1)$ when $\alpha \to 0$.

Corollary 9. *Let* $U \overset{d}{=} F(1,\alpha)$. *Then, as* $\alpha \to 0$,

$$U^{-\alpha} \xrightarrow{d} \mathrm{Exp}(1).$$

6. Numerical Experiments

A. Tail Estimators. The preceding results describe the asymptotic ratio of the logarithm of the sample mean to the logarithm of various forms of the sample variance, such as the ordinary sample variance v_n, the upper semivariance v_n^+, the local upper semivariance v_n^{+*}, and the lower semivariance v_n^- when a random sample is from an approximately stable $F(1,\alpha)$ satisfying Eq. **9**. Most of these ratios (apart from that for the lower semivariance) depend asymptotically only on α. Based on these results, we propose estimators of the index α. We define the ratios R_1, R_2, R_3, and R_L where

$$R_1 := \frac{\log v_n}{\log M_1'} \xrightarrow{p} \frac{2-\alpha}{1-\alpha}, \qquad R_2 := \frac{\log v_n^+}{\log M_1'} \xrightarrow{p} \frac{2-\alpha}{1-\alpha},$$

$$R_3 := \frac{\log v_n^{+*}}{\log M_1'} \xrightarrow{p} \frac{2-\alpha^2}{1-\alpha}, \qquad R_L := \frac{\log v_n^-}{\log M_1'} \xrightarrow{\text{a.s.}} 2.$$

The results generalize to $F(c,\alpha)$ for $c>0$ because as noted after Eq. **9**, $X/c \overset{d}{\approx} F(1,\alpha)$ if and only if $X \overset{d}{\approx} F(c,\alpha)$ for $c>0$. Applying the continuous mapping theorem to the above results for the variance, the upper semivariance, and the local upper semivariance yields three consistent estimators of α:

$$B_1 := \frac{2-R_1}{1-R_1}, \qquad B_2 := \frac{2-R_2}{1-R_2},$$

$$B_3 := \frac{R_3 - \sqrt{R_3^2 - 4(R_3-2)}}{2}.$$

The Hill estimator [Hill (44)] is a traditional tail-index estimator, which requires the largest k observations where $k \to \infty$ and $k/n \to 0$ as $n \to \infty$. However, k depends on the unknown parameters such as α and the series representation of the survival function [Hall (45)]. In practice, the number k is based on the "stable" point in the Hill plot, which may not always be available. Gomes and Guillou (46) give a comprehensive review.

Theorem 9 implies that N_n^+/n converges to zero in probability, which motivates the choice of $k = N_n^+ + 1$ in the Hill estimator:

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

where $X_{(i)}$ is the *i*th-order statistic, $1 \leq i \leq n$. We evaluate this choice of $k = N_n^+ + 1$ in the Hill estimator, denoted by HI.N, numerically. We also replace the smallest $(n-k)$ order statistics in the original Hill estimator by the sample mean M_1' to obtain a new Hill-type estimator:

$$\mathrm{HI.M} := \left(\frac{1}{N_n^+}\sum_{X_i > M_1'} \log(X_i/M_1')\right)^{-1}.$$

From Bergström (47), the survival function of the stable law for $0<\alpha<1$ is

$$\begin{aligned}\overline{F}(1,\alpha)(x) &= \int_x^\infty -\frac{1}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^k}{k!}(\sin\pi\alpha k)\frac{\Gamma(ak+1)}{t^{ak+1}}\,dt \\ &= \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k!}(\sin\pi\alpha k)\frac{\Gamma(ak)}{x^{ak}} \\ &= Cx^{-\alpha}[1 + Dx^{-\alpha} + o(x^{-\alpha})],\end{aligned}$$

where $C>0$ and $D \neq 0$. From Hall (45), it is optimal to choose k tending to infinity at a rate of order $n^{2\alpha/(2\alpha+\alpha)} = n^{2/3}$. We also consider this choice $k = n^{2/3}$ for another Hill-type estimator, denoted by HI.Opt, and we compare the behavior with other estimators.

Table 1. Bias ($\times 10^3$; average of [estimate minus true α]) for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with sample size $n = 10^4$ from $F(1, \alpha)$

α	B1	B2	B3	HI.N	HI.M	HI.Opt	MHB3
0.1	−5.24	−3.87	−3.00	10.25	135.16	−0.92	−5.82
0.2	−11.96	−6.88	−3.79	−9.31	73.52	−1.73	−9.65
0.3	−19.43	−8.55	−2.38	−25.60	30.89	−2.05	−12.03
0.4	−27.72	−9.75	0.63	−32.82	4.87	−1.54	−13.44
0.5	−35.03	−8.91	5.96	−29.40	−5.30	1.42	−12.56
0.6	−43.76	−10.44	9.41	−24.21	−8.21	6.67	−10.26
0.7	−50.27	−11.28	12.19	−10.06	0.13	19.37	−3.26
0.8	−53.49	−12.55	11.48	31.58	37.82	51.80	7.30
0.9	−50.31	−13.69	5.46	204.26	208.27	153.44	5.45

Table 2. MSE ($\times 10^3$) (mean squared [estimate minus true α]) for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with sample size $n = 10^4$ from $F(1, \alpha)$

α	B1	B2	B3	HI.N	HI.M	HI.Opt	MHB3
0.1	0.14	0.15	0.11	2.61	20.06	0.02	0.10
0.2	0.53	0.58	0.35	4.73	9.13	0.09	0.31
0.3	1.13	1.23	0.71	7.33	6.31	0.19	0.57
0.4	1.86	1.96	1.16	9.15	6.39	0.34	0.85
0.5	2.60	2.66	1.76	9.12	6.76	0.54	1.15
0.6	3.47	3.20	2.32	7.93	6.13	0.84	1.53
0.7	4.15	3.38	2.60	6.46	5.30	1.52	1.94
0.8	4.32	3.05	2.28	6.97	6.67	4.35	2.13
0.9	3.59	2.10	1.33	57.51	58.26	26.36	1.33

In our simulations, we generate 10^4 independent random samples, each with sample size *n*, from $F(1, \alpha)$ by using the rstable function from the R package stabledist with arguments for the tail-index parameter alpha $= \alpha$, the skewness parameter beta $= 1$, the scale parameter gamma $= |1 - i\tan(\pi\alpha/2)|^{-1/\alpha}$, the location parameter delta $= 0$, and parameterization pm $= 1$. Setting pm $= 1$ specifies that we use the parameterization of stable laws in Samorodnitsky and Taqqu (4). For each random sample, we calculate the six estimators B_1, B_2, B_3, HI.N, HI.M, and HI.Opt. Then, we estimate the bias as the average of the 10^4 differences between each estimator of α and the true α. We estimate the mean squared error (MSE) as the average of 10^4 squared differences between each estimator of α and the true α.

In Table 1 for bias and Table 2 for MSE, the sample size is $n = 10^4$. According to the bias estimates in Table 1, B_1 tends to underestimate α, while B_2 and B_3 reduce the bias from B_1 by introducing the upper semivariance, which focuses more on larger numbers. B_3 has smaller bias than B_2 for most of the α except $\alpha = 0.7$ and 0.8. In Table 2, B_3 has smaller MSE than B_1 and B_2. Estimators HI.N and HI.M do not perform as well as B_3.

The estimator HI.Opt with the optimal choice of $k = n^{2/3}$ for the Hill estimator has the smallest bias, when $\alpha \leq 0.6$, and MSE, when $\alpha \leq 0.7$. However, B_3 from Taylor's law of the local semivariance has better performance, especially much smaller bias, than HI.Opt for $\alpha \geq 0.8$. Since HI.Opt tends to overestimate α, especially when $\alpha \geq 0.7$, we defined the estimator MHB3 to be the minimum of B_3 and HI.Opt. This MHB3 not only reduces the bias dramatically but also improves the MSE of B_3 for α close to 1.

The advantages of B_3 and MHB3 gradually vanish when sample size increases because $k = n^{2/3}$ is an asymptotically optimal choice. However, for sample sizes smaller than 10^4, B_3 and MHB3 can improve HI.Opt even more. More comparisons are in *SI Appendix* for sample sizes $n = 10^2, 10^3$, and 10^5. On the other hand, although the behavior of B_1, B_2, and B_3 depends on *c* in $F(c, \alpha)$, one sees similar patterns in bias and MSE. B_3 and MHB3 still have better bias and MSE for $\alpha \geq 0.8$ for small sample sizes. More comparisons are in *SI Appendix* for $F(2, \alpha)$ and $F(0.5, \alpha)$.

Tables in *SI Appendix* also show that both bias and MSE decrease when sample size increases, as expected of consistent estimators and as proved in *Corollary 1*.

B. Asymptotic Distribution of N_n^+/n^α. To illustrate *Theorem 9*, we generate 10^3 independent random samples from $F(1, \alpha)$ with sample size $n = 10^6$ and calculate N_n^+/n^α for each random sample. We use the 10^3 values of N_n^+/n^α to estimate the distribution of N_n^+/n^α. To estimate the distribution of $U^{-\alpha}/\Gamma(1-\alpha)$, we generate 10^3 independent random values $U_1, \ldots, U_{10^3}$ from $F(1, \alpha)$ and calculate the corresponding $U_i^{-\alpha}/\Gamma(1-\alpha)$ for $i = 1, \ldots, 10^3$. Then, we use the 10^3 values of $U_i^{-\alpha}/\Gamma(1-\alpha)$ to estimate the distribution of $U^{-\alpha}/\Gamma(1-\alpha)$. The histograms and

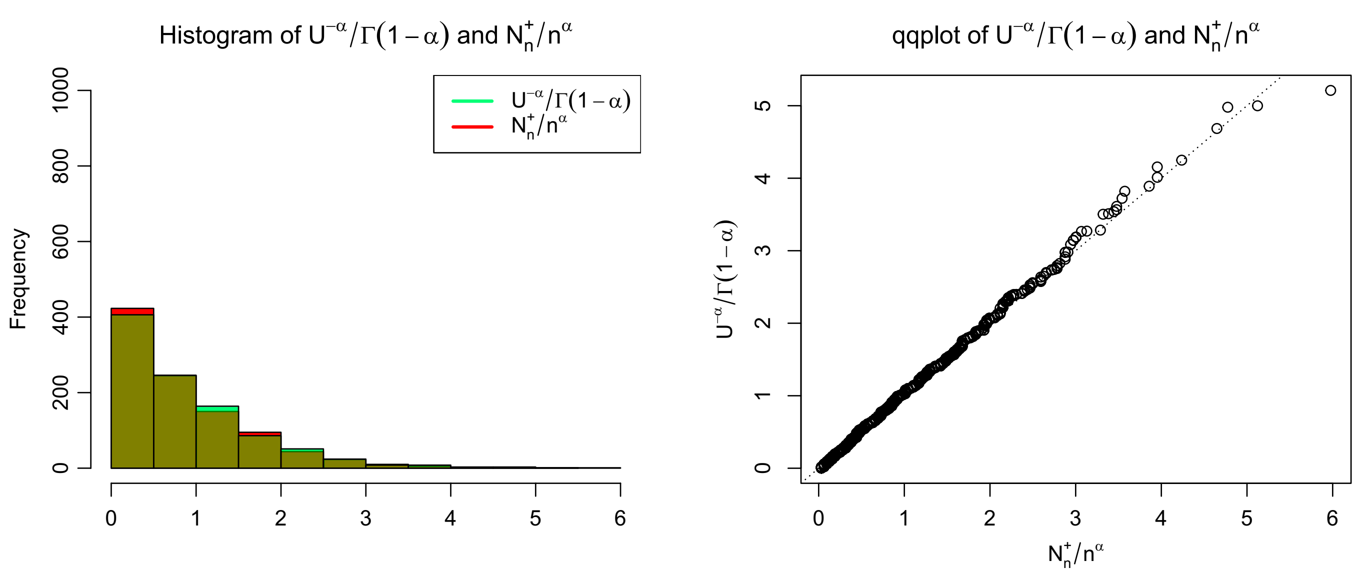

Fig. 1. Histogram and quantile–quantile plot of N_n^+/n^α and $U^{-\alpha}/\Gamma(1-\alpha)$ for $\alpha = 0.25$. The *P* value of the KS test is 0.1995.

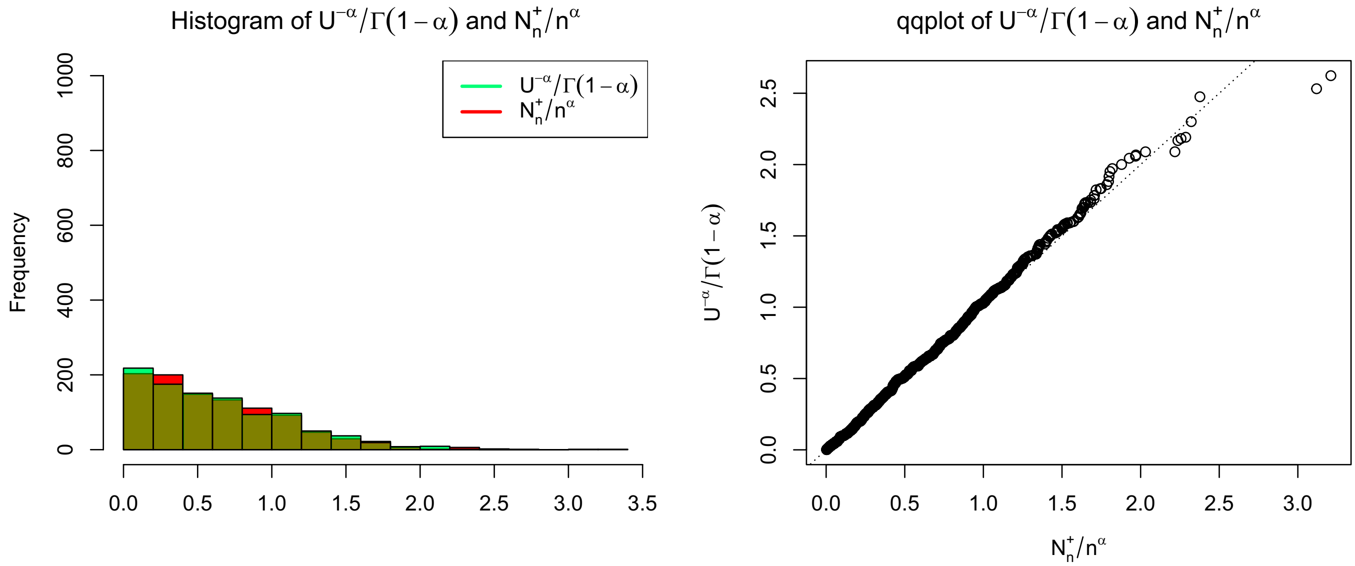

Fig. 2. Histogram and quantile–quantile plot of N_n^+/n^α and $U^{-\alpha}/\Gamma(1-\alpha)$ for $\alpha = 0.50$. The *P* value of the KS test is 0.9135.

quantile–quantile plots of N_n^+/n^α and $U^{-\alpha}/\Gamma(1-\alpha)$ with $\alpha = 0.25$ and $\alpha = 0.5$ are in Figs. 1 and 2, respectively. The histograms mostly overlap. The *P* values of the two-sample Kolmogorov–Smirnov (KS) test are 0.1995 and 0.9135, respectively. These observations support the convergence of N_n^+/n^α in distribution.

As expected, the speed of convergence of N_n^+/n^α in *Theorem 9* depends on α. Similarly, the speeds of convergence of the moment ratios in *Theorems 3* and *6* also depend on both α and the orders of the moments. We discuss the sample sizes required to see the convergence in distributions in *Theorems 3, 6,* and *9* in *SI Appendix*. From our simulation results, smaller α and higher-order moments result in faster convergence in distribution for the ratios of the moments.

Data Availability. Computer code has been deposited in GitHub (https://github.com/cftang9/TLHM). Readers can generate the tables and figures using the R code there.

ACKNOWLEDGMENTS. J.E.C. thanks Roseanne Benjamin for help during this work. S.C.P.Y. acknowledges financial support from Hong Kong General Research Fund Grants HKGRF-14300717 "New Kinds of Forward-Backward Stochastic Systems with Applications," HKGRF-14300319 "Shape-Constrained Inference: Testing for Monotonicity," and HKGRF-14301321 "General Theory for Infinite Dimensional Stochastic Control: Mean Field and Some Classical Problems."

1. R. Carmona, "Heavy tail distributions" in *Statistical Analysis of Financial Data in R* (Springer, New York, NY, ed. 2, 2014), chap. 2, 69–120.
2. W. Feller, *An Introduction to Probability Theory and Its Applications* (John Wiley & Sons, Inc., New York, NY, 1971), vol. 2.
3. S. I. Resnick, *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling* (Springer Science & Business Media, 2007).
4. G. Samorodnitsky, M. S. Taqqu, *Stable Non-Gaussian Random Processes* (Chapman & Hall, New York, NY, 1994).
5. J. Nešlehová, P. Embrechts, V. Chavez-Demoulin, Infinite mean models and the LDA for operational risk. *J. Oper. Risk* **1**, 3–25 (2006).
6. M. Campolieti, Heavy-tailed distributions and the distribution of wealth: Evidence from rich lists in Canada, 1999–2017. *Physica A* **503**, 263–272 (2018).
7. C. Schluter, Top incomes, heavy tails, and rank-size regressions. *Econometrics* **6**, 10 (2018).
8. F. M. Scherer, D. Harhoff, J. Kukies, "Uncertainty and the size distribution of rewards from innovation" in *Capitalism and Democracy in the 21st Century*, D. C. Mueller, U. Cantner, Eds. (Springer, 2001), pp. 181–206.
9. G. Silverberg, B. Verspagen, The size distribution of innovations revisited: An application of extreme value statistics to citation and value measures of patent significance. *J. Econom.* **139**, 318–339 (2007).
10. Y. Cen, "City size distribution, city growth and urbanisation in China," PhD thesis, University of Birmingham, Birmingham, United Kingdom (2015).
11. N. Bérubé, M. Sainte-Marie, P. Mongeon, V. Larivière, Words by the tail: Assessing lexical diversity in scholarly titles using frequency-rank distribution tail fits. *PLoS One* **13**, e0197775 (2018).
12. P. Embrechts, C. Klüppelberg, T. Mikosch, *Modelling Extremal Events: For Insurance and Finance* (Springer Science & Business Media, 2013), vol. 33.
13. R. Ibragimov, D. Jaffee, J. Walden, Nondiversification traps in catastrophe insurance markets. *Rev. Financ. Stud.* **22**, 959–993 (2009).
14. R. Ibragimov, Portfolio diversification and value at risk under thick-tailedness. *Quant. Finance* **9**, 565–580 (2009).
15. M. Brown, J. E. Cohen, V. H. de la Peña, Taylor's law, via ratios, for some distributions with infinite mean. *J. Appl. Probab.* **54**, 657–669 (2017).
16. L. R. Taylor, Aggregation, variance and the mean. *Nature* **189**, 732–735 (1961).
17. R. A. J. Taylor, *Taylor's Power Law: Order and Pattern in Nature* (Elsevier Academic Press, Cambridge, MA, 2019).
18. Z. Eisler, I. Bartos, J. Kertész, Fluctuation scaling in complex systems: Taylor's law and beyond. *Adv. Phys.* **57**, 89–142 (2008).
19. P. Berck, J. M. Hihn, Using the semivariance to estimate safety-first rules. *Am. J. Agric. Econ.* **64**, 298–300 (1982).
20. S. A. Bond, S. E. Satchell, Statistical properties of the sample semivariance. *Appl. Math. Finance* **9**, 219–239 (2002).
21. W. W. Hogan, J. M. Warren, Toward the development of an equilibrium capital-market model based on semivariance. *J. Financ. Quant. Anal.* **9**, 1–11 (1974).
22. H. Jin, H. Markowitz, X.Y. Zhou, A note on semivariance. *Math. Finance*. **16**, 53–61 (2006).
23. K. Liagkouras, K. Metaxiotis, The constrained mean-semivariance portfolio optimization problem with the support of a novel multiobjective evolutionary algorithm. *J. Softw. Eng. Appl.* **6**, 22–29 (2013).
24. T. J. Nantell, B. Price, An analytical comparison of variance and semivariance capital market theories. *J. Financ. Quant. Anal.* **14**, 221–242 (1979).
25. R. B. Porter, Semivariance and stochastic dominance: A comparison. *Am. Econ. Rev.* **64**, 200–204 (1974).
26. C. G. Turvey, G. Nayak, The semivariance-minimizing hedge ratio. *J. Agric. Resour. Econ.* **28**, 100–115 (2003).
27. C. Z. van de Beek, H. Leijnse, P. J. J. F. Torfs, R. Uijlenhoet, Climatology of daily rainfall semi-variance in The Netherlands. *Hydrol. Earth Syst. Sci.* **15**, 171–183 (2011).
28. A. Giometto, M. Formentin, A. Rinaldo, J. E. Cohen, A. Maritan, Sample and population exponents of generalized Taylor's law. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 7755–7760 (2015).
29. J. E. Cohen, M. Xu, W. S. F. Schuster, Allometric scaling of population variance with mean body size is predicted from Taylor's law and density-mass allometry. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 15829–15834 (2012).
30. J. E. Cohen, M. Xu, W. S. F. Schuster, Stochastic multiplicative population growth predicts and interprets Taylor's power law of fluctuation scaling. *Proc. Biol. Sci.* **280**, 20122955 (2013).
31. W. F. Sharpe, Mutual fund performance. *J. Bus.* **39**, 119–138 (1966).
32. R. E. Miller, A. K. Gehr, Sample size bias and Sharpe's performance measure: A note. *J. Financ. Quant. Anal.* **13**, 943–946 (1978).
33. M. Eling, S. Farinelli, D. Rossello, L. Tibiletti, One-size or tailor-made performance ratios for ranking hedge funds? *J. Deriv. Hedge Funds* **16**, 267–277 (2011).
34. F. A. Sortino, L. N. Price, Performance measurement in a downside risk framework. *J. Invest.* **3**, 59–64 (1994).
35. T. N. Rollinger, S. T. Hoffman, *Sortino: A 'Sharper' Ratio* (Red Rock Capital, Chicago, IL, 2013).

36. S. Farinelli, L. Tibiletti, Sharpe thinking in asset ranking with one-sided measures. *Eur. J. Oper. Res.* **185**, 1542–1547 (2008).
37. H. Albrecher, S. A. Ladoucette, J. L. Teugels, Asymptotics of the sample coefficient of variation and the sample dispersion. *J. Stat. Plan. Inference* **140**, 358–368 (2010).
38. J. E. Cohen, R. A. Davis, G. Samorodnitsky, Heavy-tailed distributions, correlations, kurtosis and Taylor's Law of fluctuation scaling. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **476**, 20200610 (2020).
39. A. Cascon, C. Keating, W. F. Shadwick, *An Introduction to Omega* (The Finance Development Centre, Fuqua-Duke University, 2002).
40. F. A. Sortino, R. Van Der Meer, A. Plantinga, The Dutch triangle. *J. Portfol. Manage.* **26**, 50–57 (1999).
41. J. H. J. Einmahl, The empirical distribution function as a tail estimator. *Stat. Neerl.* **44**, 79–82 (1990).
42. S. J. Wolfe, "On moments of probability distribution functions" in *Fractional Calculus and Its Applications*, B. Ross, Ed. (Springer-Verlag, Berlin, Germany, 1975), pp. 306–316.
43. A. W. Marshall, I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic Press, New York, NY, 1979).
44. B. M. Hill, A simple general approach to inference about the tail of a distribution. *Ann. Stat.* **3**, 1163–1174 (1975).
45. P. Hall, On some simple estimates of an exponent of regular variation. *J. R. Stat. Soc. B* **44**, 37–42 (1982).
46. M. I. Gomes, A. Guillou, Extreme value theory and statistics of univariate extremes: A review. *Int. Stat. Rev.* **83**, 263–292 (2015).
47. H. Bergström, On some expansions of stable distribution functions. *Ark. Mat.* **2**, 375–378 (1952).

APPLIED MATHEMATICS

ECONOMIC SCIENCES

Supplementary Materials: Taylor's law of fluctuation scaling for semivariances and higher moments of heavy-tailed data

Mark Brown[a], **Joel E. Cohen**[b,c,a,d], **Chuan-Fa Tang**[e], **and Sheung Chi Phillip Yam**[f]

[a]Department of Statistics, Columbia University; [b]Laboratory of Populations, The Rockefeller University; [c]Laboratory of Populations, Earth Institute, Columbia University; [d]Department of Statistics, University of Chicago; [e]Department of Mathematical Sciences, University of Texas at Dallas; [f]Department of Statistics, Chinese University of Hong Kong

This manuscript was compiled on October 25, 2021

Section A proves the assertions in Section 3 of the main text (semivariances). Section B does the same for Section 4 (higher moments). Section C does the same for Section 5 (number of observations that exceed the sample mean). Section D establishes the asymptotic properties of the modified financial ratios, such as quasi-concavity, scale-invariance, monotonicity, and sensitivity to the tail index of the distribution. Section E amplifies the results of Section 6 with more simulation results for the tail-index estimators. Section F examines the effects of sample size on the convergence of some distributions and statistics. Section G gives the references used in these Supplementary Materials.

The indicator function $I(A)$ of an event A is defined as

$$I(A) := \begin{cases} 1, & \text{if event } A \text{ occurs;} \\ 0, & \text{if event } A \text{ does not occur.} \end{cases}$$

Let $\xrightarrow{d}$ mean "converges in distribution to". Let $\xrightarrow{p}$ mean "converges in probability to". Let $\xrightarrow{\text{a.s.}}$ mean "converges almost surely to".

A. Proofs in Section 3: semivariances

Proof of Theorem 1. For $a > 0$, define

$$N_n(a) := \#\{X_i \le a | i \in \{1, \dots, n\}\}. \quad \text{[S.1]}$$

By definition,

$$I(M_1' > a) = \begin{cases} 1, & \text{if } M_1' > a; \\ 0, & \text{if } M_1' \le a. \end{cases} \quad \text{[S.2]}$$

If $X_i \le a$, then

$$(X_i - M_1')^2 = (M_1' - X_i)^2 \ge (M_1' - a)^2 I(M_1' > a),$$

hence

$$\frac{(X_i - M_1')^2}{M_1'^2} = \left(1 - \frac{X_i}{M_1'}\right)^2 \ge \left(1 - \frac{a}{M_1'}\right)^2 I(M_1' > a), \quad \text{[S.3]}$$

and therefore

$$\frac{v_n^-}{M_1'^2} = \frac{1}{n-1} \sum_{i: X_i \le M_1'} \frac{(X_i - M_1')^2}{M_1'^2} \ge \frac{N_n(a)}{n-1} \left(1 - \frac{a}{M_1'}\right)^2 I(M_1' > a). \quad \text{[S.4]}$$

The inequality in Eq. (S.4) is obtained by first omitting from the summation any term in which $a < X_i \le M_1'$, and then using the inequality Eq. (S.3) to replace each term in which $X_i \le a$ by its lower bound. Eq. (S.4) is a convenient lower bound.

By the strong law of large numbers for nonnegative random variables with mean $+\infty$, $M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$. Consequently, for fixed a, as $n \to \infty$,

$$I(M_1' > a) \xrightarrow{\text{a.s.}} 1, \quad \text{[S.5]}$$

$$\left(1 - \frac{a}{M_1'}\right)^2 \xrightarrow{\text{a.s.}} 1. \quad \text{[S.6]}$$

According to the strong law of large numbers, for fixed a, as $n \to \infty$,

$$\frac{N_n(a)}{n-1} \xrightarrow{\text{a.s.}} F_X(a) := \mathbb{P}(X \le a). \tag{S.7}$$

From Eq. (S.4), Eq. (S.5), Eq. (S.6), and Eq. (S.7), for any $a > 0$,

$$\liminf_{n\to\infty} \frac{v_n^-}{M_1'^2} \ge F_X(a) \quad \text{a.s.}. \tag{S.8}$$

Letting $a \to \infty$, Eq. (S.8) gives

$$\liminf_{n\to\infty} \frac{v_n^-}{M_1'^2} \ge 1 \quad \text{a.s.}. \tag{S.9}$$

According to a similar argument of Brown et al. (1, p. 665), for a given M_1', the maximal value of v_n^- is attained when any $n-1$ of X_is equal 0 and the one remaining X_i equals nM_1'. For such values, $v_n^- = (n-1)M_1'^2/(n-1) = M_1'^2$. Thus in all cases, $v_n^- \le M_1'^2$ and

$$\limsup_{n\to\infty} \frac{v_n^-}{M_1'^2} \le 1 \quad \text{a.s.}. \tag{S.10}$$

From Eq. (S.9) and Eq. (S.10), $v_n^-/M_1'^2 \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. ☐

Proof of Corollary 1. From Theorem 1, as $n \to \infty$,

$$\frac{\log v_n^-}{\log M_1'} = \frac{\log[v_n^-/(M_1')^2]}{\log M_1'} + \frac{\log(M_1')^2}{\log M_1'} \xrightarrow{\text{a.s.}} 0 + 2 = 2.$$

☐

Proof of Lemma 1. For any $a > 0$, by definition,

$$\frac{N_n^-}{n} \ge \frac{N_n(a)}{n} I(M_1' > a). \tag{S.11}$$

Hence, for all $a > 0$,

$$\liminf_{n\to\infty} \frac{N_n^-}{n} \ge F_X(a), \tag{S.12}$$

and, letting $a \to \infty$,

$$\liminf_{n\to\infty} \frac{N_n^-}{n} \ge 1. \tag{S.13}$$

But by definition $N_n^-/n \le 1$. Therefore, $N_n^-/n \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. ☐

Proof of Corollary 2. From Eq. (16) in Lemma 1, $N_n^-/n \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. By Corollary 1, as $n \to \infty$, we have

$$\frac{\log v_n^{-*}}{\log M_1'} = \frac{\log v_n^- + \log(n/N_n^-)}{\log M_1'} \xrightarrow{\text{a.s.}} 2 + 0 = 2.$$

☐

Proof of Corollary 3. We write

$$\frac{v_n^-}{M_1'^{\alpha^*}} = \frac{v_n^-}{M_1'^2} M_1'^{-(\alpha^*-2)}.$$

From Theorem 1, $v_n^-/M_1'^2 \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. Since $M_1' \xrightarrow{\text{a.s.}} \infty$ and $\alpha^* = (2-\alpha)/(1-\alpha) > 2$, it follows that $M_1'^{-(\alpha^*-2)} \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. Thus $v_n^-/M_1'^{\alpha^*} \xrightarrow{\text{a.s.}} 0$. From Lemma 1, $N_n^-/n \xrightarrow{\text{a.s.}} 1$, so $v_n^{-*}/M_1'^{\alpha^*} \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. ☐

By the standard definition, a sequence of random variables Y_n indexed by n is defined to be $O_p(1)$, and we write $Y_n = O_p(1)$, if for any $\epsilon > 0$, there exist M_ϵ, $0 < M_\epsilon < \infty$ and N_ϵ, $0 < N_\epsilon < \infty$ such that $\mathbb{P}(|Y_n| > M_\epsilon) < \epsilon$ for all $n > N_\epsilon$. If $Y_n \xrightarrow{d} Y$, then $Y_n = O_p(1)$, but the converse does not hold. We write $Y_n = o_p(1)$ if, for any $\epsilon > 0$, $\mathbb{P}(|Y_n| > \epsilon) \to 0$ as $n \to \infty$.

Lemma A.1. *Let $X_1,\dots,X_n$ be a random sample from $F_X \overset{d}{\approx} F(1,\alpha)$, $0<\alpha<1$, satisfying* Eq. (9). *Given $0<\epsilon<1$, define $q_\epsilon>0$ to be the quantile of $F(1,\alpha)$ such that $F(1,\alpha)(q_\epsilon)=\epsilon$. Define $b_n := q_\epsilon n^{(1-\alpha)/\alpha}$. Then as $n\to\infty$,*

$$\sup_{t\geq b_n}\left|\frac{\sum_{i=1}^n I(X_i>t)}{n^\alpha} - n^{1-\alpha}\{1-F_X(t)\}\right| \xrightarrow{p} 0.$$

Proof of Lemma A.1. Because X is non-atomic and nonnegative, $q_\epsilon>0$ so $b_n\to\infty$ since $0<\alpha<1$. From Einmahl (2, p. 80, Corollary 1), for any positive sequences $k_n\leq n$ and m_n such that $k_n\to\infty, k_n/n\to 0$ and $m_n/k_n^{1/2}\to\infty$ as $n\to\infty$,

$$\sup_{t\geq F_X^{-1}(1-k_n/n)}\frac{n}{m_n}\left|\frac{\sum_{i=1}^n I(X_i>t)}{n} - \{1-F_X(t)\}\right| \xrightarrow{p} 0,$$

where F_X^{-1} is the quantile function of F_X. We choose $k_n = n\{1-F_X(b_n)\} = n\{1-F_X(q_\epsilon n^{(1-\alpha)/\alpha})\}$ and $m_n=n^\alpha$. Then

$$\frac{m_n}{k_n^{1/2}} = \frac{n^\alpha}{[n\{1-F_X(q_\epsilon n^{(1-\alpha)/\alpha})\}]^{1/2}} = \frac{n^\alpha}{n^{\alpha/2}}\left(\frac{1}{n^{1-\alpha}\{1-F_X(q_\epsilon n^{(1-\alpha)/\alpha})\}}\right)^{1/2}.$$

From Eq. (9), by Definition 1, $n^{1-\alpha}\{1-F_X(q_\epsilon n^{(1-\alpha)/\alpha})\}\to\{q_\epsilon^\alpha\Gamma(1-\alpha)\}^{-1}>0$. Further, $n^\alpha/n^{\alpha/2}\to\infty$ as $n\to\infty$ since $0<\alpha<1$. Therefore $m_n/k_n^{1/2}\to\infty$ as $n\to\infty$. Applying Einmahl (2, p. 80, Corollary 1) gives the claimed limit. □

We clarify that the quantile q_ϵ is specific to the particular stable law $F(\alpha,1)$ with the Laplace transform in Eq. [7], so $F(\alpha,1)(q_\epsilon)=\epsilon$ for $0<\epsilon<1$. By contrast, F_X^{-1} is the quantile function of the distribution F_X of any random variable X that satisfies Eq. [9].

Lemma A.2. *Let $X_1,\dots,X_n$ and $X_1^*,\dots,X_n^*$ be two independent random samples from $F_X \overset{d}{\approx} F(1,\alpha)$ with $\alpha\in(0,1)$. Let these samples have sample means M_1' and $M_1'^*$, respectively. Then*

$$\frac{\sum_{i=1}^n I(X_i>M_1'^*)}{n^\alpha} = O_p(1), \qquad \left(\frac{\sum_{i=1}^n I(X_i>M_1'^*)}{n^\alpha}\right)^{-1} = O_p(1). \qquad \text{[S.14]}$$

We clarify that $M_1'^*$ is the sample mean of the <u>second</u> sample $\{X_i^*\}_{i=1}^n$ and $I(X_i>M_1'^*)=1$ if and only if the element X_i of the <u>first</u> sample $\{X_i\}_{i=1}^n$ exceeds the sample mean $M_1'^*$ of the <u>second</u> sample $\{X_i^*\}_{i=1}^n$.

Proof. Given $0<\epsilon<1$, define $q_\epsilon>0$ to be the quantile of $F(1,\alpha)$ such that $F(1,\alpha)(q_\epsilon)=\epsilon$. Define $b_n := q_\epsilon n^{(1-\alpha)/\alpha}$. Then $b_n\to\infty$ as $n\to\infty$ because $0<\alpha<1$. To show $n^{-\alpha}\sum_{i=1}^n I(X_i>M_1'^*)=O_p(1)$, we let $C>0$ be any positive constant and let $F_{M_1'^*}$ be the distribution of $M_1'^*$. Since $X_1,\dots,X_n$ and $M_1'^*$ are independent,

$$\begin{aligned}&\mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i>M_1'^*)}{n^\alpha}>C\right)\\ &=\int_0^{b_n}\mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i>t)}{n^\alpha}>C\right)dF_{M_1'^*}(t)+\int_{b_n}^\infty\mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i>t)}{n^\alpha}>C\right)dF_{M_1'^*}(t).\end{aligned} \qquad \text{[S.15]}$$

The first term on the right side of Eq. (S.15) is bounded above by $F_{M_1'^*}(b_n)$, where

$$F_{M_1'^*}(b_n)=\mathbb{P}\left(\frac{\sum_{i=1}X_i^*}{n}\leq q_\epsilon n^{(1-\alpha)/\alpha}\right)=\mathbb{P}\left(\frac{\sum_{i=1}X_i^*}{n^{1/\alpha}}\leq q_\epsilon\right)\to F(1,\alpha)(q_\epsilon)=\epsilon,$$

because $n^{-1/\alpha}\sum_{i=1}^n X_i^* \xrightarrow{d} F(1,\alpha)$ as $n\to\infty$ from Albrecher et al. (3, p. 362, Remark 2.1). For an upper bound for the second term on the right side of Eq. (S.15), we observe that

$$\int_{b_n}^\infty\mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i>t)}{n^\alpha}>C\right)dF_{M_1'^*}(t)\leq\mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i>b_n)}{n^\alpha}>C\right).$$

Therefore it suffices to show that $n^{-\alpha}\sum_{i=1}^n I(X_i>b_n)=O_p(1)$. From Lemma A.1, as $n\to\infty$,

$$\left|\frac{\sum_{i=1}^n I(X_i>b_n)}{n^\alpha}-n^{1-\alpha}\{1-F_X(b_n)\}\right|\xrightarrow{p}0.$$

Because $b_n=q_\epsilon n^{(1-\alpha)/\alpha}$, we have $n^{1-\alpha}\{1-F_X(b_n)\}=(b_n^\alpha/q_\epsilon^\alpha)\{1-F_X(b_n)\}\to\{q_\epsilon^\alpha\Gamma(1-\alpha)\}^{-1}$ as $n\to\infty$. Therefore

$$\frac{\sum_{i=1}^n I(X_i>b_n)}{n^\alpha}\xrightarrow{p}\{q_\epsilon^\alpha\Gamma(1-\alpha)\}^{-1}$$

as $n \to \infty$ and $n^{-\alpha}\sum_{i=1}^n I(X_i > b_n) = O_p(1)$.

A similar calculation replacing q_ϵ by $q_{1-\epsilon}$ proves the second claim in Lemma A.2 as follows. For $0 < \epsilon < 1$, recall that q_ϵ is the quantile of X such that $F(1,\alpha)(q_\epsilon) = \epsilon$. Here we define $b_n := q_{1-\epsilon} n^{(1-\alpha)/\alpha}$. Then $b_n \to \infty$ and, for a given $C > 0$, in this case

$$\mathbb{P}\left(\left(\frac{\sum_{i=1}^n I(X_i > M_1'^*)}{n^\alpha}\right)^{-1} > C\right)$$
$$= \int_0^{b_n} \mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i > t)}{n^\alpha} < \frac{1}{C}\right) dF_{M_1'^*}(t) + \int_{b_n}^{\infty} \mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i > t)}{n^\alpha} < \frac{1}{C}\right) dF_{M_1'^*}(t). \quad \text{[S.16]}$$

For the second term on the right side of Eq. (S.16),

$$\int_{b_n}^{\infty} \mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i > t)}{n^\alpha} < \frac{1}{C}\right) dF_{M_1'^*}(t) \le 1 - F_{M_1'^*}(b_n).$$

Because $n^{-1/\alpha}\sum_{i=1}^n X_i^* \xrightarrow{d} F(1,\alpha)$, we have, as $n \to \infty$,

$$1 - F_{M_1'^*}(b_n) = \mathbb{P}\left(\frac{\sum_{i=1} X_i^*}{n} > q_{1-\epsilon} n^{(1-\alpha)/\alpha}\right) = \mathbb{P}\left(\frac{\sum_{i=1} X_i^*}{n^{1/\alpha}} > q_{1-\epsilon}\right) \to 1 - F(1,\alpha)(q_{1-\epsilon}) = \epsilon.$$

For the first term on the right side of Eq. (S.16), when $0 \le t \le b_n$,

$$\mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i > t)}{n^\alpha} < \frac{1}{C}\right) \le \mathbb{P}\left(\frac{\sum_{i=1}^n I(X_i > b_n)}{n^\alpha} < \frac{1}{C}\right).$$

From Lemma A.1, as $n \to \infty$,

$$\frac{\sum_{i=1}^n I(X_i > b_n)}{n^\alpha} \xrightarrow{p} \{q_{1-\epsilon}^\alpha \Gamma(1-\alpha)\}^{-1}.$$

Therefore, for large enough $C > 0$ such that $C^{-1} < \{q_{1-\epsilon}^\alpha \Gamma(1-\alpha)\}^{-1}$, we have $\mathbb{P}\left(n^{-\alpha}\sum_{i=1}^n I(X_i > b_n) < C^{-1}\right) \to 0$ as $n \to \infty$. $\square$

Lemma A.3. *Let $X_1, \ldots, X_n$ be a random sample of size n from the distribution satisfying $X \overset{d}{\approx} F(c,\alpha)$, $c > 0$, $0 < \alpha < 1$, with sample mean M_1'. Then, with N_n^+ defined in* Eq. (5),

$$\frac{N_n^+}{n^\alpha} := \frac{\sum_{i=1}^n I(X_i > M_1')}{n^\alpha} = O_p(1), \qquad \left(\frac{N_n^+}{n^\alpha}\right)^{-1} = O_p(1).$$

Proof of Lemma A.3. Consider $k > 1$ independent random samples, each of size n, from F_X:

$$\{X_1^{(1)}, \cdots, X_n^{(1)}\}, \{X_1^{(2)}, \cdots, X_n^{(2)}\}, \ldots, \{X_1^{(k)}, \cdots, X_n^{(k)}\},$$

having sample means $M_1'^{(1)}, \ldots, M_1'^{(k)}$, respectively. Let

$$\widetilde{\mathcal{I}} := I\left(M_1' > \min_{1\le j\le k} M_1'^{(j)}\right).$$

Recall that $N_n^+ := \#\{i : X_i > M_1'\}$ (Eq. (5)) and define $\overline{N}_n(a) := \sum_{i=1}^n I(X_i > a)$ for $a \in \mathbb{R}$ with

$$\overline{N}_n\left(M_1'^{(j)}\right) = \sum_{i=1}^n I(X_i > M_1'^{(j)}), \quad j = 1, \ldots, k;$$
$$\overline{N}_n\left(\min_{1\le j\le k} M_1'^{(j)}\right) = \sum_{i=1}^n I\left(X_i > \min_{1\le j\le k} M_1'^{(j)}\right).$$

Then $N_n^+ = N_n^+\widetilde{\mathcal{I}} + N_n^+(1-\widetilde{\mathcal{I}})$ with

$$N_n^+(1-\widetilde{\mathcal{I}}) = N_n^+ I\{M_1' \le \min_{1\le j\le k} M_1'^{(j)}\}. \quad \text{[S.17]}$$

Next,

$$\begin{aligned}
N_n^+\widetilde{\mathcal{I}} &= \#\{i : X_i > M_1'\} I\{M_1' > \min_{1\le j\le k} M_1'^{(j)}\} \\
&\le \#\{i : X_i > \min_{1\le j\le k} M_1'^{(j)}\} I\{M_1' > \min_{1\le j\le k} M_1'^{(j)}\} \\
&= \left\{\max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\}\right\} I\{M_1' > \min_{1\le j\le k} M_1'^{(j)}\} \qquad \text{(see below)} \\
&= \left\{\max_{1\le j\le k} \overline{N}_n\left(M_1'^{(j)}\right)\right\} I\{M_1' > \min_{1\le j\le k} M_1'^{(j)}\} \\
&\le \max_{1\le j\le k} \overline{N}_n\left(M_1'^{(j)}\right). \qquad \text{[S.18]}
\end{aligned}$$

The equality $\#\{i : X_i > \min_{1\le j\le k} M_1'^{(j)}\} = \left\{\max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\}\right\}$ holds because $\min_{1\le j\le k} M_1'^{(j)}$ is the smallest member among $\{M_1'^{(1)}, \dots, M_1'^{(k)}\}$, so for all $1 \le j \le k$,

$$\#\{i : X_i > \min_{1\le j\le k} M_1'^{(j)}\} \ge \#\{i : X_i > M_1'^{(j)}\}.$$

Therefore, the inequality above still holds for the maximum on the right-hand side, that is

$$\#\{i : X_i > \min_{1\le j\le k} M_1'^{(j)}\} \ge \max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\}.$$

On the other hand, because $\max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\}$ is the largest number among $\{\#\{i : X_i > M_1'^{(1)}\}, \dots, \#\{i : X_i > M_1'^{(k)}\}\}$, then

$$\max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\} \ge \#\{i : X_i > M_1'^{(j)}\},$$

for all $1 \le j \le k$. Therefore,

$$\max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\} \ge \#\{i : X_i > \min_{1\le j\le k} M_1'^{(j)}\}$$

because $\min_{1\le j\le k} M_1'^{(j)}$ is still a member of $\{M_1'^{(1)}, \dots, M_1'^{(k)}\}$. Since we have proved the weak inequality in both directions, we have the equality $\max_{1\le j\le k} \#\{i : X_i > M_1'^{(j)}\} = \#\{i : X_i > \min_{1\le j\le k} M_1'^{(j)}\}$.

From Eq. (S.17) and Eq. (S.18),

$$N_n^+ \le \max\left(\max_{1\le j\le k} \overline{N}_n(M_1'^{(j)}), N_n^+ I\{M_1' \le \min_{1\le j\le k} M_1'^{(j)}\}\right). \qquad \text{[S.19]}$$

For $M > 0$, $N_n^+ I(M_1' \le \min_{1\le j\le k} M_1'^{(j)}) > n^\alpha M$ implies that $I(M_1' \le \min_{1\le j\le k} M_1'^{(j)}) = 1$. Therefore

$$\mathbb{P}\left(\frac{N_n^+ I(M_1' \le \min_{1\le j\le k} M_1'^{(j)})}{n^\alpha} > M\right) \le \mathbb{P}\left(I\{M_1' \le \min_{1\le j\le k} M_1'^{(j)}\} = 1\right).$$

Because $M_1', M_1'^{(1)}, \dots, M_1'^{(k)}$ are independent and identically distributed, $I(M_1' \le \min_{1\le j\le k} M_1'^{(j)}) = 1$ if and only if M_1' is the smallest number among $M_1', M_1'^{(1)}, \dots, M_1'^{(k)}$, which has probability $1/(k+1)$. Therefore,

$$\mathbb{P}\left(\frac{N_n^+ I(M_1' \le \min_{1\le j\le k} M_1'^{(j)})}{n^\alpha} > M\right) \le \frac{1}{k+1}. \qquad \text{[S.20]}$$

Furthermore, because the sample $\{X_1, \dots, X_n\}$ is independent of $M_1'^{(1)}, \dots, M_1'^{(k)}$, it is also true that $\overline{N}_n(M_1'^{(1)}), \dots, \overline{N}_n(M_1'^{(k)})$ are identically distributed and for $1 \le j \le k$,

$$\mathbb{P}\left(\frac{\overline{N}_n(M_1'^{(1)})}{n^\alpha} > M\right) = \mathbb{P}\left(\frac{\overline{N}_n(M_1'^{(j)})}{n^\alpha} > M\right). \qquad \text{[S.21]}$$

Since $\mathbb{P}\left(\max\{Y_l, l = 1, \dots, m\} > t\right) = \mathbb{P}\left(\cup_{1\le l\le m}\{Y_l > t\}\right) \le \sum_{l=1}^{n} \mathbb{P}\left(\{Y_l > t\}\right)$, from Eq. (S.19), Eq. (S.20), and Eq. (S.21), we have

$$\begin{aligned}
\mathbb{P}\left(\frac{N_n^+}{n^\alpha} > M\right) &\le \mathbb{P}\left(n^{-\alpha} N_n^+ I(M_1' \le \min_{1\le j\le k} M_1'^{(j)}) > M\right) + \left\{\sum_{1\le j\le k} \mathbb{P}(n^{-\alpha}\overline{N}_n(M_1'^{(j)}) > M)\right\} \\
&\le \mathbb{P}\left(I\{M_1' \le \min_{1\le j\le k} M_1'^{(j)}\} = 1\right) + \left\{\sum_{1\le j\le k} \mathbb{P}(n^{-\alpha}\overline{N}_n(M_1'^{(j)}) > M)\right\} \\
&\le \frac{1}{k+1} + k\mathbb{P}\left(n^{-\alpha}\overline{N}_n(M_1'^{(1)}) > M\right).
\end{aligned}$$

Given $\epsilon > 0$, we can choose a large enough k to make $1/(k+1)$ small enough so that $1/(k+1) < \epsilon/2$. For the chosen k, because $N_n(M_1'^{(1)}) = O_p(1)$ from Lemma A.2, we can further choose M large enough that $\mathbb{P}(n^{-\alpha}\overline{N}_n(M_1'^{(1)}) > M) < \epsilon/2k$ for a large enough n, and then $\mathbb{P}(n^{-\alpha}N_n^+ > M) < \epsilon$. ☐

Proof of Theorem 2. Since $v_n = v_n^- + v_n^+$, it follows from Eq. (18) for v_n^- and from Eq. (11) for v_n that $v_n^+/(M_1')^{\alpha^*} \xrightarrow{d} W$ as $n \to \infty$. Hence as $n \to \infty$,

$$\log v_n^+ - \alpha^* \log M_1' \xrightarrow{d} \log W. \qquad \text{[S.22]}$$

Dividing both sides of Eq. (S.22) by $\log M_1'$, and employing a version of Slutsky's theorem (Arnold (4, p. 242, Corollary 6.8(c))), gives, as $n \to \infty$,

$$\frac{\log v_n^+}{\log M_1'} - \alpha^* \xrightarrow{p} 0,$$

which is the first part of Eq. (19). According to Lemma A.3, $N_n^+/n^\alpha = O_p(1)$ and $n^\alpha/N_n^+ = O_p(1)$. Then

$$O_p(1) = \log(N_n^+/n^\alpha) = \log N_n^+ - \alpha \log n,$$
$$O_p(1) = \log(n^\alpha/N_n^+) = -\log N_n^+ + \alpha \log n,$$

so as $n \to \infty$,

$$\frac{\log N_n^+}{\log n} \xrightarrow{p} \alpha, \qquad \frac{\log \frac{N_n^+}{n}}{\log n} \xrightarrow{p} \alpha - 1, \qquad \frac{\log \frac{n}{N_n^+}}{\log n} \xrightarrow{p} 1 - \alpha. \qquad \text{[S.23]}$$

Recall that $n^{-(\frac{1-\alpha}{\alpha})}M_1' \xrightarrow{d} F(1,\alpha)$ as $n \to \infty$ [Feller (5, p. 448)]. Then $\log M_1' - (\frac{1-\alpha}{\alpha})\log n = O_p(1)$ and

$$\frac{\log M_1'}{\log n} \xrightarrow{p} \frac{1-\alpha}{\alpha}. \qquad \text{[S.24]}$$

From Eq. (S.23) and Eq. (S.24), as $n \to \infty$,

$$\frac{\log(n/N_n^+)}{\log M_1'} = \frac{\log(n/N_n^+)/\log n}{\log M_1'/\log n} \xrightarrow{p} \frac{1-\alpha}{(1-\alpha)/\alpha} = \alpha. \qquad \text{[S.25]}$$

From the definition Eq. (6) of the local upper semivariance,

$$v_n^{+*} = v_n^+ \left(\frac{n-1}{N_n^+}\right). \qquad \text{[S.26]}$$

Eq. (S.26), Eq. (S.25), and Eq. (19) give, as $n \to \infty$,

$$\frac{\log v_n^{+*}}{\log M_1'} = \frac{\log(v_n^+)}{\log M_1'} + \frac{\log((n-1)/N_n^+)}{\log M_1'} \xrightarrow{p} \frac{2-\alpha}{1-\alpha} + \alpha = \frac{2-\alpha^2}{1-\alpha},$$

which is the second part of Eq. (19). ☐

B. Proofs in Section 4: higher moments

We assume $0 < \alpha < 1$ throughout. To prove Theorem 3, we recall a standard definition and prove a lemma.

Definition B.1. *For two sequences b_n, c_n such that $b_n \to \infty$ and $c_n \to \infty$ as $n \to \infty$, define $b_n \sim c_n$ to mean that $\lim_{n\to\infty} b_n/c_n = 1$.*

Define a_n as a sequence of nonnegative numbers such that $1 - F_X(a_n) \sim n^{-1}$ where $F_X \overset{d}{\approx} F(1,\alpha)$ as in Definition 1, Eq. (9). Thus $a_n \sim \{n/\Gamma(1-\alpha)\}^{1/\alpha}$.

Lemma B.1. *Given $s > 0, t > 0, y > 0$, and $h_1, h_2 > 0$, the equation $sx^{h_2} + tx^{h_1} - y = 0$ has exactly one positive root $x = r_{h_2,h_1}(y,s,t)$ and the equation $sx^{h_2}/a_n^{h_2} + tx^{h_1}/a_n^{h_1} - y = 0$ has exactly one positive root $x = a_n r_{h_2,h_1}(y,s,t)$.*

Proof. The function $sx^{h_2}+tx^{h_1}-y$ is strictly increasing in $x > 0$ because it has a positive first derivative $s\cdot h_2 x^{h_2-1}+t\cdot h_1 x^{h_1-1} > 0$ for $x > 0$. When $x = 0$, then $sx^{h_2} + tx^{h_1} - y = -y < 0$. When $x = (y/t)^{1/h_1} > 0$, then $sx^{h_2} + tx^{h_1} - y > tx^{h_1} - y = 0$. The unique positive root must be in the interval $(0, (y/t)^{1/h_1})$. If $x = a_n r_{h_2,h_1}(y,s,t)$, then

$$\begin{aligned} sx^{h_2}/a_n^{h_2} + tx^{h_1}/a_n^{h_1} - y &= s\{a_n r_{h_2,h_1}(y,s,t)\}^{h_2}/a_n^{h_2} + t\{a_n r_{h_2,h_1}(y,s,t)\}^{h_1}/a_n^{h_1} - y \\ &= s\{r_{h_2,h_1}(y,s,t)\}^{h_2} + t\{r_{h_2,h_1}(y,s,t)\}^{h_1} - y = 0. \end{aligned}$$

☐

Proof of Theorem 3. The first part of the proof of the convergence in distribution briefly follows the proof of Albrecher et al. (3, p. 361, Theorem 2.1) and Brown et al. (1). For $X_1 \overset{d}{=} F_X$ and $h > 0$, we calculate the following integral by integration by parts:

$$1 - \mathbb{E}[e^{-\theta X_1^h - \psi X_1}] = \int_0^\infty (1 - e^{-\theta x^h - \psi x})\, dF_X(x) = \int_0^\infty \{1 - F_X(x)\}(h\theta x^{h-1} + \psi) e^{-\theta x^h - \psi x}\, dx.$$

Set $y = \theta x^h + \psi x$. Then

$$1 - \mathbb{E}[e^{-\theta X_1^h - \psi X_1}] = \int_0^\infty [1 - F_X\{r_h(y, \theta, \psi)\}] e^{-y}\, dy, \tag{S.27}$$

where $x = r_h(y, \theta, \psi)$ is the only positive root of $\theta x^h + \phi x - y = 0$. Setting $\theta = s/a_n^h$ and $\psi = t/a_n$, we write

$$\mathbb{E}(e^{-s(1/a_n^h)\sum_{i=1}^n X_i^h - t(1/a_n)\sum_{i=1}^n X_i}) = \exp\{n \log \mathbb{E}(e^{-s(1/a_n^h)X_1^h - t(1/a_n)X_1})\}.$$

Because $a_n \to \infty$ as $n \to \infty$, we have $\mathbb{E}(e^{-s(1/a_n^h)X_1^h - t(1/a_n)X_1}) \to 1$ as $n \to \infty$. Then by Taylor's expansion of the log function,

$$\begin{aligned}\exp\{n \log \mathbb{E}(e^{-s(1/a_n^h)X_1^h - t(1/a_n)X_1})\} &\sim \exp\{-n(1 - \mathbb{E}[e^{-s(1/a_n^h)X_1^h - t(1/a_n)X_1}])\} \\ &= \exp\left\{-n \int_0^\infty \left[1 - F_X\left\{a_n r_h(y, s/a_n^h, t/a_n)\right\}\right] e^{-y}\, dy\right\}.\end{aligned}$$

The last equality holds because of Eq. (S.27). Lemma B.1 shows that $r_h(y, s/a_n^h, t/a_n)/a_n = r_h(y, s, t)$, where $x = r_h(y, s, t)$ is the unique positive root of $sx^h + tx - y = 0$. Then for every positive integer n,

$$1 - F_X\left\{r_h(y, s/a_n^h, t/a_n)\right\} = 1 - F_X\left\{a_n \frac{r_h(y, s/a_n^h, t/a_n)}{a_n}\right\} = 1 - F_X\left\{a_n r_h(y, s, t)\right\}.$$

From Eq. (9) and $a_n \sim \{n/\Gamma(1-\alpha)\}^{1/\alpha}$, we have

$$1 - F_X\left\{a_n r_h(y, s, t)\right\} \sim \frac{\{a_n r_h(y, s, t)\}^{-\alpha}}{\Gamma(1-\alpha)} \sim \frac{\frac{\Gamma(1-\alpha)}{n}\{r_h(y, s, t)\}^{-\alpha}}{\Gamma(1-\alpha)} = \frac{1}{n}\{r_h(y, s, t)\}^{-\alpha}.$$

On the other hand, $sx^h + tx - y < 0$ when $x = 0$ and $sx^h + tx - y > tx - y = 0$ when $x = y/t$. Then $0 < r_h(y, s, t) < y/t$ because $sx^h + tx - y$ is strictly increasing in $x > 0$. Because $\int_0^\infty \{y/t\}^{-\alpha} e^{-y}\, dy = t^\alpha \Gamma(1-\alpha) < \infty$, by the dominated convergence theorem, the limit of the joint Laplace transform of $(\frac{n}{a_n^h} M_h', \frac{n}{a_n} M_1')$ exists and is given by

$$\lim_{n\to\infty} \mathbb{E}\left(e^{-(s/a_n^h)\sum_{i=1}^n X_i^h - (t/a_n)\sum_{i=1}^n X_i}\right) = \exp\left\{-\int_0^\infty \{r_h(y, s, t)\}^{-\alpha} e^{-y}\, dy\right\}. \tag{S.28}$$

We conclude that $\left(\frac{n}{a_n^h} M_h', \frac{n}{a_n} M_1'\right) \xrightarrow{d} (U_h, V)$ as $n \to \infty$ where (U_h, V) has the joint Laplace transform Eq. (S.28). Therefore by Slutsky's theorem, as $n \to \infty$, for $h > \alpha$,

$$\frac{M_h'}{(M_1')^{\alpha(h,1)}} = \frac{a_n^h/n}{(a_n/n)^{\alpha(h,1)}} \frac{\frac{n}{a_n^h} M_h'}{(\frac{n}{a_n} M_1')^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \cdot \frac{U_h}{V^{\alpha(h,1)}}. \qquad \square$$

Lemma B.2. *Under the assumptions of Theorem 3,*

$$\frac{M_h'}{(M_1')^{\alpha(h,1)}} = O_p(1), \qquad \frac{(M_1')^{\alpha(h,1)}}{M_h'} = O_p(1).$$

Proof. From Theorem 3, as $n \to \infty$,

$$\frac{M_h'}{(M_1')^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}}, \qquad \frac{(M_1')^{\alpha(h,1)}}{M_h'} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{1-h}{1-\alpha}} \frac{V^{\alpha(h,1)}}{U_h}.$$

Then it suffices to show that, for any $\epsilon > 0$,

$$\lim_{\epsilon\to 0} \mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} < \epsilon\right) = 0, \qquad \lim_{\epsilon\to 0} \mathbb{P}\left(\frac{V^{\alpha(h,1)}}{U_h} < \epsilon\right) = 0.$$

Given $\epsilon > 0$ and $c > 0$,

$$\begin{aligned}\mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} \le \epsilon\right) &= \mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} \le \epsilon, V^{\alpha(h,1)} \le c\right) + \mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} \le \epsilon, V^{\alpha(h,1)} > c\right)\\ &= \mathbb{P}\left(U_h \le V^{\alpha(h,1)}\epsilon, V^{\alpha(h,1)} \le c\right) + \mathbb{P}\left(U_h \le V^{\alpha(h,1)}\epsilon, V^{\alpha(h,1)} > c\right)\\ &\le \mathbb{P}\left(U_h \le c\epsilon, V^{\alpha(h,1)} \le c\right) + \mathbb{P}\left(U_h \le V^{\alpha(h,1)}\epsilon, V^{\alpha(h,1)} > c\right)\\ &\le \mathbb{P}\left(U_h \le c\epsilon\right) + \mathbb{P}\left(V^{\alpha(h,1)} > c\right) = \mathbb{P}\left(U_h \le c\epsilon\right) + \mathbb{P}\left(V > c^{1/\alpha(h,1)}\right).\end{aligned}$$

Choose $c = 1/\epsilon^{1/2}$. Then

$$\begin{aligned}\mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} \le \epsilon\right) &\le \mathbb{P}\left(U_h \le \epsilon^{1/2}\right) + \mathbb{P}\left(V > \epsilon^{-1/2\alpha(h,1)}\right)\\ &= \mathbb{P}\left(\frac{U_h}{\{\Gamma(1-\alpha/h)\}^{h/\alpha}} \le \frac{\epsilon^{1/2}}{\{\Gamma(1-\alpha/h)\}^{h/\alpha}}\right) + \mathbb{P}\left(\frac{V}{\{\Gamma(1-\alpha)\}^{1/a}} > \frac{\epsilon^{-1/2\alpha(h,1)}}{\{\Gamma(1-\alpha)\}^{1/\alpha}}\right).\end{aligned}$$

Recall Remark 1 that $U_h \overset{d}{=} F(\{\Gamma(1-\alpha/h)\}^{h/\alpha}, \alpha/h)$ and $V \overset{d}{=} F(\{\Gamma(1-\alpha)\}^{1/\alpha}, \alpha)$. Therefore, $U_h/\{\Gamma(1-\alpha/h)\}^{h/\alpha} \overset{d}{=} F(1, \alpha/h)$, $V/\{\Gamma(1-\alpha)\}^{1/\alpha} \overset{d}{=} F(1, \alpha)$, and

$$\mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} \le \epsilon\right) \le F(1, \alpha/h)\left(\frac{\epsilon^{1/2}}{\{\Gamma(1-\alpha/h)\}^{h/\alpha}}\right) + \left\{1 - F(1,\alpha)\left(\frac{\epsilon^{-1/2\alpha(h,1)}}{\{\Gamma(1-\alpha)\}^{1/\alpha}}\right)\right\}.$$

Feller (5, p. 448, XIII(6.1), (6.2)) states that $e^{x^{-\alpha/h}} F(1, \alpha/h)(x) \to 0$ as $x \to 0$ and $x^\alpha\{1 - F(1,\alpha)(x)\} \to \frac{1}{\Gamma(1-\alpha)}$ as $x \to \infty$. Therefore $F(1, \alpha/h)\left(\frac{\epsilon^{1/2}}{\{\Gamma(1-\alpha/h)\}^{h/\alpha}}\right) \to 0$, $1 - F(1, \alpha)\left(\frac{\epsilon^{-1/2\alpha(h,1)}}{\{\Gamma(1-\alpha)\}^{1/\alpha}}\right) \to 0$, and hence $\mathbb{P}\left(\frac{U_h}{V^{\alpha(h,1)}} \le \epsilon\right) \to 0$ as $\epsilon \to 0$. Using the similar arguments, we also have $\mathbb{P}\left(\frac{V^{\alpha(h,1)}}{U_h} \le \epsilon\right) \to 0$ as $\epsilon \to 0$. □

Proof of Theorem 4. From Theorem 3, for $h > \alpha$, by Slutsky's theorem, as $n \to \infty$,

$$\frac{\log M'_h}{\log M'_1} = \frac{\log\{M'_h/(M'_1)^{\alpha(h,1)}\}}{\log M'_1} + \frac{\log(M'_1)^{\alpha(h,1)}}{\log M'_1} \xrightarrow{p} \alpha(h, 1)$$

because M'_1 and $\log M'_1$ diverge to infinity, and $\log M'_h/(M'_1)^{\alpha(h,1)}$ is bounded in probability from Lemma B.2. Using this result for $h_1, h_2 > \alpha$ gives, by Slutsky's theorem, as $n \to \infty$,

$$\frac{\log M'_{h_2}}{\log M'_{h_1}} = \frac{\log M'_{h_2}}{\log M'_1}\frac{\log M'_1}{\log M'_{h_1}} \xrightarrow{p} \frac{h_2-\alpha}{1-\alpha}\frac{1-\alpha}{h_1-\alpha} = \frac{h_2-\alpha}{h_1-\alpha}.$$

□

To prove Corollary 4, we need a lemma.

Lemma B.3. *For all positive integers $j \le h$ and $h > 1$, and for any $0 < \alpha < 1$, $j - (h-j)\alpha(j,1) - \alpha(h,1) < 0$.*

Proof. (i) When $1 < j < h$,

$$\begin{aligned}j - (h-j)\alpha(j,1) - \alpha(h,1) &= \frac{j(1-\alpha)}{1-\alpha} - \frac{(h-j)(j-\alpha)}{1-\alpha} - \frac{h-\alpha}{1-\alpha}\\ &< \frac{h(1-\alpha)}{1-\alpha} - \frac{hj - h\alpha - j^2 + j\alpha}{1-\alpha} - \frac{h-\alpha}{1-\alpha} \quad \text{(because } j < h, (1-\alpha) > 0)\\ &= (1-\alpha)^{-1}\{h - h\alpha - hj + h\alpha + j^2 - j\alpha - h + \alpha\}\\ &= (1-\alpha)^{-1}\{h - hj + j^2 - j\alpha - h + \alpha\}\\ &< (1-\alpha)^{-1}\{h - j^2 + j^2 - j\alpha - h + \alpha\} \quad \text{(because } j < h)\\ &= (1-\alpha)^{-1}\{-j\alpha + \alpha\}\\ &= (1-\alpha)^{-1}\alpha\{-j + 1\}\\ &< 0.\end{aligned}$$

(ii) When $j = 1 \le h$, $\alpha(j, 1) = 1$. Since we assume $h > 1$, $2 - h \le 0$. Then

$$j - (h-j)\alpha(j,1) - \alpha(h,1) = 2 - h - \frac{h-\alpha}{1-\alpha} \le -\frac{h-\alpha}{1-\alpha} < 0.$$

(iii) When $0 < j = h$,

$$j - (h-j)\alpha(j,1) - \alpha(h,1) = h - \alpha(h,1) = h - \frac{h-\alpha}{1-\alpha} = \frac{h - h\alpha - h + \alpha}{1-\alpha} = \frac{\alpha(1-h)}{1-\alpha} < 0.$$

□

Proof of Corollary 4. Assuming that $h > 1$ is an integer, the binomial expansion gives

$$\frac{M_h}{(M_1')^{\alpha(h,1)}} = \frac{M_h'}{(M_1')^{\alpha(h,1)}} + \sum_{j=1}^{h}(-1)^j \binom{h}{j} \frac{(M_j')^{h-j}(M_1')^j}{(M_1')^{\alpha(h,1)}},$$

where

$$\frac{(M_j')^{h-j}(M_1')^j}{(M_1')^{\alpha(h,1)}} = (M_1')^{j-(h-j)\alpha(j,1)-\alpha(h,1)} \left(\frac{M_j'}{(M_1')^{\alpha(j,1)}} \right)^{h-j}.$$

Lemma B.3 shows that the exponent of M_1', namely $j - (h-j)\alpha(j,1) - \alpha(h,1)$, is negative for integers $1 \le j \le h$ and $h > 1$. Therefore as $n \to \infty$, $(M_1')^{j-(h-j)\alpha(j,1)-\alpha(h,1)} \xrightarrow{p} 0$. From Lemma B.2, $\{M_j'/(M_1')^{\alpha(j,1)}\}^{h-j}$ is bounded in probability. Therefore as $n \to \infty$

$$\left| \frac{M_h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}} \right| \le \sum_{j=1}^{h} \binom{h}{j} \frac{(M_j')^{h-j}(M_1')^j}{(M_1')^{\alpha(h,1)}} \xrightarrow{p} 0 \qquad [S.29]$$

and

$$\frac{M_h}{(M_1')^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}}.$$

52 □

Proof of Theorem 5. From Corollary 4, by Slutsky's theorem, as $n \to \infty$,

$$\frac{\log M_h}{\log M_1'} = \frac{\log\{M_h/(M_1')^{\alpha(h,1)}\}}{\log M_1'} + \frac{\log(M_1')^{\alpha(h,1)}}{\log M_1'} \xrightarrow{p} \alpha(h,1)$$

because M_1' and $\log M_1'$ diverge to infinity and $\log\{M_h/(M_1')^{\alpha(h,1)}\}$ is bounded in probability from Lemma B.2. Since $h_i, i = 1, 2$ are assumed to be positive integers, $h_i > 1, i = 1, 2$, so from the first result in Theorem 5, we derive the second result in Theorem 5 because, by Slutsky's theorem, as $n \to \infty$,

$$\frac{\log M_{h_2}}{\log M_{h_1}} = \frac{\log M_{h_2}}{\log M_1'} \frac{\log M_1'}{\log M_{h_1}} \xrightarrow{p} \frac{h_2 - \alpha}{1-\alpha} \frac{1-\alpha}{h_1 - \alpha} = \frac{h_2-\alpha}{h_1-\alpha}.$$

53 □

54 To prove Theorem 6, we need the following lemma.

Lemma B.4. *If $X \stackrel{d}{=} F_X$ such that $1 - F_X(x) \sim x^{-\alpha}\ell(x)$ where ℓ is a slowly varying function, i.e., $\lim_{x\to\infty} \ell(tx)/\ell(x) = 1$ for any $t > 0$, and further ℓ is such that $\lim_{x\to\infty} \ell(x) = L$, and if $h_2 \ge h_1 > \alpha$ are two positive real numbers, then as $n \to \infty$,*

$$\frac{M_{h_2}'}{(M_{h_1}')^{\alpha(h_2,h_1)}} \xrightarrow{d} L^{\frac{h_1-h_2}{h_1-\alpha}} \frac{U_{h_2}}{(U_{h_1})^{\alpha(h_2,h_1)}},$$

where the random vector (U_{h_1}, U_{h_2}) has a joint Laplace transform with $s > 0$, $t > 0$,

$$\mathbb{E}(e^{-sU_{h_2}-tU_{h_1}}) = \exp\left\{ -\int_0^\infty \{r_{h_2,h_1}(y,s,t)\}^{-\alpha} e^{-y}\, dy \right\},$$

55 *and $x = r_{h_2,h_1}(y,s,t)$ is the unique positive root of $sx^{h_2} + tx^{h_1} - y = 0$ for $y > 0$.*

Proof of Lemma B.4. Define the sequence $a_n, n = 1, 2, \ldots$ as the solutions of $n^{-1} = 1 - F_X(a_n)$. Then $a_n \to \infty$. Following the line of argument in the proof of Theorem 3, we integrate by parts, with $\theta > 0, \psi > 0$:

$$\begin{aligned} 1 - \mathbb{E}\left[e^{-\theta X_1^{h_2} - \psi X_1^{h_1}}\right] &= \int_0^\infty (1 - e^{-\theta x^{h_2} - \psi x^{h_1}})\, dF_X(x) \\ &= \int_0^\infty \{1 - F_X(x)\}(h_2\theta x^{h_2-1} + h_1\psi x^{h_1-1})e^{-\theta x^{h_2} - \psi x^{h_1}}\, dx. \end{aligned}$$

Define $y := \theta x^{h_2} + \psi x^{h_1}$. If $x > 0$, then $y > 0$ and

$$1 - \mathbb{E}\left[e^{-\theta X_1^{h_2} - \psi X_1^{h_1}}\right] = \int_0^\infty \left(1 - F_X\{r_{h_2,h_1}(y,\theta,\psi)\}\right) e^{-y}\, dy,$$

where $x = r_{h_2,h_1}(y,\theta,\psi)$ is the positive root of $\theta x^{h_2} + \phi x^{h_1} - y = 0$. Set $\theta = s/a_n^{h_2}$ and $\psi = t/a_n^{h_1}$. Because $\mathbb{E}(e^{-s(1/a_n^{h_2})X_1^{h_2} - t(1/a_n^{h_1})X_1^{h_1}}) \to 1$ as $n \to \infty$, we approximate $\log \mathbb{E}(e^{-s(1/a_n^{h_2})X_1^{h_2} - t(1/a_n^{h_1})X_1^{h_1}})$ by $\mathbb{E}(e^{-s(1/a_n^{h_2})X_1^{h_2} - t(1/a_n^{h_1})X_1^{h_1}}) - 1$ (by the first-order Taylor expansion). Then

$$\begin{aligned}
&\mathbb{E}\left(\exp\left\{-(s/a_n^{h_2})\sum_{i=1}^n X_i^{h_2} - (t/a_n^{h_1})\sum_{i=1}^n X_i^{h_1}\right\}\right) \\
&= \exp\{n \log \mathbb{E}(e^{-s(1/a_n^{h_2})X_1^{h_2} - t(1/a_n^{h_1})X_1^{h_1}})\} \\
&\sim \exp\{-n(1 - \mathbb{E}[e^{-s(1/a_n^{h_2})X_1^{h_2} - t(1/a_n^{h_1})X_1^{h_1}}])\} \\
&= \exp\left\{-\int_0^\infty n\left[1 - F_X\left\{r_{h_2,h_1}(y, s/a_n^{h_2}, t/a_n^{h_1})\right\}\right] e^{-y}\, dy\right\}.
\end{aligned}$$

From Lemma B.1, $a_n^{-1} r_{h_2,h_1}(y, s/a_n^{h_2}, t/a_n^{h_1}) = r_{h_2,h_1}(y,s,t)$, which implies that

$$1 - F_X\left\{r_{h_2,h_1}(y, s/a_n^{h_2}, t/a_n^{h_1})\right\} = 1 - F_X\{a_n a_n^{-1} r_{h_2,h_1}(y, s/a_n^{h_2}, t/a_n^{h_1})\} = 1 - F_X\{a_n r_{h_2,h_1}(y,s,t)\}.$$

Because $1 - F_X(x) \sim x^{-\alpha}\ell(x)$,

$$n[1 - F_X\{a_n r_{h_2,h_1}(y,s,t)\}] \sim n\{a_n r_{h_2,h_1}(y,s,t)\}^{-\alpha}\ell\{a_n r_{h_2,h_1}(y,s,t)\}.$$

From the definition of the slowly varying function, because the constant $r_{h_2,h_1}(y,s,t) > 0$ and $a_n \to \infty$ as $n \to \infty$, we have

$$n\{a_n r_{h_2,h_1}(y,s,t)\}^{-\alpha}\ell\{a_n r_{h_2,h_1}(y,s,t)\} \sim n\{a_n r_{h_2,h_1}(y,s,t)\}^{-\alpha}\ell(a_n) = n\{r_{h_2,h_1}(y,s,t)\}^{-\alpha} a_n^{-\alpha}\ell(a_n).$$

By our earlier definition, $a_n^{-\alpha}\ell(a_n) \sim 1 - F_X(a_n) = n^{-1}$. Therefore

$$n\{r_{h_2,h_1}(y,s,t)\}^{-\alpha} a_n^{-\alpha}\ell(a_n) \sim n\{r_{h_2,h_1}(y,s,t)\}^{-\alpha} n^{-1} = \{r_{h_2,h_1}(y,s,t)\}^{-\alpha}.$$

Hence,

$$\lim_{n\to\infty} \mathbb{E}\left(\exp\left\{-(s/a_n^{h_2})\sum_{i=1}^n X_i^{h_2} - (t/a_n^{h_1})\sum_{i=1}^n X_i^{h_1}\right\}\right) = \exp\left\{-\int_0^\infty \{r_{h_2,h_1}(y,s,t)\}^{-\alpha} e^{-y}\, dy\right\}.$$

Therefore, $\left(\frac{n}{a_n^{h_2}} M'_{h_2}, \frac{n}{a_n^{h_1}} M'_{h_1}\right)$ converges to (U_{h_2}, U_{h_1}) in distribution as $n \to \infty$. By Slutsky's theorem, for $h_1, h_2 > \alpha$,

$$\frac{M'_{h_2}}{(M'_{h_1})^{\alpha(h_2,h_1)}} = \frac{a_n^{h_2}/n}{\left(a_n^{h_1}/n\right)^{\alpha(h_2,h_1)}} \frac{\frac{n}{a_n^{h_2}} M'_{h_2}}{\left(\frac{n}{a_n^{h_1}} M'_{h_1}\right)^{\alpha(h_2,h_1)}} \xrightarrow{d} L^{\frac{h_1-h_2}{h_1-\alpha}} \frac{U_{h_2}}{(U_{h_1})^{\alpha(h_2,h_1)}}$$

as $n \to \infty$ because $\lim_{n\to\infty} \ell(a_n) = \lim_{x\to\infty} \ell(x) = L$. □

Proof of Theorem 6. Theorem 6 is a special case of Lemma B.4 with $L = \{\Gamma(1-\alpha)\}^{-1}$. □

Proof of Corollary 5. We first show that, for a positive integer $h > 1$, as $n \to \infty$,

$$\frac{M_h}{M'_h} \xrightarrow{p} 1. \qquad \text{[S.30]}$$

Indeed, the binomial expansion of M_h gives

$$\frac{M_h}{M'_h} = 1 + \sum_{j=1}^h (-1)^j \binom{h}{j} \frac{(M'_j)^{h-j}(M'_1)^j}{M'_h},$$

where

$$\frac{(M'_j)^{h-j}(M'_1)^j}{M'_h} = \frac{(M'_j)^{h-j}(M'_1)^j}{(M'_1)^{\alpha(h,1)}} \frac{(M'_1)^{\alpha(h,1)}}{M'_h}.$$

From Eq. (S.29), we have $(M'_j)^{h-j}(M'_1)^j/(M'_1)^{\alpha(h,1)} = o_p(1)$. From Lemma B.2, we also have $(M'_1)^{\alpha(h,1)}/M'_h = O_p(1)$. Therefore, $(M'_j)^{h-j}(M'_1)^j/M'_h = o_p(1)$ and

$$\left|\frac{M_h}{M'_h} - 1\right| \le \sum_{j=1}^h \binom{h}{j} \frac{(M'_j)^{h-j}(M'_1)^j}{M'_h} = \sum_{j=1}^h \binom{h}{j} o_p(1) O_p(1) = o_p(1).$$

This proves Eq. (S.30). We showed in Lemma B.4 and Theorem 6 that $\left(\frac{n}{a_n^{h_1}}M'_{h_1}, \frac{n}{a_n^{h_2}}M'_{h_2}\right) \xrightarrow{d} (U_1, U_2)$ as $n \to \infty$. So by the continuous mapping theorem, as $n \to \infty$,

$$\frac{nM'_{h_2}}{(nM'_{h_1})^{\frac{h_2}{h_1}}} = \frac{a_n^{h_2}}{(a_n^{h_1})^{\frac{h_2}{h_1}}} \frac{\frac{n}{a_n^{h_2}}M'_{h_2}}{\left(\frac{n}{a_n^{h_1}}M'_{h_1}\right)^{\frac{h_2}{h_1}}} = 1 \cdot \frac{\frac{n}{a_n^{h_2}}M'_{h_2}}{\left(\frac{n}{a_n^{h_1}}M'_{h_1}\right)^{\frac{h_2}{h_1}}} \xrightarrow{d} \frac{U_1}{(U_2)^{\frac{h_2}{h_1}}}.$$

Therefore, for positive integers $h_2 \geq h_1 > 1$, applying Eq. (S.30) and the result immediately above gives

$$n^{\frac{h_1-h_2}{h_1}} \frac{M_{h_2}}{(M_{h_1})^{h_2/h_1}} = \frac{M_{h_2}}{M'_{h_2}} \cdot \left(\frac{nM'_{h_2}}{(nM'_{h_1})^{\frac{h_2}{h_1}}}\right) \cdot \left(\frac{M'_{h_1}}{M_{h_1}}\right)^{\frac{h_2}{h_1}} \xrightarrow{d} 1 \cdot \frac{U_{h_1}}{(U_{h_2})^{\frac{h_2}{h_1}}} \cdot 1 = \frac{U_{h_1}}{(U_{h_2})^{\frac{h_2}{h_1}}},$$

as $n \to \infty$ by Slutsky's Theorem. Now

$$\frac{\log|M_{h_2}|}{\log|M_{h_1}|} = \frac{\log(|M_{h_2}/M'_{h_2}|) + \log M'_{h_2}}{\log(|M_{h_1}/M'_{h_1}|) + \log M'_{h_1}} = \frac{\frac{\log(|M_{h_2}/M'_{h_2}|)}{\log M'_{h_1}} + \frac{\log M'_{h_2}}{\log M'_{h_1}}}{\frac{\log(|M_{h_1}/M'_{h_1}|)}{\log M'_{h_1}} + 1}.$$

By the continuous mapping theorem, Eq. (S.30) implies $\log(|M_{h_1}/M'_{h_1}|) \xrightarrow{p} 0$ while $\log M'_{h_1} \xrightarrow{a.s.} \infty$ as $n \to \infty$. Therefore $\log(|M_{h_1}/M'_{h_1}|)/\log M'_{h_1} \xrightarrow{p} 0$ and $\log(|M_{h_2}/M'_{h_2}|)/\log M'_{h_1} \xrightarrow{p} 0$ as both are close to 0 with probability approaching to 1 as $n \to \infty$. Theorem 4 gives $(\log M'_{h_2})/\log M'_{h_1} \xrightarrow{p} \alpha(h_1, h_2)$ as $n \to \infty$. Therefore, another application of Slutsky's Theorem gives, as $n \to \infty$,

$$\frac{\log|M_{h_2}|}{\log|M_{h_1}|} \xrightarrow{p} \frac{0 + \alpha(h_2, h_1)}{0+1} = \alpha(h_2, h_1).$$

□

Proof of Theorem 7. This proof is a general version of the proof of Theorem 1. Denote $c_+ = \max(c, 0)$ for $c \in \mathbb{R}$. If $0 < X_i \leq a$, then, because Theorem 7 assumes $h > 0$,

$$[(M'_1 - X_i)_+]^h \geq [(M'_1 - a)_+]^h I(M'_1 > a) = (M'_1 - a)^h I(M'_1 > a),$$

hence

$$\frac{[(M'_1 - X_i)_+]^h}{M_1'^h} = \left[\left(1 - \frac{X_i}{M'_1}\right)_+\right]^h \geq \left(1 - \frac{a}{M'_1}\right)^h I(M'_1 > a). \quad \text{[S.31]}$$

Therefore

$$\begin{aligned} \frac{M_h^-}{M_1'^h} &= \frac{1}{n}\sum_{i=1}^{n} \frac{[(M'_1 - X_i)_+]^h}{(M'_1)^h} \\ &\geq \frac{N_n(a)}{n}\left(1 - \frac{a}{M'_1}\right)^h I(M'_1 > a). \end{aligned} \quad \text{[S.32]}$$

The inequality in Eq. (S.32) is obtained by omitting from the summation any term in which $a < X_i \leq M'_1$ and then using the inequality Eq. (S.31) to replace each term in which $X_i \leq a$ by its lower bound. Eq. (S.32) is a convenient lower bound.

By the strong law of large numbers for random variables with an infinite mean, $M'_1 \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$. Consequently, for fixed a, as $n \to \infty$, $I(M'_1 > a) \xrightarrow{\text{a.s.}} 1$, and

$$\left(1 - \frac{a}{M'_1}\right)^h \xrightarrow{\text{a.s.}} 1. \quad \text{[S.33]}$$

From Eq. (S.32), Eq. (S.5), Eq. (S.33), and Eq. (S.7), for all $a > 0$,

$$\liminf_{n\to\infty} \frac{M_h^-}{M_1'^h} \geq F(a) \quad \text{a.s.}. \quad \text{[S.34]}$$

Letting $a \to \infty$ in Eq. (S.34) gives

$$\liminf_{n\to\infty} \frac{M_h^-}{M_1'^h} \geq 1 \quad \text{a.s.}. \quad \text{[S.35]}$$

Brown et al. (1, p. 665, for $h = 2$) give an argument that is independent of $h > 0$ to show that for a given $M_1' > 0$, the maximal value of M_h^- is attained when any $n-1$ of the X_is equal 0 and the one remaining X_i equals nM_1'. For such values, $M_h^- = (n-1)M_1'^h/n \le M_1'^h$. Thus in all cases, $M_h^- \le M_1'^h$ and

$$\limsup_{n\to\infty} \frac{M_h^-}{(M_1')^h} \le 1 \quad \text{a.s..} \tag{S.36}$$

From Eq. (S.35) and Eq. (S.36), $M_h^-/(M_1')^h \xrightarrow{\text{a.s.}} 1$ and $\log M_n^- - h \log M_1' \xrightarrow{\text{a.s.}} 0$ as $n \to \infty$. □

Proof of Corollary 6. The first claim is a special case in Theorem 7 when $h = 1$. For the second claim, note that

$$\frac{M_h^-}{(M_1')^{\alpha(h,1)}} = \frac{M_h^-}{(M_1')^h} \frac{1}{(M_1')^{\alpha(h,1)-h}}. \tag{S.37}$$

From Theorem 7, the first factor on the right side of Eq. (S.37), $\frac{M_h^-}{(M_1')^h}$, converges a.s. to 1 as $n \to \infty$. In the second factor on the right side of Eq. (S.37), the exponent in the denominator is $\alpha(h,1) - h = \alpha \cdot (h-1)/(1-\alpha) > 0$ because $h > 1$ and $0 < \alpha < 1$. Hence $\frac{1}{(M_1')^{\alpha(h,1)-h}}$ converges to 0 a.s. as $n \to \infty$. According to Slutsky's theorem, the ratio in Eq. (S.37) converges to 0 a.s. as $n \to \infty$.

For the third claim, we write

$$\frac{\log M_h^-}{\log M_1'} = \frac{\log M_h^- - h\log M_1'}{\log M_1'} + \frac{h \log M_1'}{\log M_1'} = \frac{\log M_h^- - h\log M_1'}{\log M_1'} + h.$$

From Theorem 7, $\frac{\log M_h^- - h\log M_1'}{\log M_1'}$ converges to 0 a.s. as $n \to \infty$. Again, according to Slutsky's theorem, $\log M_h^-/\log M_1'$ converges to h a.s. as $n \to \infty$. By definition, $M_h^{-*} = nM_h^-/N_n^-$. Hence

$$\frac{\log M_h^{-*}}{\log M_1'} = \frac{\log n/N_n^-}{\log M_1'} + \frac{\log M_h^-}{\log M_1'}.$$

From Lemma 1, N_n^-/n converges to 0 a.s. as $n \to \infty$. Due to Slutsky's theorem, $\log M_h^{-*}/\log M_1'$ converges to h a.s. as $n \to \infty$. □

Theorem B.1. *Consider a random sample $X_1, \ldots, X_n$ from F_X satisfying* Eq. (9), *i.e., such that $X \overset{d}{\approx} F(c, \alpha)$. Then as $n \to \infty$,*

1. *for $0 < h < 1$,*

$$\frac{\frac{1}{n}\sum_{i=1}^n |X_i - M_1'|^h}{(M_1')^h} - 1 = o_p(1) \qquad \textit{and} \qquad \frac{M_h^+}{(M_1')^h} = o_p(1);$$

2. *for $h = 1$,*

$$\frac{\frac{1}{n}\sum_{i=1}^n |X_i - M_1'|^h}{(M_1')^h} \xrightarrow{a.s.} 1 \qquad \textit{and} \qquad \frac{M_1^+}{M_1'} \xrightarrow{a.s.} 1;$$

3. *for $h > 1$,*

$$\frac{\frac{1}{n}\sum_{i=1}^n |X_i - M_1'|^h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}} = o_p(1) \qquad \textit{and} \qquad \frac{M_h^+}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}} = o_p(1).$$

To prove this result, we establish a useful lemma.

Lemma B.5. *For any real numbers $0 < r \le 1$ and x and y, $\big||x|^r - |y|^r\big| \le \big|x - y\big|^r$.*

Proof. Here we apply the c_r-inequality (6, pp. 319-320, Theorem (8)): for any real numbers $x, y \in \mathbb{R}$ and $r > 0$, $|x+y|^r \le c_r(|x|^r + |y|^r)$ where $c_r = 1$ when $0 < r \le 1$ and $c_r = 2^{r-1}$ when $1 \le r < \infty$.

Lemma B.5 assumes $0 < r \le 1$. From the c_r-inequality, $|x|^r = |x - y + y|^r \le |x-y|^r + |y|^r$ and then $|x|^r - |y|^r \le |x-y|^r$. Exchanging x and y gives $|y|^r - |x|^r \le |x-y|^r$. The two inequalities imply that $\big||x|^r - |y|^r\big| \le \big|x-y\big|^r$. □

Proof of Theorem B.1. 1. When $0 < h < 1$, we write

$$\left|\frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^h}{(M_1')^h} - 1\right| = \frac{1}{(M_1')^h}\left|\frac{1}{n}\sum_{i=1}^{n}(|X_i - M_1'|^h - (M_1')^h)\right|.$$

From Lemma B.5,

$$\begin{aligned}\left|\frac{1}{n}\sum_{i=1}^{n}(|X_i - M_1'|^h - (M_1')^h)\right| &\le \frac{1}{n}\sum_{i=1}^{n}\left||M_1' - X_i|^h - (M_1')^h\right| \\ &\le \frac{1}{n}\sum_{i=1}^{n}\left|M_1' - X_i - M_1'\right|^h = \frac{1}{n}\sum_{i=1}^{n}X_i^h = M_h',\end{aligned}$$

where the penultimate equality follows because all $X_i \ge 0$. From Lemma B.2, $M_h'/(M_1')^{\alpha(h,1)} = O_p(1)$. Because $M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$,

$$\frac{M_h'}{(M_1')^h} = \frac{M_h'}{(M_1')^{\alpha(h,1)}}\frac{(M_1')^{\alpha(h,1)}}{(M_1')^h} = O_p(1)(M_1')^{\alpha(h,1)-h} = O_p(1)(M_1')^{\frac{\alpha(h-1)}{1-\alpha}} = o_p(1),$$

as the exponent is negative.

2. When $h = 1$, the identities $|X_i - M_1'| = (X_i - M_1')_+ + (M_1' - X_i)_+$ and $X_i - M_1' = (X_i - M_1')_+ - (M_1' - X_i)_+$ imply that $|X_i - M_1'| = (X_i - M_1')_+ + (M_1' - X_i)_+ = \{X_i - M_1' + (M_1' - X_i)_+\} + (M_1' - X_i)_+ = X_i - M_1' + 2(M_1' - X_i)_+$. Then

$$\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^h = \frac{1}{n}\sum_{i=1}^{n}(X_i - M_1') + \frac{2}{n}\sum_{i=1}^{n}(M_1' - X_i)_+ = \frac{2}{n}\sum_{i=1}^{n}(M_1' - X_i)_+ = 2M_1^-.$$

From Theorem 7, $M_1^-/M_1' \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$. So $n^{-1}\sum_{i=1}^{n}|X_i - M_1'|/M_1' \xrightarrow{\text{a.s.}} 2$ as $n \to \infty$. But $M_1^+ = M_1^-$ because $M_1 = M_1^+ - M_1^-$ and $M_1 = 0$ by definition Eq. (2). Thus $M_1^+/M_1' \xrightarrow{\text{a.s.}} 1$ as $n \to \infty$.

3. When $h > 1$, let $\lfloor h \rfloor$ be the largest integer not greater than h. If $h = \lfloor h \rfloor$, then as above,

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^h &= \frac{1}{n}\sum_{i=1}^{n}\{(X_i - M_1')_+\}^h + \frac{1}{n}\sum_{i=1}^{n}\{(M_1' - X_i)_+\}^h \\ &= \frac{1}{n}\sum_{i=1}^{n}(X_i - M_1')^h + \frac{1}{n}\sum_{i=1}^{n}\{1 + (-1)^h\}(M_1' - X_i)_+^h = M_h + \{1 + (-1)^h\}M_h^-.\end{aligned}$$

From Eq. (S.29) in the proof of Corollary 4 , we have $\left|\frac{M_h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}}\right| = o_p(1)$. On the other hand, from Theorem 7, $M_h^-/M_1'^h \xrightarrow{p} 1$ as $n \to \infty$. Therefore, as $n \to \infty$,

$$\left|\frac{2M_h^-}{(M_1')^{\alpha(h,1)}}\right| = \frac{2M_h^-}{(M_1')^{\alpha(h,1)}} = 2\frac{M_h^-}{(M_1')^h}M_1'^{h-\alpha(h,1)} \xrightarrow{p} 2 \cdot 1 \cdot 0 = 0$$

since $M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$ and $h - \alpha(h,1) < 0$ for $h > 1$. Hence

$$\begin{aligned}\left|\frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}}\right| &= \left|\frac{M_h + \{1 + (-1)^h\}M_h^-}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}}\right| \\ &\le \left|\frac{M_h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}}\right| + \left|\frac{2M_h^-}{(M_1')^{\alpha(h,1)}}\right| = o_p(1). \qquad \text{[S.38]}\end{aligned}$$

But if $h > \lfloor h \rfloor$, then

$$\begin{aligned}
&\left| \frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}} \right| = \frac{1}{n(M_1')^{\alpha(h,1)}} \left| \sum_{i=1}^{n}(|X_i - M_1'|^h - X_i^h) \right| \\
&= \frac{1}{n(M_1')^{\alpha(h,1)}} \left| \sum_{i=1}^{n}(|X_i - M_1'|^h - X_i^{h-\lfloor h \rfloor}|X_i - M_1'|^{\lfloor h \rfloor}) + (X_i^{h-\lfloor h \rfloor}|X_i - M_1'|^{\lfloor h \rfloor} - X_i^h) \right| \\
&= \frac{1}{n(M_1')^{\alpha(h,1)}} \left| \sum_{i=1}^{n} \left\{ (|X_i - M_1'|^{h-\lfloor h \rfloor} - X_i^{h-\lfloor h \rfloor})|X_i - M_1'|^{\lfloor h \rfloor} \right\} + \left\{ X_i^{h-\lfloor h \rfloor}(|X_i - M_1'|^{\lfloor h \rfloor} - X_i^{\lfloor h \rfloor}) \right\} \right| \\
&\leq \frac{1}{n(M_1')^{\alpha(h,1)}} \left| \sum_{i=1}^{n}(|X_i - M_1'|^{h-\lfloor h \rfloor} - X_i^{h-\lfloor h \rfloor})|X_i - M_1'|^{\lfloor h \rfloor} \right| + \frac{1}{n(M_1')^{\alpha(h,1)}} \left| \sum_{i=1}^{n} X_i^{h-\lfloor h \rfloor}(|X_i - M_1'|^{\lfloor h \rfloor} - X_i^{\lfloor h \rfloor}) \right| \\
&\leq \frac{1}{n(M_1')^{\alpha(h,1)}} \sum_{i=1}^{n} \left| |X_i - M_1'|^{h-\lfloor h \rfloor} - X_i^{h-\lfloor h \rfloor} \right| |X_i - M_1'|^{\lfloor h \rfloor} + \frac{1}{n(M_1')^{\alpha(h,1)}} \sum_{i=1}^{n} X_i^{h-\lfloor h \rfloor} \left| |X_i - M_1'|^{\lfloor h \rfloor} - X_i^{\lfloor h \rfloor} \right| \text{(because } X_i \geq 0) \\
&\leq \frac{1}{n(M_1')^{\alpha(h,1)}} \sum_{i=1}^{n} \left\{ \left| -M_1' \right|^{h-\lfloor h \rfloor} |X_i - M_1'|^{\lfloor h \rfloor} \right\} + \frac{1}{n(M_1')^{\alpha(h,1)}} \sum_{i=1}^{n} X_i^{h-\lfloor h \rfloor} | - M_1'|^{\lfloor h \rfloor} \\
&= \frac{(M_1')^{h-\lfloor h \rfloor}}{(M_1')^{\alpha(h,1)}} \left(\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^{\lfloor h \rfloor} \right) + \frac{(M_1')^{\lfloor h \rfloor}}{(M_1')^{\alpha(h,1)}} \left(\frac{1}{n}\sum_{i=1}^{n} X_i^{h-\lfloor h \rfloor} \right) \\
&= \frac{(M_1')^{h-\lfloor h \rfloor}(M_1')^{\alpha(\lfloor h \rfloor,1)}}{(M_1')^{\alpha(h,1)}} \left(\frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^{\lfloor h \rfloor}}{(M_1')^{\alpha(\lfloor h \rfloor,1)}} \right) + \frac{(M_1')^{\lfloor h \rfloor}(M_1')^{\alpha(h-\lfloor h \rfloor,1)}}{(M_1')^{\alpha(h,1)}} \left(\frac{\frac{1}{n}\sum_{i=1}^{n} X_i^{h-\lfloor h \rfloor}}{(M_1')^{\alpha(h-\lfloor h \rfloor,1)}} \right).
\end{aligned}$$

Since $\lfloor h \rfloor$ is a positive integer, from Eq. (S.38) and Lemma B.2, we have

$$\frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^{\lfloor h \rfloor}}{(M_1')^{\alpha(\lfloor h \rfloor,1)}} = \left(\frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^{\lfloor h \rfloor}}{(M_1')^{\alpha(\lfloor h \rfloor,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}} \right) + \frac{M_h'}{(M_1')^{\alpha(h,1)}} = o_p(1) + O_p(1) = O_p(1).$$

In addition, since $h - \lfloor h \rfloor > 0$, from Lemma B.2, we also have

$$O_p(1) = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i^{h-\lfloor h \rfloor}}{(M_1')^{\alpha(h-\lfloor h \rfloor,1)}} =: \frac{M_{h-\lfloor h \rfloor}'}{(M_1')^{\alpha(h-\lfloor h \rfloor,1)}}.$$

Therefore, we claim

$$\left| \frac{\frac{1}{n}\sum_{i=1}^{n}|X_i - M_1'|^h}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}} \right| \leq (M_1')^{h-\lfloor h \rfloor+\alpha(\lfloor h \rfloor,1)-\alpha(h,1)} O_p(1) + (M_1)^{\lfloor h \rfloor+\alpha(h-\lfloor h \rfloor,1)-\alpha(h,1)} O_p(1) = o_p(1).$$

To verify this claim, we check that $h - \lfloor h \rfloor + \alpha(\lfloor h \rfloor, 1) - \alpha(h, 1) < 0$ and $\lfloor h \rfloor + \alpha(h - \lfloor h \rfloor, 1) - \alpha(h, 1) < 0$ for $h > \lfloor h \rfloor$:

$$\begin{aligned}
h - \lfloor h \rfloor + \alpha(\lfloor h \rfloor, 1) - \alpha(h, 1) &= h - \lfloor h \rfloor + \frac{\lfloor h \rfloor - \alpha}{1-\alpha} - \frac{h-\alpha}{1-\alpha} \\
&= \frac{(h - \lfloor h \rfloor)(1-\alpha)}{1-\alpha} + \frac{\lfloor h \rfloor - \alpha}{1-\alpha} - \frac{h-\alpha}{1-\alpha} \\
&= \frac{h - h\alpha - \lfloor h \rfloor + \alpha\lfloor h \rfloor + \lfloor h \rfloor - \alpha - h + \alpha}{1-\alpha} \\
&= \frac{-h\alpha + \alpha\lfloor h \rfloor}{1-\alpha} = \frac{\alpha(\lfloor h \rfloor - h)}{1-\alpha} < 0
\end{aligned}$$

and

$$\begin{aligned}
\lfloor h \rfloor + \alpha(h - \lfloor h \rfloor, 1) - \alpha(h, 1) &= \lfloor h \rfloor + \frac{h - \lfloor h \rfloor - \alpha}{1-\alpha} - \frac{h-\alpha}{1-\alpha} \\
&= \frac{\lfloor h \rfloor - \alpha\lfloor h \rfloor + h - \lfloor h \rfloor - \alpha - h + \alpha}{1-\alpha} = \frac{-\alpha\lfloor h \rfloor}{1-\alpha} < 0.
\end{aligned}$$

Because $M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$, the last claimed equality holds.

□

Proof of Theorem 8. For $0 < h < 1$, from Theorem B.1, $M_h^+/(M_1')^h = o_p(1)$. For $h = 1$, from Theorem B.1, $M_h^+/M_1' \xrightarrow{p} 1$ as $n \to \infty$. For $h > 1$, from Theorem B.1, $M_h^+/(M_1')^{\alpha(h,1)} - M_h'/(M_1')^{\alpha(h,1)} = o_p(1)$. From Theorem 3, $M_h'/(M_1')^{\alpha(h,1)} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} U_h/V^{\alpha(h,1)}$ as $n \to \infty$. Therefore, by Slutsky's theorem, for $h > 1$, as $n \to \infty$,

$$\frac{M_h^+}{(M_1')^{\alpha(h,1)}} = \left(\frac{M_h^+}{(M_1')^{\alpha(h,1)}} - \frac{M_h'}{(M_1')^{\alpha(h,1)}}\right) + \frac{M_h'}{(M_1')^{\alpha(h,1)}} \xrightarrow{d} 0 + \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}}.$$

Consequently,

$$\log M_h^+ - \alpha(h,1)\log M_1' = O_p(1). \qquad \text{[S.39]}$$

Dividing both sides of Eq. (S.39) by $\log M_1'$, and employing Slutsky's theorem, (Arnold (4, p. 242, Corollary 6.8(c)))

$$\frac{\log M_h^+}{\log M_1'} - \alpha(h,1) \xrightarrow{p} 0, \qquad \text{[S.40]}$$

because $\log M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$. Definition 3 of the local upper centered moments is

$$M_h^{+*} = M_h^+ \left(\frac{n}{N_n^+}\right). \qquad \text{[S.41]}$$

Eq. (S.25) gives $[\log(n/N_n^+)]/\log M_1' \xrightarrow{p} \alpha$ as $n \to \infty$. Therefore, from Eq. (S.25) and Eq. (S.40), as $n \to \infty$,

$$\frac{\log M_h^{+*}}{\log M_1'} = \frac{\log M_h^+}{\log M_1'} + \frac{\log(n/N_n^+)}{\log M_1'} \xrightarrow{p} \alpha(h,1) + \alpha = \frac{h-\alpha^2}{1-\alpha}.$$

□

Proof of Corollary 7. For fixed $c > 0$, $p > 1$, and $q > 1$, if $M_1' > c$, then

$$\Phi_{\mathrm{FT}}^n(c,p,q) := \frac{[\frac{1}{n}\sum_{i=1}^n[(X_i - c)_+]^p]^{1/p}}{[\frac{1}{n}\sum_{i=1}^n[(c - X_i)_+]^q]^{1/q}} \geq \frac{[\frac{1}{n}\sum_{i=1}^n[(X_i - M_1')_+]^p]^{1/p}}{[\frac{1}{n}\sum_{i=1}^n[(M_1' - X_i)_+]^q]^{1/q}} =: \Phi_{\mathrm{FT}}^n(M_1',p,q).$$

Then, given $C > 0$, we have

$$\begin{aligned}
\mathbb{P}\left(\Phi_{\mathrm{FT}}^n(c,p,q) > C\right) &= \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(c,p,q) > C, M_1' > c\right) + \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(c,p,q) > C, M_1' \leq c\right) \\
&\geq \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(M_1',p,q) > C, M_1' > c\right) + \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(c,p,q) > C, M_1' \leq c\right) \\
&\geq \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(M_1',p,q) > C\right) + \mathbb{P}\left(M_1' > c\right) - 1 \\
&\quad + \mathbb{P}\left(M_1' \leq c\right) - \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(c,p,q) \leq C, M_1' \leq c\right) \\
&\geq \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(M_1',p,q) > C\right) - \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(c,p,q) \leq C, M_1' \leq c\right) \\
&\geq \mathbb{P}\left(\Phi_{\mathrm{FT}}^n(M_1',p,q) > C\right) - \mathbb{P}\left(M_1' \leq c\right).
\end{aligned}$$

Because $M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$, $\mathbb{P}(M_1' \leq c) \to 0$. Therefore, to show that $\mathbb{P}(\Phi_{\mathrm{FT}}^n(c,p,q) > C) \to 1$ as $n \to \infty$, it suffices to show that $\Phi_{\mathrm{FT}}^n(M_1',p,q) \xrightarrow{p} \infty$ as $n \to \infty$. To do so, we write

$$\Phi_{\mathrm{FT}}^n(M_1',p,q) = \frac{(M_p^+)^{1/p}}{(M_q^-)^{1/q}} = \left(\frac{M_p^+}{M_1'^{\alpha(p,1)}}\right)^{1/p} \cdot (M_1')^{\frac{\alpha(p,1)}{p}-1} \cdot \left(\frac{M_1'^q}{M_q^-}\right)^{1/q}.$$

From Lemma B.2, $(M_p^+/M_1'^{\alpha(p,1)})^{1/p} = O_p(1)$. From Theorem 7, $\left(M_1'^q/M_q^-\right)^{1/q} \xrightarrow{\text{a.s.}} 1$. Because $\{\alpha(p,1)/p\} - 1 = \alpha(p-1)/\{p(1-\alpha)\} > 0$ and because $M_1' \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$, we have $(M_1')^{\frac{\alpha(p,1)}{p}-1} \xrightarrow{\text{a.s.}} \infty$ as $n \to \infty$. Thus $\Phi_{\mathrm{FT}}^n(M_1',p,q) \xrightarrow{p} \infty$ as $n \to \infty$, as desired. □

Proof of Corollary 8. Theorem 8 gives $(\log M_p^+)/\log M_1' \xrightarrow{p} \alpha(p,1)$ as $n \to \infty$. Because $q > 0$, Corollary 6.2 gives $(\log M_1')/\log M_q^- \xrightarrow{p} 1/q$ as $n \to \infty$. Then by Slutsky's theorem, as $n \to \infty$,

$$\frac{\log M_p^+}{\log M_q^-} = \frac{\log M_p^+}{\log M_1'}\frac{\log M_1'}{\log M_q^-} \xrightarrow{p} \alpha(p,1) \cdot \frac{1}{q} = \frac{p-\alpha}{q(1-\alpha)}.$$

□

C. Proofs in Section 5: number of observations that exceed the sample mean

Lemma C.1. *Assume that two independent random samples* $\{X_i\}_{i=1}^n$ *and* $\{X_i^*\}_{i=1}^n$ *from* $F_X \overset{d}{\approx} F(1,\alpha)$ *with* $\alpha \in (0,1)$ *have sample means,* M_1' *and* $M_1'^*$*, respectively. Define* $\overline{N}_n^*(a) := \#\{X_i^* > a | i \in \{1,\ldots,n\}\}$ *and recall the definition* $\overline{N}_n(a) := \#\{X_i > a | i \in \{1,\ldots,n\}\}$*. Then*

$$\frac{\overline{N}_n(M_1')}{n^\alpha} - \frac{\overline{N}_n^*(M_1')}{n^\alpha} = o_p(1).$$

We clarify that, in <u>*both*</u> *terms on the left side above,* M_1' *is the sample mean of the* <u>*first*</u> *sample* $\{X_i\}_{i=1}^n$ *and* $\overline{N}_n^*(M_1')$ *counts how many members of the* <u>*second*</u> *sample* $\{X_i^*\}_{i=1}^n$ *exceed the sample mean* M_1' *of the* <u>*first*</u> *sample.*

Proof of Lemma C.1. Define $\overline{F}_n$ and $\overline{F}_n^*$ to be the empirical survival functions of the random samples $\{X_i\}_{i=1}^n$ and $\{X_i^*\}_{i=1}^n$, respectively. By the definitions, $\overline{N}_n(M_1')/n^\alpha = \overline{F}_n(M_1')n^{1-\alpha}$ and $\overline{N}_n^*(M_1')/n^\alpha = \overline{F}_n^*(M_1')n^{1-\alpha}$. Given $\epsilon > 0$, recall $b_n := q_\epsilon n^{(1-\alpha)/\alpha}$ as defined in Lemma A.1, where $q_\epsilon := F^{-1}(1,\alpha)(\epsilon)$ and $F^{-1}(1,\alpha)$ is the quantile function for $F(1,\alpha)$. Then

$$\begin{aligned}
&\mathbb{P}\left(\left|\overline{N}_n(M_1')/n^\alpha - \overline{N}_n^*(M_1')/n^\alpha\right| > \epsilon\right) \\
&= \mathbb{P}\left(\left|\overline{F}_n(M_1')n^{1-\alpha} - \overline{F}_n^*(M_1')n^{1-\alpha}\right| > \epsilon\right) \\
&= \mathbb{P}\left(\left|\overline{F}_n(M_1')n^{1-\alpha} - \overline{F}_n^*(M_1')n^{1-\alpha}\right| > \epsilon, M_1' \le b_n\right) \\
&\quad + \mathbb{P}\left(\left|\overline{F}_n(M_1')n^{1-\alpha} - \overline{F}_n^*(M_1')n^{1-\alpha}\right| > \epsilon, M_1' > b_n\right) \\
&\le \mathbb{P}\left(M_1' \le b_n\right) + \mathbb{P}\left(\left|\overline{F}_n(M_1')n^{1-\alpha} - \overline{F}_n^*(M_1')n^{1-\alpha}\right| > \epsilon, M_1' > b_n\right).
\end{aligned}$$

As $n \to \infty$,

$$\mathbb{P}\left(M_1' \le b_n\right) = \mathbb{P}\left(\frac{\sum_{i=1}^n X_i}{n} \le q_\epsilon n^{\frac{1-\alpha}{\alpha}}\right) = \mathbb{P}\left(\frac{\sum_{i=1}^n X_i}{n^{1/\alpha}} \le q_\epsilon\right) \to F(1,\alpha)(q_\epsilon) = \epsilon$$

because $\sum_{i=1}^n X_i/n^{1/\alpha} \overset{d}{\to} F(1,\alpha)$ as $n \to \infty$ (Albrecher et al. (3, Remark 2.1)). On the other hand,

$$\begin{aligned}
&\mathbb{P}\left(\left|\overline{F}_n(M_1')n^{1-\alpha} - \overline{F}_n^*(M_1')n^{1-\alpha}\right| > \epsilon, M_1' > b_n\right) \\
&\le \mathbb{P}\left(\sup_{t>b_n}\left|\overline{F}_n(t)n^{1-\alpha} - \overline{F}_n^*(t)n^{1-\alpha}\right| > \epsilon, M_1' > b_n\right) \\
&\le \mathbb{P}\left(\sup_{t>b_n}\left|\overline{F}_n(t)n^{1-\alpha} - \overline{F}_n^*(t)n^{1-\alpha}\right| > \epsilon\right) \\
&\le \mathbb{P}\left(\sup_{t>b_n}\left|\overline{F}_n(t)n^{1-\alpha} - \overline{F}_X(t)n^{1-\alpha}\right| > \epsilon/2\right) + \mathbb{P}\left(\sup_{t>b_n}\left|\overline{F}_X(t)n^{1-\alpha} - \overline{F}_n^*(t)n^{1-\alpha}\right| > \epsilon/2\right) \\
&= 2\mathbb{P}\left(\sup_{t>b_n}\left|\overline{F}_n(t)n^{1-\alpha} - \overline{F}_X(t)n^{1-\alpha}\right| > \epsilon/2\right).
\end{aligned}$$

The last equality holds because the two samples are identically distributed. From Lemma A.1, $\sup_{t>b_n}\left|\overline{F}_n(t)n^{1-\alpha} - \overline{F}_X(t)n^{1-\alpha}\right| \overset{p}{\to} 0$ as $n \to \infty$. Hence $2\mathbb{P}\left(\sup_{t>b_n}\left|\overline{F}_n(t)n^{1-\alpha} - \overline{F}_X(t)n^{1-\alpha}\right| > \epsilon/2\right) \to 0$ as $n \to \infty$. The claim follows. □

Proof of Theorem 9. Let $\{X_1,\ldots,X_n\}$ and $\{X_1^*,\ldots,X_n^*\}$ be two independent random samples from the same distribution F_X satisfying Eq. (9) with sample means M_1' and $M_1'^*$, respectively. Lemma C.1 shows that

$$\frac{\sum I(X_i > M_1')}{n^\alpha} - \frac{\sum I(X_i^* > M_1')}{n^\alpha} = o_p(1).$$

Because both samples have the same distribution, for all $t \in \mathbb{R}$,

$$\mathbb{P}\left(\frac{\sum I(X_i^* > M_1')}{n^\alpha} \le t\right) = \mathbb{P}\left(\frac{\sum I(X_i > M_1'^*)}{n^\alpha} \le t\right).$$

Consequently, Lemma C.1 justifies the second equality below:

$$\begin{aligned}
\frac{\sum I(X_i > M_1')}{n^\alpha} &= \frac{\sum I(X_i > M_1')}{n^\alpha} - \frac{\sum I(X_i^* > M_1')}{n^\alpha} + \frac{\sum I(X_i^* > M_1')}{n^\alpha} \\
&= o_p(1) + \frac{\sum I(X_i^* > M_1')}{n^\alpha}.
\end{aligned}$$

Because $n^{-\alpha}\sum I(X_i > M_1')$ and $n^{-\alpha}\sum I(X_i^* > M_1')$ are identically distributed for every finite sample size n, they must have the same limiting distribution.

Define

$$
\begin{aligned}
U_n &:= n^{-\left(\frac{1-\alpha}{\alpha}\right)} M_1'^*, \\
Z_n^* &:= n^{-\alpha} \sum_{i=1}^n I\left(X_i > n^{\left(\frac{1-\alpha}{\alpha}\right)} U_n\right) = n^{-\alpha} \sum_{i=1}^n I(X_i > M_1'^*).
\end{aligned}
$$

Then as $n \to \infty$, $U_n \xrightarrow{d} F(1, \alpha)$. The conditional distribution of $n^\alpha Z_n^*$ given $U_n = u$ is

$$
\text{Binomial}\left(n, \overline{F}_X\left(n^{\frac{1-\alpha}{\alpha}} u\right)\right).
$$

For any $t > 0$ and for any $\epsilon > 0$, the Laplace transform of Z_n^* satisfies

$$
\mathbb{E}\left[e^{-tZ_n^*}\right] = \mathbb{E}\left[e^{-tZ_n^*} I(U_n \le \epsilon)\right] + \mathbb{E}\left[e^{-tZ_n^*} I(U_n > \epsilon)\right].
$$

For the first term on the right side, $Z_n^* \ge 0$ implies $\mathbb{E}\left[e^{-tZ_n^*} I(U_n \le \epsilon)\right] \le \mathbb{E}\left[I(U_n \le \epsilon)\right] = \mathbb{P}(U_n \le \epsilon)$. For the second term on the right side, using the tower property,

$$
\mathbb{E}\left[e^{-tZ_n^*} I(U_n > \epsilon)\right] = \mathbb{E}\left[\mathbb{E}\left[e^{-tZ_n^*} I(U_n > \epsilon)\middle| U_n\right]\right].
$$

If $u \le \epsilon$, $\mathbb{E}\left[e^{-tZ_n^*} I(U_n > \epsilon) | U_n = u\right] = 0$; and if $u > \epsilon$, then $\mathbb{E}\left[e^{-tZ_n^*} I(U_n > \epsilon) | U_n = u\right] = \mathbb{E}\left[e^{-tZ_n^*} | U_n = u\right]$ and

$$
\begin{aligned}
& \mathbb{E}\left[e^{-tZ_n^*}\middle| U_n = u\right] = \mathbb{E}\left[e^{-\frac{t}{n^\alpha}(n^\alpha Z_n^*)}\middle| U_n = u\right] \\
&= \left(F_X(n^{\frac{1-\alpha}{\alpha}} u) + \overline{F}_X(n^{\frac{1-\alpha}{\alpha}} u) e^{-\frac{t}{n^\alpha}}\right)^n \\
&= \left(1 - (1 - e^{-\frac{t}{n^\alpha}}) \overline{F}_X(n^{\frac{1-\alpha}{\alpha}} u)\right)^n \text{ (due to the conditional Binomial distribution of } Z_n^* | U_n = u) \\
&= \left[1 - \left\{\frac{t}{n^\alpha} + R_{1n}(t)\right\}\left\{\frac{n^{\alpha-1} u^{-\alpha}}{\Gamma(1-\alpha)} + R_{2n}(u)\right\}\right]^n \\
&= \left[1 - \frac{t u^{-\alpha}}{n \Gamma(1-\alpha)} - \frac{t}{n^\alpha} R_{2n}(u) - \frac{n^{\alpha-1} u^{-\alpha}}{\Gamma(1-\alpha)} R_{1n}(t) - R_{1n}(t) R_{2n}(u)\right]^n.
\end{aligned}
$$

Here $|R_{1n}(t)| \le t^2/(2n^{2\alpha})$ by the mean value theorem and $|R_{2n}(u)| \le \epsilon n^{\alpha-1} u^{-\alpha}/\Gamma(1-\alpha)$ because of Eq. (9) for large enough n (also see Albrecher et al. (3)). Then

$$
\left|\frac{t}{n^\alpha} R_{2n}(u)\right| \le \frac{\epsilon t}{\Gamma(1-\alpha)} u^{-\alpha} n^{-1}, \qquad \left|\frac{n^{\alpha-1} u^{-\alpha}}{\Gamma(1-\alpha)} R_{1n}(t)\right| \le \frac{t^2}{2\Gamma(1-\alpha)} u^{-\alpha} n^{-\alpha-1},
$$

$$
|R_{1n}(t) R_{2n}(u)| \le \frac{\epsilon t^2}{2\Gamma(1-\alpha)} u^{-\alpha} n^{-\alpha-1}.
$$

Let $R_{3n}(u,t) := n\left\{\frac{t}{n^\alpha} R_{2n}(u) + \frac{n^{\alpha-1} u^{-\alpha}}{\Gamma(1-\alpha)} R_{1n}(t) + R_{1n}(t) R_{2n}(u)\right\}$. Then

$$
|R_{3n}(u,t)| \le \frac{2t\epsilon + t^2 n^{-\alpha} + t^2 \epsilon n^{-\alpha}}{2\Gamma(1-\alpha)} u^{-\alpha} \le \frac{2t\epsilon + t\epsilon + t\epsilon}{2\Gamma(1-\alpha)} u^{-\alpha} = \frac{4t\epsilon}{\Gamma(1-\alpha)} u^{-\alpha}
$$

for a large enough n such that $tn^{-\alpha} \le \min\{\epsilon, 1\}$. Therefore

$$
\mathbb{E}\left[e^{-tZ_n^*}\middle| U_n = u\right] = \left[1 - \frac{1}{n}\left\{\frac{t u^{-\alpha}}{\Gamma(1-\alpha)} + R_{3n}(u,t)\right\}\right]^n,
$$

and

$$
\left[1 - \frac{u^{-\alpha}(t + 4t\epsilon)}{\Gamma(1-\alpha)}\right]^n \le \mathbb{E}\left[e^{-tZ_n^*}\middle| U_n = u\right] \le \left[1 - \frac{u^{-\alpha}(t - 4t\epsilon)}{\Gamma(1-\alpha)}\right]^n.
$$

Hence

$$
\mathbb{E}\left\{\left[1 - \frac{U_n^{-\alpha}(t + 4t\epsilon)}{\Gamma(1-\alpha)}\right]^n I(U_n > \epsilon)\right\} \le \mathbb{E}\left[e^{-tZ_n^*} I(U_n > \epsilon)\right] \le \mathbb{E}\left\{\left[1 - \frac{U_n^{-\alpha}(t - 4t\epsilon)}{\Gamma(1-\alpha)}\right]^n I(U_n > \epsilon)\right\}.
$$

Applying the monotone convergence theorem gives

$$\lim_{n\to\infty} \mathbb{E}\left\{\left[1-\frac{U_n^{-\alpha}(t+4t\epsilon)}{\Gamma(1-\alpha)}\right]^n I(U_n>\epsilon)\right\} \le \lim_{n\to\infty} \mathbb{E}\left[e^{-tZ_n^*} I(U_n>\epsilon)\right] \le \lim_{n\to\infty} \mathbb{E}\left\{\left[1-\frac{U_n^{-\alpha}(t-4t\epsilon)}{\Gamma(1-\alpha)}\right]^n I(U_n>\epsilon)\right\}.$$

Because $(1-t/n)^n \to \exp(-t)$ uniformly over $[0,\infty)$ and $U_n \xrightarrow{d} U \overset{d}{=} F(1,\alpha)$ as $n\to\infty$, the bounded convergence theorem yields

$$\mathbb{E}\left[\exp\left\{-\frac{U^{-\alpha}(t+4t\epsilon)}{\Gamma(1-\alpha)}\right\} I(U>\epsilon)\right] \le \lim_{n\to\infty} \mathbb{E}\left[e^{-tZ_n^*} I(U_n>\epsilon)\right] \le \mathbb{E}\left[\exp\left\{-\frac{U^{-\alpha}(t-4t\epsilon)}{\Gamma(1-\alpha)}\right\} I(U>\epsilon)\right].$$

Then, for ϵ small enough that $t-4t\epsilon>0$, as $n\to\infty$ and $\epsilon\to 0$,

$$\begin{aligned}
0 &\le \mathbb{E}\left[\exp\left\{-\frac{U^{-\alpha}(t+4t\epsilon)}{\Gamma(1-\alpha)} I(U\le\epsilon)\right\}\right] \le\ \mathbb{E}\left[I(U\le\epsilon)\right] = \mathbb{P}(U\le\epsilon)\to 0,\\
0 &\le \mathbb{E}\left[\exp\left\{-\frac{U^{-\alpha}(t-4t\epsilon)}{\Gamma(1-\alpha)} I(U\le\epsilon)\right\}\right] \le\ \mathbb{E}\left[I(U\le\epsilon)\right] = \mathbb{P}(U\le\epsilon)\to 0,\\
0 &\le \mathbb{E}\left[e^{-tZ_n^*} I(U_n\le\epsilon)\right] \le \mathbb{E}\left[I(U_n\le\epsilon)\right] = \mathbb{P}(U_n\le\epsilon)\to\mathbb{P}(U\le\epsilon)\to 0.
\end{aligned}$$

By the bounded convergence theorem, as $\epsilon\to 0$,

$$\begin{aligned}
\mathbb{E}\left[\exp\left\{-\frac{U^{-\alpha}(t+4t\epsilon)}{\Gamma(1-\alpha)}\right\} I(U>\epsilon)\right] &\to \mathbb{E}\left[\exp\left\{-t\left(\frac{U^{-\alpha}}{\Gamma(1-\alpha)}\right)\right\}\right],\\
\mathbb{E}\left[\exp\left\{-\frac{U^{-\alpha}(t-4t\epsilon)}{\Gamma(1-\alpha)}\right\} I(U>\epsilon)\right] &\to \mathbb{E}\left[\exp\left\{-t\left(\frac{U^{-\alpha}}{\Gamma(1-\alpha)}\right)\right\}\right].
\end{aligned}$$

Thus we have shown that

$$\lim_{n\to\infty} \mathbb{E}\left[e^{-tZ_n^*}\right] = \mathbb{E}\left[\exp\left\{-t\left(\frac{U^{-\alpha}}{\Gamma(1-\alpha)}\right)\right\}\right].$$

Because the Laplace transform is unique, we conclude that, as $n\to\infty$,

$$Z_n^* \xrightarrow{d} \frac{U^{-\alpha}}{\Gamma(1-\alpha)}.$$

□

Proof of Theorem 10. To prove claim 1, from $\varepsilon \overset{d}{=} \mathrm{Exp}(1)$, we have, for any $x\ge 0$, $\mathbb{P}(\varepsilon>x)=e^{-x}$. Hence

$$\mathbb{P}\left(\frac{\varepsilon}{U}>t\right) = \mathbb{P}(\varepsilon>tU) = \mathbb{E}\left[e^{-tU}\right] = \mathscr{L}(t) = e^{-t^\alpha},$$

since $U \overset{d}{=} F(1,\alpha)$. Because U and ε are independent, it follows that

$$\mathbb{P}\left(\left(\frac{\varepsilon}{U}\right)^\alpha > t\right) = \mathbb{P}\left(\frac{\varepsilon}{U}>t^{1/\alpha}\right) = e^{-\left(t^{1/\alpha}\right)^\alpha} = e^{-t}.$$

Thus $U^{-\alpha}\varepsilon^\alpha \overset{d}{=} \mathrm{Exp}(1)$, which proves claim 1.

To prove claim 2, define $W_\alpha := \varepsilon^\alpha$. Then, for any $x\ge 0$, x not necessarily an integer, by the definition of the gamma function,

$$\mathbb{E}(W_\alpha^x) = \mathbb{E}(\varepsilon^{x\alpha}) = \int_0^\infty t^{x\alpha} e^{-t} dt = \Gamma(1+x\alpha).$$

Now, from claim 1, $(\mathrm{Exp}(1))^x \overset{d}{=} U^{-x\alpha}\varepsilon^{x\alpha}$, and by independence,

$$\begin{aligned}
&\mathbb{E}(U^{-x\alpha}W_\alpha^x) = \mathbb{E}(U^{-x\alpha})\mathbb{E}(W_\alpha^x)\\
&= \mathbb{E}(U^{-x\alpha})\Gamma(1+x\alpha) = \mathbb{E}[(\mathrm{Exp}(1))^x] = \Gamma(1+x)
\end{aligned}$$

or

$$\mathbb{E}(U^{-x\alpha}) = \frac{\Gamma(1+x)}{\Gamma(1+x\alpha)} = \frac{\Gamma(1+x)}{\mathbb{E}(W_\alpha^x)}, \quad x\ge 0.$$

To prove claim 3, again defining $W_\alpha := \varepsilon^\alpha$, we have

$$\frac{\mathbb{E}[V^2]}{2(\mathbb{E}[V])^2} = \frac{\{\Gamma(1+\alpha)\}^2}{\Gamma(1+2\alpha)} = \frac{(\mathbb{E}[W_\alpha])^2}{\mathbb{E}[W_\alpha^2]} < 1.$$

The final inequality holds because $W \geq 0$ and for $K \geq 2$, $\mathbb{E}[W_\alpha^K] \geq (\mathbb{E}[W_\alpha])^K$ with equality if and only if W is constant. Since W is not constant, the inequality is strict. Thus

$$\begin{aligned}\mathbb{E}[V^2] < 2(\mathbb{E}[V])^2 &\iff \mathrm{Var}(V) < (\mathbb{E}[V])^2 \\ &\iff \mathrm{SD}(V) < \mathbb{E}[V].\end{aligned}$$

To prove claim 4, since $\mathbb{E}(W_\alpha^K) = \Gamma(1+K\alpha)$, we have, for $K \geq 2$,

$$\frac{\mathbb{E}[V^K]}{K!(\mathbb{E}[V])^K} = \frac{\Gamma(1+\alpha)^K}{\Gamma(1+K\alpha)} = \frac{(\mathbb{E}[W_\alpha])^K}{\mathbb{E}[W_\alpha^K]} < 1.$$

The argument for the strict inequality is the same as in the proof of claim 3.

To prove claim 5, we use a random series representation of stable laws (Samorodnitsky and Taqqu (7, p. 22, their Prop. 1.4.1)),

$$U \overset{d}{=} \{\Gamma(1-\alpha)\}^{-1/\alpha} \sum_{j=1}^{\infty} S_j^{-1/\alpha},$$

where $\{S_j\}$ are the event times from a Poisson process with rate 1 and $S_1 \overset{d}{=} \varepsilon$ is Exp(1). It follows that

$$U \geq_{st} \{\Gamma(1-\alpha)\}^{-1/\alpha} S_1^{-1/\alpha} = \{\Gamma(1-\alpha)\}^{-1/\alpha} \varepsilon^{-1/\alpha}.$$

Then $U^{-\alpha} \leq_{st} \Gamma(1-\alpha)\varepsilon$, which implies $V = U^{-\alpha}/\Gamma(1-\alpha) \leq_{st} \varepsilon$. Thus $\mathbb{P}(V > t) \leq e^{-t}$ for all $t > 0$. □

Proof of Corollary 9. Claim 1 of Theorem 10 shows that $U^{-\alpha}\varepsilon^\alpha \overset{d}{\to} \mathrm{Exp}(1)$. As $\alpha \to 0$, we have $\varepsilon^\alpha \overset{p}{\to} 1$. By Slutsky's theorem, $U^{-\alpha} \overset{d}{\to} \mathrm{Exp}(1)$. □

D. Asymptotic properties of the modified financial ratios

A real-valued function f of real n-vectors ($n \geq 1$) is defined to be quasi-concave if, and only if, for all $\underline{X} := \{X_1, \ldots, X_n\} \in \mathbb{R}^n, \underline{Y} := \{Y_1, \ldots, Y_n\} \in \mathbb{R}^n, \lambda \in [0,1]$, we have $f(\lambda \underline{X} + (1-\lambda)\underline{Y}) \geq \min\{f(\underline{X}), f(\underline{Y})\}$. For any two random samples $\underline{X} = \{X_1, \ldots, X_n\}$ and $\underline{Y} = \{Y_1, \ldots, Y_n\}$ such that $X_i, Y_j \overset{d}{\approx} F(1,\alpha)$ for some $\alpha \in (0,1)$, $i, j = 1, \ldots, n$, a real-valued function f of real n-vectors ($n \geq 1$) satisfying, for any $\epsilon \in (0,1)$ and $\lambda \in [0,1]$,

$$\mathbb{P}\left(\frac{f(\lambda \underline{X} + (1-\lambda)\underline{Y})}{\min\{f(\underline{X}), f(\underline{Y})\}} \geq 1 - \epsilon\right) \to 1$$

as $n \to \infty$, is said to be F_α-asymptotically quasi-concave. A real-valued function f of a real random n-vector $\underline{X}$ as just defined that satisfies, for any $\epsilon \in (0,1)$ and $c > 0$,

$$\mathbb{P}\left(\left|\frac{f(c\underline{X})}{f(\underline{X})} - 1\right| < \epsilon\right) \to 1$$

as $n \to \infty$, is said to be F_α-asymptotically scale-invariant.

Proposition D.1. *Let $\underline{X}$ be a random sample from $F(1,\alpha)$ for some $\alpha \in (0,1)$. Let $\underline{X}$ have sample variance $\sigma_n^2(\underline{X}) := v_n$, sample mean $\overline{X} := M_1'$, and sample size n. Then, for the sample $\underline{X}$, as $n \to \infty$, the modified Sharpe ratio $\log(M_1' - r_f)/\log(v_n)$ for $r_f > 0$ is (i) F_α-asymptotically quasi-concave, (ii) F_α-asymptotically scale-invariant, (iii) asymptotically dependent only on the tail index α of the distribution, and (iv) asymptotically monotonic with respect to a shift by a constant ($\underline{Y} = \underline{X} + d$, $0 < d < \infty$).*

Proof. **(i) F_α-asymptotic quasi-concavity**: Let $\sigma_n^2(\underline{X})$ be the sample variance of a random sample $\underline{X}$ with sample mean $\overline{X}$ and sample size n. Consider another random sample $\underline{Y}$ with sample mean $\overline{Y}$, sample variance $\sigma_n^2(\underline{Y})$, and the same sample size n. For $\lambda \in [0,1]$,

$$\begin{aligned}&\sigma_n^2(\lambda \underline{X} + (1-\lambda)\underline{Y}) \\ &= \lambda^2 \sigma_n^2(\underline{X}) + 2\lambda(1-\lambda) Cov_n(\underline{X}, \underline{Y}) + (1-\lambda)^2 \sigma_n^2(\underline{Y}) \\ &\leq \lambda^2 \max\{\sigma_n^2(\underline{X}), \sigma_n^2(\underline{Y})\} + 2\lambda(1-\lambda) \max\{\sigma_n^2(\underline{X}), \sigma_n^2(\underline{Y})\} + (1-\lambda)^2 \max\{\sigma_n^2(\underline{X}), \sigma_n^2(\underline{Y})\} \\ &= \max\{\sigma_n^2(\underline{X}), \sigma_n^2(\underline{Y})\},\end{aligned}$$

where Cov_n is the sample covariance between $\underline{X}$ and $\underline{Y}$. The inequality above holds because $|Cov_n(\underline{X},\underline{Y})| \le \sqrt{\sigma_n^2(\underline{X})\sigma_n^2(\underline{Y})} \le \max\{\sigma_n^2(\underline{X}),\sigma_n^2(\underline{Y})\}$. For $\lambda \in [0,1]$, we also have

$$\frac{\log(\lambda\overline{X}+(1-\lambda)\overline{Y}-r_f)}{\log\sigma_n^2(\lambda\underline{X}+(1-\lambda)\underline{Y})} = \frac{\log\{\lambda(\overline{X}-r_f)+(1-\lambda)(\overline{Y}-r_f)\}}{\log\sigma_n^2(\lambda\underline{X}+(1-\lambda)\underline{Y})}.$$

If $\min\{\sigma_n^2(\underline{X}),\sigma_n^2(\underline{Y})\} > 1$ and $\min\{\overline{X},\overline{Y}\} > 1+r_f$, which hold asymptotically a.s., then because of $\sigma_n^2(\lambda\underline{X}+(1-\lambda)\underline{Y}) \le \max\{\sigma_n^2(\underline{X}),\sigma_n^2(\underline{Y})\}$, we have

$$\begin{aligned}\frac{\log\{\lambda(\overline{X}-r_f)+(1-\lambda)(\overline{Y}-r_f)\}}{\log\sigma_n^2(\lambda\underline{X}+(1-\lambda)\underline{Y})} &\ge \min\left\{\frac{\log\{\lambda(\overline{X}-r_f)+(1-\lambda)(\overline{Y}-r_f)\}}{\log\sigma_n^2(\underline{X})}, \frac{\log\{\lambda(\overline{X}-r_f)+(1-\lambda)(\overline{Y}-r_f)\}}{\log\sigma_n^2(\underline{Y})}\right\} \\ &\ge \min\left\{\frac{\log(\overline{X}-r_f)}{\log\sigma_n^2(\underline{X})}, \frac{\log(\overline{Y}-r_f)}{\log\sigma_n^2(\underline{Y})}\right\}.\end{aligned}$$

The last inequality holds because the log function is concave, so $\log\{\lambda(\overline{X}-r_f)+(1-\lambda)(\overline{Y}-r_f)\}$ is larger than or equal to $\log(\overline{X}-r_f)$ when $\lambda=1$ or is larger than or equal to $\log(\overline{Y}-r_f)$ when $\lambda=0$. Because $\underline{X}$ and $\underline{Y}$ are assumed to be positive and drawn from $F(1,\alpha)$ with an infinite expectation, as $n\to\infty$, $\mathbb{P}(\min\{\sigma_n^2(\underline{X}),\sigma_n^2(\underline{Y})\} > r_f+1, \min\{\overline{X},\overline{Y}\} > 1) \to 1$ because all $\sigma_n^2(\underline{X})$, $\sigma_n^2(\underline{Y})$, $\overline{X}$, and $\overline{Y}$ converge to infinity a.s.. Therefore, given $0<\epsilon<1$, as $n\to\infty$,

$$\mathbb{P}\left(\frac{\log(\lambda\overline{X}+(1-\lambda)\overline{Y}-r_f)}{\log\sigma_n^2(\lambda\underline{X}+(1-\lambda)\underline{Y})}\Big/\min\left\{\frac{\log(\overline{X}-r_f)}{\log\sigma_n^2(\underline{X})}, \frac{\log(\overline{Y}-r_f)}{\log\sigma_n^2(\underline{Y})}\right\} \ge 1-\epsilon\right) \to 1.$$

Hence this modified Sharpe ratio is F_α-asymptotically quasi-concave.

(ii) F_α-asymptotic scale-invariance: Since $\underline{X}$ is from $F(1,\alpha)$, as $n\to\infty$, $\overline{X} \overset{a.s.}{\to} \infty$; for any $c>0$, $c\overline{X} \overset{a.s.}{\to} \infty$; and $\sigma_n^2(\underline{X}) \overset{a.s.}{\to} \infty$. Then

$$\left|\frac{\log(\overline{X}-r_f)}{\log\sigma_n^2(\underline{X})}\Big/\frac{\log(c\overline{X}-r_f)}{\log\sigma_n^2(c\underline{X})}\right| = \left\{\left|\frac{\log(\overline{X}-r_f)}{\log(\overline{X})+\log c}\right|\cdot\left|\frac{\log(c\overline{X})}{\log(c\overline{X}-r_f)}\right|\right\}\Big/\left|\frac{\log\sigma_n^2(\underline{X})+\log c}{\log\sigma_n^2(\underline{X})}\right| \overset{p}{\to} 1.$$

Hence the modified Sharpe ratio is F_α-asymptotically scale-invariant.

(iii) Distribution-based: The modified Sharpe ratio converges to $(2-\alpha)/(1-\alpha)$, which depends only on the tail index α of $F(1,\alpha)$.

(iv) Monotonicity under a constant shift: Consider $\underline{Y} = \underline{X}+d$ with $0<d<\infty$. Then $\underline{X} \le \underline{Y}$ a.s. by definition. When $\min\{\overline{X}-r_f, \overline{Y}-r_f, \sigma_n^2(\underline{X}), \sigma_n^2(\underline{Y})\} > 1$, we have

$$\frac{\log(\overline{X}-r_f)}{\log\sigma_n^2(\underline{X})} \le \frac{\log(\overline{X}+d-r_f)}{\log\sigma_n^2(\underline{X}+d)} = \frac{\log(\overline{X}+d-r_f)}{\log\sigma_n^2(\underline{X})} = \frac{\log(\overline{Y}-r_f)}{\log\sigma_n^2(\underline{Y})}.$$

Since $\overline{X} \overset{a.s.}{\to} \infty$ and $\sigma_n^2(\underline{X}) \overset{a.s.}{\to} \infty$ as $n\to\infty$, and since r_f is finite, we have $\mathbb{P}(\min\{\overline{X}-r_f, \overline{Y}-r_f, \sigma_n^2(\underline{X}), \sigma_n^2(\underline{Y})\} > 1) \to 1$. Therefore, the modified Sharpe ratio is monotonic with a probability converging to 1 as $n\to\infty$. □

Proposition D.2. *For samples from $F(1,\alpha), \alpha \in (0,1)$, the modified Sortino ratios $\log(M_1'-r_f)/\log v_n^-$ and $\log(M_1'-r_f)/\log v_n^{-*}$ for $r_f>0$ are F_α-asymptotically quasi-concave and F_α-asymptotically scale-invariant, but do not depend on the distribution, asymptotically as $n\to\infty$. They are asymptotically monotonic with respect to a constant shift.*

Proof. **F_α-asymptotic quasi-concavity**: Let $v_n^-(\underline{X})$ denote the lower semivariance of a random sample $\underline{X}$ with sample mean $\overline{X}$. Consider another random sample $\underline{Y}$ with sample mean $\overline{Y}$ and lower semivariance $v_n^-(\underline{Y})$. According to Example 1 in Rockafellar et al. (8), the square root of lower semivariance has the following properties:

$$\sqrt{v_n^-(\underline{X}+\underline{Y})} \le \sqrt{v_n^-(\underline{X})} + \sqrt{v_n^-(\underline{Y})} \text{ and } \sqrt{v_n^-(c\underline{X})} = c\sqrt{v_n^-(\underline{X})},$$

for $c>0$. Therefore, whenever $\lambda\in[0,1]$, we have

$$\sqrt{v_n^-(\lambda\underline{X}+(1-\lambda)\underline{Y})} \le \sqrt{v_n^-(\lambda\underline{X})} + \sqrt{v_n^-\{(1-\lambda)\underline{Y}\}} = \lambda\sqrt{v_n^-(\underline{X})} + (1-\lambda)\sqrt{v_n^-(\underline{Y})} \le \max\left\{\sqrt{v_n^-(\underline{X})}, \sqrt{v_n^-(\underline{Y})}\right\}.$$

Following the lines of the proof in Proposition D.1 that the modified Sharpe ratio is F_α-asymptotically quasi-concave, if $\min\{\overline{X},\overline{Y}\} > r_f+1$ and $\min\{v_n^-(\underline{X}), v_n^-(\underline{Y})\} > 1$, one can see again that, for $0<\epsilon<1$, the probability that

$$\frac{\log(\lambda\overline{X}+(1-\lambda)\overline{Y}-r_f)}{\log v_n^-(\lambda\underline{X}+(1-\lambda)\underline{Y})} \ge (1-\epsilon)\min\left\{\frac{\log(\overline{X}-r_f)}{\log v_n^-(\underline{X})}, \frac{\log(\overline{Y}-r_f)}{\log v_n^-(\underline{Y})}\right\} \qquad \text{[S.42]}$$

converges to 1 as $n\to\infty$.

Because $\underline{X}$ and $\underline{Y}$ are positive and drawn from distributions with infinite means, it follows that $\mathbb{P}(\min\{v_n^-(\underline{X}), v_n^-(\underline{Y})\} > 1, \min\{\overline{X},\overline{Y}\} > 1+r_f) \to 1$ as $n\to\infty$ because $\overline{X}$ and $\overline{Y}$ converge to infinity a.s. and $v_n^-(\underline{X})$ and $v_n^-(\underline{Y})$ converge to infinity a.s. as $n\to\infty$ by Corollary 1. Therefore, the modified Sortino ratio $\log(M_1'-r_f)/\log v_n^-$ is F_α-asymptotically quasi-concave.

For the modified Sortino ratio $\log(M_1' - r_f)/\log v_n^{-*}$, we need some properties of the local lower semivariance. Let $N_n^-(\underline{X})$ denote the number of observations in $\underline{X}$ less than or equal to the sample mean $\overline{X}$. From the definition of the local lower semivariance, we have

$$v_n^{-*}(\lambda\underline{X} + (1-\lambda)\underline{Y}) = \frac{nv_n^-(\lambda\underline{X} + (1-\lambda)\underline{Y})}{N_n^-(\lambda\underline{X} + (1-\lambda)\underline{Y})}, \quad v_n^{-*}(\underline{X}) = \frac{nv_n^-(\underline{X})}{N_n^-(\underline{X})}, \text{ and } v_n^{-*}(\underline{Y}) = \frac{nv_n^-(\underline{Y})}{N_n^-(\underline{Y})}.$$

We aim to show that, for $\lambda \in [0,1]$, for $0 < \epsilon < 1$, as $n \to \infty$, the probability converges to 1 that

$$\frac{\log(\lambda\overline{X} + (1-\lambda)\overline{Y})}{\log v_n^{-*}(\lambda\underline{X} + (1-\lambda)\underline{Y})} \geq (1-\epsilon)\min\left\{\frac{\log(\overline{X} - r_f)}{\log v_n^{-*}(\underline{X})}, \frac{\log(\overline{Y} - r_f)}{\log v_n^{-*}(\underline{Y})}\right\}.$$

Now

$$\frac{\log(\lambda\overline{X} + (1-\lambda)\overline{Y})}{\log v_n^{-*}(\lambda\underline{X} + (1-\lambda)\underline{Y})} = \frac{\log(\lambda\overline{X} + (1-\lambda)\overline{Y})}{\log v_n^{-*}(\lambda\underline{X} + (1-\lambda)\underline{Y}) + \log\{n/N_n^-(\lambda\underline{X} + (1-\lambda)\underline{Y})\}}, \tag{S.43}$$

$$\min\left\{\frac{\log(\overline{X} - r_f)}{\log v_n^{-*}(\underline{X})}, \frac{\log(\overline{Y} - r_f)}{\log v_n^{-*}(\underline{Y})}\right\} = \min\left\{\frac{\log(\overline{X} - r_f)}{\log v_n^-(\underline{X}) + \log\{n/N_n^-(\underline{X})\}}, \frac{\log(\overline{Y} - r_f)}{\log v_n^-(\underline{Y}) + \log\{n/N_n^-(\underline{Y})\}}\right\}. \tag{S.44}$$

Therefore, if we can show that, as $n \to \infty$, $\log\{n/N_n^-(\lambda\underline{X} + (1-\lambda)\underline{Y})\} \xrightarrow{p} 0$ in Eq. (S.43), and both $\log\{n/N_n^-(\underline{X})\}$ and $\log\{n/N_n^-(\underline{Y})\}$ in Eq. (S.44) converge to 0 in probability, then the modified Sortino ratio $\log(M_1' - r_f)/\log v_n^{-*}$ is F_α-asymptotically quasi-concave. Since both samples $\underline{X}$ and $\underline{Y}$ are drawn from $F(1,\alpha), \alpha \in (0,1)$, by Lemma 1, both $\log\{n/N_n^-(\underline{X})\}$ and $\log\{n/N_n^-(\underline{Y})\}$ converge to 0 a.s. as $n \to \infty$. Therefore, $\log\{n/N_n^-(\lambda\underline{X} + (1-\lambda)\underline{Y})\}$ also converges to 0 a.s. as $n \to \infty$ from Lemma 1 because $\lambda\underline{X} + (1-\lambda)\underline{Y}$ has an infinite mean.

F_α-asymptotic scale-invariance: The proof of F_α-asymptotic quasi-concavity above gives $\sqrt{v_n^-(c\underline{X})} = c\sqrt{v_n^-(\underline{X})}$ and $\sqrt{v_n^{-*}(c\underline{X})} = c\sqrt{v_n^{-*}(\underline{X})}$. Then the proofs of the F_α-asymptotic scale-invariance for the modified ratios $\log(M_1' - r_f)/\log v_n^-$ and $\log(M_1' - r_f)/\log v_n^{-*}$ are the same as that for the F_α-asymptotic scale-invariance for the modified Sharpe ratio in Proposition D.1 if we replace σ_n^2 by v_n^- or v_n^{-*}, respectively.

Not distribution-based: Because both $\log(\overline{X} - r_f)/\log v_n^-(\underline{X})$ and $\log(\overline{X} - r_f)/\log v_n^{-*}(\underline{X})$ converge to 2 a.s. when $\underline{X}$ comes from $F(1,\alpha), \alpha \in (0,1)$, the modified ratios are not sensitive to α.

Monotonicity under a constant shift: Consider $\underline{Y} = \underline{X} + d$ with $0 < d < \infty$. Then $\underline{X} \leq \underline{Y}$ a.s. by definition. When $\min\{\overline{X} - r_f, \overline{Y} - r_f, v_n^-(\underline{X}), v_n^-(\underline{Y})\} > 1$, we have

$$\frac{\log(\overline{X} - r_f)}{\log v_n^-(\underline{X})} \leq \frac{\log(\overline{X} + d - r_f)}{\log v_n^-(\underline{X} + d)} = \frac{\log(\overline{X} + d - r_f)}{\log v_n^-(\underline{X})} = \frac{\log(\overline{Y} - r_f)}{\log v_n^-(\underline{Y})}. \tag{S.45}$$

Since $\overline{X} \xrightarrow{a.s.} \infty$ and $v_n^-(\underline{X}) \xrightarrow{a.s.} \infty$ (from Corollary 1) as $n \to \infty$, and since r_f is finite, we have $\mathbb{P}(\min\{\overline{X} - r_f, \overline{Y} - r_f, v_n^-(\underline{X}), v_n^-(\underline{Y})\} > 1) \to 1$ as $n \to \infty$. Also, because $v_n^{-*} \xrightarrow{a.s.} \infty$ (from Corollary 2) as $n \to \infty$, we may replace v_n^- in Eq. (S.45) by v_n^{-*}. Therefore, the probability that the modified Sortino ratios are monotonic converges to 1 as $n \to \infty$. ☐

Proposition D.3. *For samples $\underline{X}$ of size n from $F(1,\alpha), \alpha \in (0,1)$, as $n \to \infty$,*
a) the ratio $\log(M_1' - r_f)/\log v_n^+$ for $r_f > 0$ (i) is F_α-asymptotically quasi-concave; (ii) is F_α-asymptotically scale-invariant; (iii) depends only on the tail index α of the distribution; and (iv) is asymptotically monotonic with respect to a constant shift; and
b) the ratio $\log(M_1' - r_f)/\log v_n^{+*}$ for $r_f > 0$ (i) satisfies, for $0 < \epsilon < 1$, $0 \leq \lambda \leq 1$, $\underline{Y} = \underline{X} + d$ for $0 < d < \infty$,

$$\mathbb{P}\left(\frac{\log(\lambda\overline{X} + (1-\lambda)\overline{Y})}{\log v_n^{+*}(\lambda\underline{X} + (1-\lambda)\underline{Y})} \geq (1-\epsilon)\min\left\{\frac{\log(\overline{X} - r_f)}{\log v_n^{+*}(\underline{X})}, \frac{\log(\overline{Y} - r_f)}{\log v_n^{+*}(\underline{Y})}\right\}\right) \to 1; \tag{S.46}$$

(ii) is F_α-asymptotically scale-invariant; (iii) depends on α; and (iv) is asymptotically monotonic with respect to a constant shift.

Proof. **a) (i, ii) F_α-asymptotic quasi-concavity and scale-invariance**: Let $v_n^+(\underline{X})$ denote the upper semivariance of the random sample $\underline{X}$. From Example 1 in Rockafellar et al. (8), parallel to the lower semivariance, we have

$$\sqrt{v_n^+(\underline{X} + \underline{Y})} \leq \sqrt{v_n^+(\underline{X})} + \sqrt{v_n^+(\underline{Y})} \text{ and } \sqrt{v_n^+(c\underline{X})} = c\sqrt{v_n^+(\underline{X})}.$$

Therefore, the proofs of quasi-concavity and scale-invariance for $\log(M_1' - r_f)/\log v_n^+$ for $r_f > 0$ are the same as those for the modified Sortino ratio $\log(M_1' - r_f)/\log v_n^-$ in Proposition D.2. Simply replace v_n^- by v_n^+.

a) (iii) Distribution-based: The ratio $\log(M_1' - r_f)/\log v_n^+$ converges to $(2-\alpha)/(1-\alpha)$, which depends only on α.

a) (iv) Monotonicity under a constant shift: Consider $\underline{Y} = \underline{X} + d$ with $0 < d < \infty$. Then $\underline{X} \leq \underline{Y}$ a.s. by definition. When $\min\{\overline{X} - r_f, \overline{Y} - r_f, v_n^+(\underline{X}), v_n^+(\underline{Y})\} > 1$, we have

$$\frac{\log(\overline{X} - r_f)}{\log v_n^+(\underline{X})} \leq \frac{\log(\overline{X} + d - r_f)}{\log v_n^+(\underline{X} + d)} = \frac{\log(\overline{X} + d - r_f)}{\log v_n^+(\underline{X})} = \frac{\log(\overline{Y} - r_f)}{\log v_n^+(\underline{Y})}. \tag{S.47}$$

Since $\overline{X} \xrightarrow{a.s.} \infty$ and $v_n^+(\underline{X}) \xrightarrow{p} \infty$ (from Theorem 2) as $n \to \infty$, and since r_f is finite, we have $\mathbb{P}(\min\{\overline{X} - r_f, \overline{Y} - r_f, v_n^-(\underline{X}), v_n^-(\underline{Y})\} > 1) \to 1$. Therefore, the ratio $\log(M_1' - r_f)/\log v_n^+$ for $r_f > 0$ is asymptotically monotonic with respect to a constant shift.

b)(i) Quasi-concavity under a constant shift: Consider $\underline{Y} = \underline{X} + d$ with $0 < d < \infty$. Then $Y_i = X_i + d$ and

$$\sum_{i=1}^{n} I(X_i > \overline{X}) = \sum_{i=1}^{n} I(Y_i - d > \overline{Y} - d) = \sum_{i=1}^{n} I(Y_i > \overline{Y}),$$
$$\sum_{i=1}^{n} I(\lambda X_i + (1-\lambda)Y_i > \lambda\overline{X} + (1-\lambda)\overline{Y}) = \sum_{i=1}^{n} I(\lambda(Y_i - d) + (1-\lambda)Y_i > \lambda(\overline{Y} - d) + (1-\lambda)\overline{Y}) = \sum_{i=1}^{n} I(Y_i > \overline{Y}).$$

Hence $N_n^+(\underline{X} + \underline{Y}) = N_n^+(\underline{X}) = N_n^+(\underline{Y})$.

Because $c\sqrt{v_n^+(\underline{X})} = \sqrt{v_n^+(c\underline{X})}$ for $c > 0$ and $\sqrt{v_n^+(\underline{X} + \underline{Y})} \leq \sqrt{v_n^+(\underline{X})} + \sqrt{v_n^+(\underline{Y})}$ from the definition $v_n^{+*}(\underline{X}) := nv_n^+(\underline{X})/N_n^+(\underline{X})$, we have

$$\frac{1}{2}\sqrt{v_n^{+*}\left(\underline{X} + \underline{Y}\right)} = \sqrt{v_n^{+*}\left(\frac{1}{2}\underline{X} + \frac{1}{2}\underline{Y}\right)} = \sqrt{\frac{nv_n^+\left(\frac{1}{2}\underline{X} + \frac{1}{2}\underline{Y}\right)}{N_n^+\left(\frac{1}{2}\underline{X} + \frac{1}{2}\underline{Y}\right)}} \leq \sqrt{\frac{n}{N_n^+\left(\frac{1}{2}\underline{X} + \frac{1}{2}\underline{Y}\right)}}\left(\sqrt{v_n^+\left(\frac{1}{2}\underline{X}\right)} + \sqrt{v_n^+\left(\frac{1}{2}\underline{Y}\right)}\right).$$

One can check that

$$N_n^+\left(\frac{1}{2}\underline{X} + \frac{1}{2}\underline{Y}\right) = N_n^+\left(\frac{1}{2}(\underline{X} + \underline{Y})\right) = N_n^+\left(\underline{X} + \underline{Y}\right),$$
$$N_n^+\left(\frac{1}{2}\underline{X}\right) = N_n^+(\underline{X}),$$
$$N_n^+\left(\frac{1}{2}\underline{Y}\right) = N_n^+(\underline{Y}).$$

Then

$$\sqrt{\frac{n}{N_n^+\left(\frac{1}{2}\underline{X} + \frac{1}{2}\underline{Y}\right)}}\left(\sqrt{v_n^+\left(\frac{1}{2}\underline{X}\right)} + \sqrt{v_n^+\left(\frac{1}{2}\underline{Y}\right)}\right) = \frac{1}{2}\sqrt{\frac{n}{N_n^+\left(\underline{X} + \underline{Y}\right)}}\left(\sqrt{v_n^+\left(\underline{X}\right)} + \sqrt{v_n^+\left(\underline{Y}\right)}\right)$$
$$= \frac{1}{2}\sqrt{\frac{n}{N_n^+\left(\underline{X}\right)}}\left(\sqrt{v_n^+\left(\underline{X}\right)}\right) + \frac{1}{2}\sqrt{\frac{n}{N_n^+\left(\underline{Y}\right)}}\left(\sqrt{v_n^+\left(\underline{Y}\right)}\right) = \frac{1}{2}\left(\sqrt{v_n^{+*}\left(\underline{X}\right)}\right) + \frac{1}{2}\left(\sqrt{v_n^{+*}\left(\underline{Y}\right)}\right)$$

and therefore

$$\sqrt{v_n^{+*}\left(\underline{X} + \underline{Y}\right)} \leq \left(\sqrt{v_n^{+*}\left(\underline{X}\right)}\right) + \left(\sqrt{v_n^{+*}\left(\underline{Y}\right)}\right).$$

As in the proof of Proposition D.3 a) (i), using the fact that $\sqrt{v_n^{+*}(c\underline{X})} = c\sqrt{v_n^{+*}(\underline{X})}$, the ratio $\log(M_1' - r_f)/\log v_n^{+*}$ satisfies Eq. (S.46) as $n \to \infty$.

b)(ii) F_α-asymptotic scale-invariance: Because $\sqrt{v_n^{+*}(c\underline{X})} = c\sqrt{v_n^{+*}(\underline{X})}$, the proof of F_α-asymptotic scale-invariance for $\log(M_1' - r_f)/\log v_n^{+*}$ is the same as that for the modified Sortino ratio in Proposition D.2, after replacing v_n^- by v_n^{+*}.

b)(iii) Distribution-based: As $n \to \infty$, both $\log(M_1' - r_f)/\log v_n^+$ and $\log(M_1' - r_f)/\log v_n^{+*}$ converge in probability to $(2-\alpha)/(1-\alpha)$, which depends on the tail index α of the distribution.

b)(iv) Monotonicity under a constant shift: Theorem 2 gives that $v_n^{+*} \xrightarrow{p} \infty$ as $n \to \infty$. Replacing v_n^+ in Eq. (S.47) by v_n^{+*} gives asymptotic monotonicity for large n. ☐

Proposition D.4. *For a random sample of size n from $F(1, \alpha), \alpha \in (0, 1)$, as $n \to \infty$, the modified Farinelli-Tibiletti ratio $\Phi_{\mathrm{FTlog}}(p, q) := p\log M_q^-/(q\log M_p^+)$ for $p \geq 1, q \geq 1$ (i) is F_α-asymptotically quasi-concave; (ii) is F_α-asymptotically scale-invariant; and (iii) depends on the tail index α of the distribution; but (iv) is not monotonic with respect to a constant shift.*

Proof. **(i) F_α-asymptotic quasi-concavity**: We shall show that, for any $1 \leq p < \infty$ and for any $0 \leq \lambda \leq 1$,

$$\{M_p^+(\lambda\underline{X} + (1-\lambda)\underline{Y})\}^{1/p} \leq \max\{\{M_p^+(\underline{X})\}^{1/p}, \{M_p^+(\underline{Y})\}^{1/p}\},$$

where $\underline{X} = \{X_1, \ldots, X_n\}$ and $\underline{Y} = \{Y_1, \ldots, Y_n\}$ as before. To see this, we write the left side as

$$
\begin{aligned}
&\{M_p^+(\lambda \underline{X} + (1-\lambda)\underline{Y})\}^{1/p} \\
&\quad = \left(\frac{1}{n} \sum_{i=1}^n \Big[\{(\lambda X_i + (1-\lambda)Y_i - (\lambda \overline{X} + (1-\lambda)\overline{Y})\} I(\lambda X_i + (1-\lambda)Y_i - (\lambda \overline{X} + (1-\lambda)\overline{Y}) > 0) \Big]^p \right)^{1/p} \\
&\quad = \left(\frac{1}{n} \sum_{i=1}^n \Big[\{\lambda(X_i - \overline{X}) + (1-\lambda)(Y_i - \overline{Y})\} I(\lambda(X_i - \overline{X}) + (1-\lambda)(Y_i - \overline{Y}) > 0) \Big]^p \right)^{1/p} \\
&\quad = \left(\frac{1}{n} \sum_{i=1}^n \Big\{ (\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big\}^p \right)^{1/p},
\end{aligned}
$$

where $U_i := X_i - \overline{X}$ and $V_i := Y_i - \overline{Y}$, and therefore $\sum_{i=1}^n U_i = \sum_{i=1}^n V_i = 0$ by definition. For any $p \geq 1$, we have

$$
\begin{aligned}
0 \leq & \sum_{i=1}^n \Big\{ (\lambda U_i + (1-\lambda)V_i) \cdot I(\lambda U_i + (1-\lambda)V_i > 0) \Big\}^p \\
= & \sum_{i=1}^n \{\lambda U_i + (1-\lambda)V_i\} \cdot I(\lambda U_i + (1-\lambda)V_i > 0) \cdot \Big\{ (\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big\}^{p-1} \\
= & \sum_{i=1}^n \Big\{ [\lambda U_i \{I(U_i > 0) + I(U_i \leq 0)\} + (1-\lambda)V_i \{I(V_i > 0) + I(V_i \leq 0)\}] \cdot I(\lambda U_i + (1-\lambda)V_i > 0) \Big\} \\
& \qquad \cdot \Big\{ (\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big\}^{p-1} \\
= & \sum_{i=1}^n \Big[\{\lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0)\} I(\lambda U_i + (1-\lambda)V_i > 0) \Big] \cdot \Big[(\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big]^{p-1} \\
& + \sum_{i=1}^n \Big[\{\lambda U_i I(U_i \leq 0) + (1-\lambda)V_i I(V_i \leq 0)\} I(\lambda U_i + (1-\lambda)V_i > 0) \Big] \cdot \Big[(\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big]^{p-1} \\
\leq & \sum_{i=1}^n \Big[\{\lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0)\} I(\lambda U_i + (1-\lambda)V_i > 0) \Big] \cdot \Big[\{(\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big]^{p-1},
\end{aligned}
$$

where the inequality holds since the obvious inequalities $U_i I(U_i \leq 0) \leq 0$ and $V_i I(V_i \leq 0) \leq 0$ imply

$$
\sum_{i=1}^n \Big[\{\lambda U_i I(U_i \leq 0) + (1-\lambda)V_i I(V_i \leq 0)\} I(\lambda U_i + (1-\lambda)V_i > 0) \Big] \cdot \Big[(\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \Big]^{p-1} \leq 0.
$$

For $i = 1, 2, \ldots, n$, and any $0 \leq \lambda \leq 1$, $\{\lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0)\} \geq 0$ and $I(\lambda U_i + (1-\lambda)V_i > 0) \leq 1$. Hence

$$
\{\lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0)\} I(\lambda U_i + (1-\lambda)V_i > 0) \leq \lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0). \qquad \text{[S.48]}
$$

We next show that, for $i = 1, 2, \ldots, n$, and any $0 \leq \lambda \leq 1$,

$$
(\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \leq \lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0). \qquad \text{[S.49]}
$$

When $U_i > 0$ and $V_i > 0$, Eq. (S.49) is trivial. When $U_i \leq 0$ and $V_i \leq 0$, Eq. (S.49) also holds since $0 \leq 0$. When $U_i \leq 0$ and $V_i > 0$, then $\lambda U_i + (1-\lambda)V_i > 0$ implies $V_i > 0$ and Eq. (S.49) holds by noting that

$$
\begin{aligned}
0 \leq (\lambda U_i + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) &\leq (\lambda \cdot 0 + (1-\lambda)V_i) I(\lambda U_i + (1-\lambda)V_i > 0) \\
&= \{\lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0)\} I(\lambda U_i + (1-\lambda)V_i > 0) \\
&\leq \lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0),
\end{aligned}
$$

where the last inequality holds since $\{\lambda U_i I(U_i > 0) + (1-\lambda)V_i I(V_i > 0)\} \geq 0$ and $I(\lambda U_i + (1-\lambda)V_i > 0) \leq 1$. Finally, by symmetry, Eq. (S.49) also holds when their roles are interchanged, i.e. $U_i > 0$ and $V_i \leq 0$. Therefore, Eq. (S.49) holds for all possible cases.

Using Eq. (S.48) and Eq. (S.49), we eventually have the following inequality:

$$\sum_{i=1}^{n}\Big[\{\lambda U_i I(U_i>0)+(1-\lambda)V_i I(V_i>0)\}I(\lambda U_i+(1-\lambda)V_i>0)\Big]\cdot\Big[(\lambda U_i+(1-\lambda)V_i)I(\lambda U_i+(1-\lambda)V_i>0)\Big]^{p-1}$$
$$\le \sum_{i=1}^{n}\Big\{\lambda U_i I(U_i>0)+(1-\lambda)V_i I(V_i>0)\Big\}\cdot\Big\{\lambda U_i I(U_i>0)+(1-\lambda)V_i I(V_i>0)\Big\}^{p-1}$$
$$= \sum_{i=1}^{n}\Big\{\lambda U_i I(U_i>0)+(1-\lambda)V_i I(V_i>0)\Big\}^{p}.$$

Applying the Minkowski inequality and using $\lambda U_i I(U_i>0)\ge 0$ and $(1-\lambda)V_i I(V_i>0)\ge 0$ for $0\le\lambda\le 1$ gives

$$\left(\sum_{i=1}^{n}\Big\{\lambda U_i I(U_i>0)+(1-\lambda)V_i I(V_i>0)\Big\}^{p}\right)^{1/p} \le \left(\sum_{i=1}^{n}\{\lambda U_i I(U_i>0)\}^{p}\right)^{1/p}+\left(\sum_{i=1}^{n}\{(1-\lambda)V_i I(V_i>0)\}^{p}\right)^{1/p}$$
$$=\lambda\left(\sum_{i=1}^{n}\{U_i I(U_i>0)\}^{p}\right)^{1/p}+(1-\lambda)\left(\sum_{i=1}^{n}\{V_i I(V_i>0)\}^{p}\right)^{1/p}.$$

By construction, $\{U_i I(U_i>0)\}^p=\{(X_i-\overline{X})I(X_i-\overline{X}>0)\}^p$ and $\{V_i I(V_i>0)\}^p=\{(Y_i-\overline{Y})I(Y_i-\overline{Y}>0)\}^p$. Hence

$$\{M_p^+(\lambda\underline{X}+(1-\lambda)\underline{Y})\}^{1/p}\le\lambda\{M_p^+(\underline{X})\}^{1/p}+(1-\lambda)\{M_p^+(\underline{Y})\}^{1/p},$$

which easily yields

$$\{M_p^+(\lambda\underline{X}+(1-\lambda)\underline{Y})\}^{1/p}\le\max\{\{M_p^+(\underline{X})\}^{1/p},\{M_p^+(\underline{Y})\}^{1/p}\}.$$

It is straightforward that $\{M_p^+(c\underline{X})\}^{1/p}=c\{M_p^+(\underline{X})\}^{1/p}$ for $c>0$. Therefore, by replacing v_n^- by M_p^+, we can use the arguments for the modified Sortino ratio in Proposition D.2 to show that $\log M_1'/\log M_p^+$ is F_α-asymptotically quasi-concave. The modified Farinelli-Tibiletti ratio can be expressed as

$$\frac{p\log M_q^-}{q\log M_p^+}=\frac{p\log M_q^-}{q\log M_1'}\frac{\log M_1'}{\log M_p^+}.$$

From Corollary 6, $\frac{p\log M_q^-}{q\log M_1'}\xrightarrow{a.s.}p$ as $n\to\infty$. Therefore, the modified Farinelli-Tibiletti ratio is F_α-asymptotically quasi-concave because $\log M_1'/\log M_p^+$ is F_α-asymptotically quasi-concave while the constant p does not affect the quasi-concavity.

(ii)F_α-asymptotic scale-invariance: $\underline{X}$ is drawn from $F(1,\alpha),\alpha\in(0,1)$. So $\log M_q^-(\underline{X})\xrightarrow{p}\infty$ as $n\to\infty$ because $\log M_q^-(\underline{X})/\log\overline{X}\xrightarrow{p}q>0$ as $n\to\infty$ from Corollary 6.3, and $\overline{X}\xrightarrow{a.s.}\infty$ as $n\to\infty$. From Theorem 8.3, $\log M_p^+(\underline{X})\xrightarrow{p}\infty$ as $n\to\infty$ because $\log M_p^+(\underline{X})/\log\overline{X}\xrightarrow{p}\alpha(h,1)$ as $n\to\infty$ and $\overline{X}\xrightarrow{a.s.}\infty$ as $n\to\infty$. Hence, for any $c>0$,

$$\left|\frac{p\log M_q^-(\underline{X})}{q\log M_p^+(\underline{X})}\Big/\frac{p\log M_q^-(c\underline{X})}{q\log M_p^+(c\underline{X})}\right|=\left|\frac{\log M_p^+(\underline{X})}{\log M_p^+(\underline{X})+\log c}\right|\Big/\left|\frac{\log M_q^+(\underline{X})+\log c}{\log M_q^-(\underline{X})}\right|,$$

which converges to 1 in probability as $n\to\infty$ because both $\log M_p^+$ and $\log M_q^-$ diverge to infinity in probability as $n\to\infty$. This proves F_α-asymptotic scale-invariance.

(iii) Distribution-based: The modified ratio $\frac{\log M_p^+(\underline{X})}{\log M_q^-(\underline{X})}\xrightarrow{p}(p-\alpha)/\{q(1-\alpha)\}$ as $n\to\infty$, which depends on α.

(iv) Invariance and monotonicity under a constant shift and a positive scaling: Let $M_p^+(\underline{X})$ and $M_p^-(\underline{X})$ denote the upper and the lower pth central partial moments of sample $\underline{X}$. Define $\underline{Y}=c\underline{X}+d$, where $c>0$ and $d>0$. Then

$$\frac{p\log M_q^-(\underline{Y})}{q\log M_p^+(\underline{Y})}=\frac{p\log M_q^-(c\underline{X}+d)}{q\log M_p^+(c\underline{X}+d)}=\frac{p\log M_q^-(c\underline{X})}{q\log M_p^+(c\underline{X})}=\frac{p\log M_q^-(\underline{X})+p\log c}{q\log M_p^+(\underline{X})+q\log c}=\frac{\frac{p\log M_q^-(\underline{X})}{q\log M_p^+(\underline{X})}+\frac{p\log c}{q\log M_p^+(\underline{X})}}{1+\frac{\log c}{\log M_p^+(\underline{X})}}. \qquad \text{[S.50]}$$

When $\underline{X}$ is drawn from $F(1,\alpha)$, according to Corollary 8, $\frac{p\log M_q^-(\underline{X})}{q\log M_p^+(\underline{X})}\xrightarrow{p}\{p(1-\alpha)\}/(p-\alpha)$ as $n\to\infty$. Thus, from Eq. (S.50), $\frac{p\log M_q^-(\underline{Y})}{q\log M_p^+(\underline{Y})}\xrightarrow{p}\{p(1-\alpha)\}/(p-\alpha)$, the same limit, because $\log c$ is a constant and $\log M_p^+$ converges to infinity a.s. (from Theorem 8 and the fact that $\log M_1'\xrightarrow{a.s.}\infty$) as $n\to\infty$.

When $c=1$ and $d>0$, then $\underline{Y}>\underline{X}$ element-wise. From Eq. (S.50), $\frac{p\log M_q^-(\underline{X})}{q\log M_p^+(\underline{X})}=\frac{p\log M_q^-(\underline{Y})}{q\log M_p^+(\underline{Y})}$. □

E. More simulation results for the tail-index estimators

Here we provide additional comparisons of the tail-index estimators. The main text defines

$$B_1 := \frac{2 - R_1}{1 - R_1}, \qquad B_2 := \frac{2 - R_2}{1 - R_2}, \qquad B_3 := \frac{R_3 - \sqrt{R_3^2 - 4(R_3 - 2)}}{2}$$

with $R_1 := \log v_n / \log M_1'$, $R_2 := \log v_n^+ / \log M_1'$ and $R_3 := \log v_n^{+*} / \log M_1'$. The estimator HI.N considers the k largest order statistics in the Hill estimator, where $k = N_n^+ + 1$, which is the number of observations larger than the sample mean plus one. The estimator HI.M replaces the smallest $(n - k)$ order statistics in the Hill estimator by the sample mean M_1'. The estimator HI.Opt is the Hill estimator with $k = n^{2/3}$. The estimator MHB3 is the minimum of B_3 and HI.Opt.

Tables 1, 3, and 5 are the biases and Tables 2, 4, 6 are the MSEs for the estimators when the underlying distribution is $F(c, \alpha)$ with $c = 0.5, 1, 2$, $\alpha = 0.1, 0.2, \ldots, 0.9$ with respective sample sizes $n = 10^2, 10^3, 10^4, 10^5$. HI.N, HI.M, and HI.Opt are independent of the choice of c. One sees that for larger sample sizes, the choice of c has negligible influence on the asymptotic behavior of the estimators.

Table 7 and Table 8 provide the bias and MSE of R_L which converges to 2 a.s. as $n \to \infty$, by Corollary 1. R_L converges much more slowly when α is near 1 and much faster when α is close to 0.

Table 1. Bias $(\times 10^3)$ (average of [estimated α minus true α]) for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with 10^4 Monte Carlo independent samples from $F(1, \alpha)$, for sample sizes $n = 10^2, 10^3, 10^4, 10^5$.

α	n	B1	B2	B3	HI.N	HI.M	HI.Opt	MHB3
0.1	10^2	-8.19	-5.77	-4.85	28.71	355.02	-1.67	-10.00
0.2	10^2	-19.08	-9.96	-5.72	21.34	286.41	-1.79	-16.44
0.3	10^2	-32.94	-13.59	-3.65	5.62	227.71	0.08	-20.40
0.4	10^2	-48.74	-17.40	1.22	-11.37	179.72	7.34	-19.83
0.5	10^2	-67.66	-23.07	5.33	-22.42	143.29	19.30	-19.00
0.6	10^2	-87.89	-30.11	8.06	-20.93	127.39	48.98	-12.55
0.7	10^2	-109.73	-36.21	9.34	12.50	145.87	110.65	-3.19
0.8	10^2	-143.81	-44.43	4.65	115.54	241.48	248.13	1.16
0.9	10^2	-147.70	-56.17	-6.63	479.82	607.47	661.17	-6.74
0.1	10^3	-6.26	-4.48	-3.65	18.72	208.08	-1.51	-7.52
0.2	10^3	-14.84	-8.19	-4.81	1.12	139.58	-2.87	-12.78
0.3	10^3	-24.26	-9.75	-2.44	-15.69	86.75	-2.96	-15.36
0.4	10^3	-34.98	-11.78	1.36	-31.90	46.89	-0.84	-16.02
0.5	10^3	-47.37	-13.71	5.84	-40.13	20.39	3.55	-15.44
0.6	10^3	-57.48	-14.36	11.59	-35.06	11.99	16.86	-9.48
0.7	10^3	-67.08	-16.48	14.25	-15.90	20.19	46.62	1.32
0.8	10^3	-75.76	-20.44	11.29	44.06	71.90	109.68	7.95
0.9	10^3	-82.40	-23.57	2.94	279.43	302.82	313.66	2.91
0.1	10^4	-5.24	-3.87	-3.00	10.25	135.16	-0.92	-5.82
0.2	10^4	-11.96	-6.88	-3.79	-9.31	73.52	-1.73	-9.65
0.3	10^4	-19.43	-8.55	-2.38	-25.60	30.89	-2.05	-12.03
0.4	10^4	-27.72	-9.75	0.63	-32.82	4.87	-1.54	-13.44
0.5	10^4	-35.03	-8.91	5.96	-29.40	-5.30	1.42	-12.56
0.6	10^4	-43.76	-10.44	9.41	-24.21	-8.21	6.67	-10.26
0.7	10^4	-50.27	-11.28	12.19	-10.06	0.13	19.37	-3.26
0.8	10^4	-53.49	-12.55	11.48	31.58	37.82	51.80	7.30
0.9	10^4	-50.31	-13.69	5.46	204.26	208.27	153.44	5.45
0.1	10^5	-4.25	-3.16	-2.32	4.62	93.62	-0.49	-4.43
0.2	10^5	-9.90	-5.89	-3.07	-14.72	38.99	-0.93	-7.40
0.3	10^5	-16.11	-7.62	-2.05	-25.85	6.25	-1.05	-9.19
0.4	10^5	-22.43	-8.15	0.40	-25.06	-6.58	-0.85	-10.65
0.5	10^5	-28.54	-7.56	4.50	-18.40	-8.70	0.33	-10.79
0.6	10^5	-34.65	-7.82	8.08	-12.02	-6.73	2.83	-9.99
0.7	10^5	-40.22	-8.75	10.27	-3.33	-0.69	8.95	-6.54
0.8	10^5	-41.39	-8.03	11.38	23.97	25.24	24.25	3.14
0.9	10^5	-36.82	-8.54	6.66	155.70	156.34	73.70	6.65

F. Effects of sample size on convergence

In this section, we study the convergences of Theorems 3, 6, and 9 in distributions with sample sizes $n = 10^2, 10^3, 10^4, 10^5$ and $\alpha = \{0.1, \ldots, 0.9\}$. We also examine the convergence of the Sortino ratio $(M_1' - r_f)/v_n^-$ discussed after Theorem 1 with sample size $n = 10^8$.

Table 2. MSE $(\times 10^3)$ (mean squared [estimated α minus true α]) for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with 10^4 Monte Carlo independent samples from $F(1, \alpha)$, for sample sizes $n = 10^2, 10^3, 10^4, 10^5$.

α	n	B1	B2	B3	HI	HI.M	HI.Opt	MHB3
0.1	10^2	0.44	0.48	0.43	6.69	129.57	0.47	0.46
0.2	10^2	1.62	1.85	1.59	11.42	89.75	1.93	1.71
0.3	10^2	3.43	3.90	3.31	15.82	64.35	4.61	3.68
0.4	10^2	5.70	6.03	5.22	21.89	50.79	8.50	5.84
0.5	10^2	8.66	8.06	7.02	29.28	46.20	14.68	8.14
0.6	10^2	12.25	9.53	8.01	39.71	52.45	25.28	9.58
0.7	10^2	16.68	10.13	7.89	57.73	75.42	50.26	9.31
0.8	10^2	29.07	9.83	6.43	116.70	160.64	131.64	7.08
0.9	10^2	53.46	8.62	3.95	556.50	708.05	646.60	3.98
0.1	10^3	0.23	0.24	0.20	4.11	45.76	0.10	0.19
0.2	10^3	0.85	0.93	0.67	6.36	24.47	0.41	0.66
0.3	10^3	1.75	1.96	1.31	9.85	15.10	0.88	1.26
0.4	10^3	2.92	3.11	2.11	13.37	12.23	1.59	1.97
0.5	10^3	4.44	4.39	3.16	16.97	12.84	2.60	2.91
0.6	10^3	5.73	5.07	3.87	18.33	14.29	4.32	3.68
0.7	10^3	6.94	5.30	4.08	19.80	17.00	8.28	4.07
0.8	10^3	7.91	4.86	3.38	28.47	28.39	21.43	3.49
0.9	10^3	8.54	3.48	1.91	144.27	153.30	119.16	1.91
0.1	10^4	0.14	0.15	0.11	2.61	20.05	0.02	0.10
0.2	10^4	0.53	0.58	0.35	4.73	9.13	0.09	0.30
0.3	10^4	1.13	1.23	0.71	7.33	6.31	0.19	0.57
0.4	10^4	1.86	1.96	1.16	9.15	6.39	0.34	0.85
0.5	10^4	2.60	2.66	1.76	9.12	6.76	0.54	1.15
0.6	10^4	3.47	3.20	2.31	7.93	6.13	0.85	1.53
0.7	10^4	4.15	3.38	2.60	6.46	5.30	1.52	1.94
0.8	10^4	4.32	3.05	2.28	6.97	6.67	4.35	2.13
0.9	10^4	3.59	2.10	1.33	57.51	58.26	26.36	1.33
0.1	10^5	0.09	0.10	0.07	1.91	10.16	0.01	0.06
0.2	10^5	0.37	0.39	0.22	3.88	4.49	0.02	0.18
0.3	10^5	0.75	0.80	0.40	5.51	4.00	0.04	0.29
0.4	10^5	1.25	1.31	0.71	5.66	4.03	0.07	0.46
0.5	10^5	1.75	1.77	1.11	4.38	3.39	0.12	0.61
0.6	10^5	2.29	2.15	1.54	2.86	2.30	0.18	0.81
0.7	10^5	2.81	2.40	1.86	1.88	1.63	0.32	1.06
0.8	10^5	2.77	2.11	1.69	1.85	1.78	0.92	1.27
0.9	10^5	2.09	1.36	0.97	28.91	28.99	5.90	0.97

From Theorems 3 and 6, where $\alpha(h_1, h_2) = (h_1 - \alpha)/(h_2 - \alpha)$ for $h_1 > \alpha$ and $h_2 > \alpha$, as $n \to \infty$,

$$\frac{M'_h}{(M'_1)^{\alpha(h,1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}}, \qquad [S.51]$$

$$\frac{M'_{h_2}}{(M'_{h_1})^{\alpha(h_2,h_1)}} \xrightarrow{d} \{\Gamma(1-\alpha)\}^{\frac{h_2-h_1}{h_1-\alpha}} \frac{U_{h_2}}{(U_{h_1})^{\alpha(h_2,h_1)}}. \qquad [S.52]$$

From Theorem 9, as $n \to \infty$,

$$\frac{N_n^+}{n^\alpha} \xrightarrow{d} \frac{U^{-\alpha}}{\Gamma(1-\alpha)}. \qquad [S.53]$$

To evaluate the convergence rate in Theorems 3 and 6, we generate random samples $\{X_1, \ldots, X_n\}$ from $F(1, \alpha)$ and calculate $\frac{M'_h}{(M'_1)^{\alpha(h,1)}}$ in Eq. (S.51) and $\frac{M'_{h_2}}{(M'_{h_1})^{\alpha(h_2,h_1)}}$ in Eq. (S.52). We repeat the process independently 1000 times to estimate their marginal distributions. To simulate the ratios $\{\Gamma(1-\alpha)\}^{\frac{h-1}{1-\alpha}} \frac{U_h}{V^{\alpha(h,1)}}$ and $\{\Gamma(1-\alpha)\}^{\frac{h_2-h_1}{h_1-\alpha}} \frac{U_{h_2}}{(U_{h_1})^{\alpha(h_2,h_1)}}$, we independently generate 1000 random vectors (U_h, V) and (U_{h_1}, U_{h_2}) satisfying the moment generating functions defined in Theorems 3 and 6, respectively. To test the null hypotheses that the left sides of Eq. (S.51) and Eq. (S.52) have the same distribution as the respective right sides of Eq. (S.51) and Eq. (S.52), we perform the two-sample Kolmogorov-Smirnov test 100 times independently and estimate the probabilities of rejecting the null hypotheses. The margin of error, assuming a 99% confidence level, is approximately 0.05. The probabilities of rejecting the null hypothesis are provided in Table 9 for $(h_1, h_2) = (1, 2), (1, 3), (1, 4), (2, 3), (2, 4), (3, 4)$, $n = 10^2, 10^3, 10^4, 10^5$ and $\alpha = \{0.1, 0.2, \ldots, 0, 9\}$.

To generate random vectors (U_h, V) and (U_{h_1}, U_{h_2}) that have the moment generating functions defined in Theorems 3 and 6, we use the approximations by sequences of independent, identically distributed random variables from an exponential

Table 3. Bias $(\times 10^3)$ for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with 10^4 Monte Carlo independent samples from $F(0.5, \alpha)$, for sample sizes $n = 10^2, 10^3, 10^4, 10^5$.

α	n	B1	B2	B3	HI	HI.M	HI.Opt	MHB3
0.1	10^2	-6.79	-4.34	-3.33	28.71	355.02	-1.67	-9.30
0.2	10^2	-13.35	-3.99	0.96	21.34	286.41	-1.79	-13.68
0.3	10^2	-19.58	0.33	12.66	5.62	227.71	0.08	-14.18
0.4	10^2	-23.68	8.19	31.93	-11.37	179.72	7.34	-8.49
0.5	10^2	-25.34	18.39	54.39	-22.42	143.29	19.30	-0.00
0.6	10^2	-18.83	32.75	78.12	-20.93	127.39	48.98	20.43
0.7	10^2	9.67	57.41	102.09	12.50	145.87	110.65	54.94
0.8	10^2	53.85	83.92	112.70	115.54	241.48	248.13	91.91
0.9	10^2	-48.91	68.65	78.37	479.82	607.47	661.17	76.85
0.1	10^3	-5.33	-3.53	-2.65	18.72	208.08	-1.51	-7.09
0.2	10^3	-11.08	-4.32	-0.51	1.12	139.58	-2.87	-11.20
0.3	10^3	-15.59	-0.80	7.95	-15.69	86.75	-2.96	-12.10
0.4	10^3	-19.09	4.41	20.72	-31.90	46.89	-0.84	-10.39
0.5	10^3	-21.37	12.10	36.84	-40.13	20.39	3.55	-6.96
0.6	10^3	-17.15	23.89	56.10	-35.06	11.99	16.86	4.44
0.7	10^3	-5.73	37.57	72.57	-15.90	20.19	46.62	27.85
0.8	10^3	21.66	54.96	83.07	44.06	71.90	109.68	62.02
0.9	10^3	63.31	67.40	75.98	279.43	302.82	313.66	75.38
0.1	10^4	-4.54	-3.16	-2.24	10.25	135.16	-0.92	-5.49
0.2	10^4	-9.13	-3.99	-0.58	-9.31	73.52	-1.73	-8.44
0.3	10^4	-12.94	-1.91	5.30	-25.60	30.89	-2.05	-9.54
0.4	10^4	-15.92	2.23	14.91	-32.82	4.87	-1.54	-9.41
0.5	10^4	-15.83	10.21	28.97	-29.40	-5.30	1.42	-6.76
0.6	10^4	-14.85	17.42	42.24	-24.21	-8.21	6.67	-2.02
0.7	10^4	-7.93	27.42	55.22	-10.06	0.13	19.37	9.31
0.8	10^4	9.10	39.89	64.16	31.58	37.82	51.80	35.52
0.9	10^4	50.68	58.14	67.22	204.26	208.27	153.44	66.55
0.1	10^5	-3.69	-2.58	-1.71	4.62	93.62	-0.49	-4.15
0.2	10^5	-7.62	-3.58	-0.50	-14.72	38.99	-0.93	-6.37
0.3	10^5	-10.92	-2.33	4.05	-25.85	6.25	-1.05	-7.05
0.4	10^5	-13.00	1.40	11.75	-25.06	-6.58	-0.85	-7.28
0.5	10^5	-13.34	7.61	22.75	-18.40	-8.70	0.33	-6.16
0.6	10^5	-11.94	14.27	34.27	-12.02	-6.73	2.83	-3.90
0.7	10^5	-7.62	21.66	44.62	-3.33	-0.69	8.95	1.48
0.8	10^5	5.49	32.73	53.49	23.97	25.24	24.25	15.99
0.9	10^5	34.91	46.44	56.17	155.70	156.34	73.70	51.34

distribution with the mean 1 in LePage et al. (9) and Cohen et al. (10). From LePage et al. (9, Theorem 2) and Cohen et al. (10, equation (2.8)), it follows that, for $h_1, h_2 > \alpha$,

$$\left(\sum_{j=1}^{\infty} Z_j^{h_1}, \sum_{j=1}^{\infty} Z_j^{h_2} \right) \overset{d}{=} (U_{h_1}, U_{h_2}), \qquad \text{[S.54]}$$

where $Z_j = \Gamma_j^{-1/\alpha}$, $\Gamma_j = \sum_{i=1}^{j} E_i$ for $j \geq 1$ and $E_1, E_2, \ldots$, is a sequence of independent and identically distributed exponential variables with mean 1.

To prove Eq. (S.54), we check that a_n defined (in Section B) such that $1 - F_X(a_n) = n^{-1}$ also satisfies LePage et al. (9, equation (3)) so that we can apply the results in LePage et al. (9). Since F_X satisfies Eq. (9) in the main text and $a_n \sim \{n/\Gamma(1-\alpha)\}^{1/\alpha}$, we have

$$\lim_{n\to\infty} n\{1 - F_X(a_n t)\} = \lim_{n\to\infty} n \left\{ \frac{n^{1/\alpha} t}{\Gamma(1-\alpha)^{1/\alpha}} \right\}^{-\alpha} / \Gamma(1-\alpha) = t^{-\alpha},$$

which satisfies LePage et al. (9, equation (3)).

Next, we replace the function $\phi(z, d)$ in the proof of LePage et al. (9, Theorem 1) by (ψ_1, ψ_2) where $\psi_i(z, d) = \{\sum_{j=1}^{\infty} [z_j I\{z_j \in (\epsilon, \infty)\}]^{h_i}\}^{1/h_i}$ for $i = 1, 2$. Therefore, as in LePage et al. (9, Theorem 2) and Cohen et al. (10, (2.8)), we have, as $n \to \infty$,

$$\left(\frac{1}{a_n} \left\{ \sum_{i=1}^{n} X_i^{h_1} \right\}^{1/h_1}, \frac{1}{a_n} \left\{ \sum_{i=1}^{n} X_i^{h_2} \right\}^{1/h_2} \right) \xrightarrow{d} \left(\left\{ \sum_{j=1}^{\infty} Z_j^{h_1} \right\}^{1/h_1}, \left\{ \sum_{j=1}^{\infty} Z_j^{h_2} \right\}^{1/h_2} \right).$$

Table 4. MSE $(\times 10^3)$ for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with 10^4 Monte Carlo independent samples from $F(0.5, \alpha)$, for sample sizes $n = 10^2, 10^3, 10^4, 10^5$.

α	n	B1	B2	B3	HI	HI.M	HI.Opt	MHB3
0.1	10^2	0.44	0.49	0.45	6.69	129.57	0.47	0.46
0.2	10^2	1.60	1.99	1.79	11.42	89.75	1.93	1.76
0.3	10^2	3.28	4.50	4.32	15.82	64.35	4.61	3.99
0.4	10^2	5.11	7.53	8.21	21.89	50.79	8.50	6.85
0.5	10^2	7.08	11.04	13.36	29.28	46.20	14.68	10.76
0.6	10^2	9.24	14.78	18.68	39.71	52.45	25.28	15.28
0.7	10^2	13.81	19.67	23.41	57.73	75.42	50.26	20.52
0.8	10^2	22.36	21.79	22.30	116.70	160.64	131.64	21.61
0.9	10^2	57.75	10.69	9.40	556.50	708.05	646.60	9.44
0.1	10^3	0.22	0.24	0.20	4.11	45.76	0.10	0.18
0.2	10^3	0.81	0.96	0.71	6.36	24.47	0.41	0.63
0.3	10^3	1.56	2.09	1.59	9.85	15.10	0.88	1.19
0.4	10^3	2.41	3.50	3.05	13.37	12.23	1.59	1.89
0.5	10^3	3.32	5.29	5.44	16.97	12.84	2.60	2.91
0.6	10^3	3.81	6.83	8.22	18.33	14.29	4.32	4.14
0.7	10^3	4.25	8.32	10.67	19.80	17.00	8.28	6.23
0.8	10^3	5.91	9.83	11.64	28.47	28.39	21.43	9.45
0.9	10^3	8.12	8.27	7.90	144.27	153.30	119.16	7.87
0.1	10^4	0.14	0.15	0.11	2.61	20.05	0.02	0.09
0.2	10^4	0.50	0.57	0.36	4.73	9.13	0.09	0.27
0.3	10^4	1.00	1.27	0.82	7.33	6.31	0.19	0.49
0.4	10^4	1.50	2.10	1.59	9.15	6.39	0.34	0.69
0.5	10^4	1.91	3.09	2.97	9.12	6.76	0.54	0.90
0.6	10^4	2.24	4.00	4.60	7.93	6.13	0.85	1.22
0.7	10^4	2.40	4.80	6.19	6.46	5.30	1.52	1.77
0.8	10^4	2.64	5.39	6.90	6.97	6.67	4.35	3.53
0.9	10^4	5.45	6.11	6.23	57.51	58.26	26.36	6.15
0.1	10^5	0.09	0.10	0.07	1.91	10.16	0.01	0.05
0.2	10^5	0.34	0.39	0.22	3.88	4.49	0.02	0.15
0.3	10^5	0.65	0.81	0.46	5.51	4.00	0.04	0.23
0.4	10^5	1.00	1.36	0.96	5.66	4.03	0.07	0.33
0.5	10^5	1.26	1.98	1.81	4.38	3.39	0.12	0.40
0.6	10^5	1.47	2.61	2.96	2.86	2.30	0.18	0.51
0.7	10^5	1.62	3.22	4.12	1.88	1.63	0.32	0.69
0.8	10^5	1.62	3.58	4.77	1.85	1.78	0.92	1.12
0.9	10^5	2.78	3.87	4.32	28.91	28.99	5.90	3.73

Applying the continuous mapping theorem gives, as $n \to \infty$,

$$\left(\frac{1}{a_n^{h_1}} \sum_{i=1}^{n} X_i^{h_1}, \frac{1}{a_n^{h_2}} \sum_{i=1}^{n} X_i^{h_2} \right) \xrightarrow{d} \left(\sum_{j=1}^{\infty} Z_j^{h_1}, \sum_{j=1}^{\infty} Z_j^{h_2} \right).$$

On the other hand, from the proof of Lemma B.4 in the supplement, we also have, as $n \to \infty$,

$$\left(\frac{1}{a_n^{h_1}} \sum_{i=1}^{n} X_i^{h_1}, \frac{1}{a_n^{h_2}} \sum_{i=1}^{n} X_i^{h_2} \right) \xrightarrow{d} (U_{h_1}, U_{h_2}),$$

where (U_{h_1}, U_{h_2}) has the joint moment generating function defined in Theorem 6. Because the asymptotic distribution is unique, the identity in distribution in Eq. (S.54) has been proved.

Since $(U_{h_1}, U_{h_2}) \stackrel{d}{=} \left(\sum_{j=1}^{\infty} Z_j^{h_1}, \sum_{j=1}^{\infty} Z_j^{h_2} \right)$, it is straightforward to approximate $\left(\sum_{j=1}^{\infty} Z_j^{h_1}, \sum_{j=1}^{\infty} Z_j^{h_2} \right)$ by

$$\left(\sum_{j=1}^{m} Z_j^{h_1}, \sum_{j=1}^{m} Z_j^{h_2} \right) = \left(\sum_{j=1}^{m} \Gamma_j^{-h_1/\alpha}, \sum_{j=1}^{m} \Gamma_j^{-h_2/\alpha} \right) \qquad \text{[S.55]}$$

for large m. To examine the sensitivity of results to the choice of m, we set $m = 1000, m = 500,$ and $m = 2000$ in Tables 9, 10 and 11, respectively. The probability of rejecting the convergence in distributions in Theorems 3 and 6 is similar for $m = 500, 1000, 2000$ except when $\alpha = 0.7$ with pairs of moments $(1, 2), (1, 3), (1, 4)$. In these simulations, at $\alpha = 0.7$, a larger value of m leads to a lower power of the Kolmogorov-Smirnov test to reject the identity of the distributions for pairs of moments $(1, 2), (1, 3), (1, 4)$. Even larger values of m may be required to demonstrate the convergence in distributions for $\alpha > 0.7$.

Table 5. Bias $(\times 10^3)$ for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with 10^4 Monte Carlo independent samples from $F(2,\alpha)$, for sample sizes $n = 10^2, 10^3, 10^4, 10^5$.

α	n	B1	B2	B3	HI	HI.M	HI.Opt	MHB3
0.1	10^2	-9.54	-7.16	-6.32	28.71	355.02	-1.67	-10.74
0.2	10^2	-24.46	-15.56	-11.96	21.34	286.41	-1.79	-19.60
0.3	10^2	-45.03	-26.22	-18.34	5.62	227.71	0.08	-28.12
0.4	10^2	-70.47	-39.86	-25.60	-11.37	179.72	7.34	-35.16
0.5	10^2	-102.51	-58.16	-36.76	-22.42	143.29	19.30	-45.39
0.6	10^2	-140.81	-81.19	-51.87	-20.93	127.39	48.98	-57.23
0.7	10^2	-190.21	-108.89	-71.53	12.50	145.87	110.65	-73.57
0.8	10^2	-267.83	-149.30	-101.92	115.54	241.48	248.13	-102.21
0.9	10^2	-296.19	-234.98	-159.99	479.82	607.47	661.17	-160.00
0.1	10^3	-7.17	-5.41	-4.63	18.72	208.08	-1.51	-7.98
0.2	10^3	-18.44	-11.90	-8.92	1.12	139.58	-2.87	-14.66
0.3	10^3	-32.41	-18.16	-12.16	-15.69	86.75	-2.96	-19.72
0.4	10^3	-49.57	-26.70	-16.40	-31.90	46.89	-0.84	-24.49
0.5	10^3	-70.61	-37.01	-22.19	-40.13	20.39	3.55	-29.84
0.6	10^3	-92.39	-48.13	-28.53	-35.06	11.99	16.86	-34.09
0.7	10^3	-118.01	-63.05	-38.82	-15.90	20.19	46.62	-40.71
0.8	10^3	-151.10	-83.45	-55.39	44.06	71.90	109.68	-55.62
0.9	10^3	-211.64	-113.75	-81.87	279.43	302.82	313.66	-81.87
0.1	10^4	-5.93	-4.57	-3.73	10.25	135.16	-0.92	-6.17
0.2	10^4	-14.71	-9.68	-6.90	-9.31	73.52	-1.73	-11.08
0.3	10^4	-25.62	-14.89	-9.68	-25.60	30.89	-2.05	-15.30
0.4	10^4	-38.80	-21.03	-12.76	-32.82	4.87	-1.54	-19.40
0.5	10^4	-52.74	-26.65	-15.37	-29.40	-5.30	1.42	-22.25
0.6	10^4	-69.88	-35.88	-20.95	-24.21	-8.21	6.67	-25.99
0.7	10^4	-87.61	-46.04	-27.86	-10.06	0.13	19.37	-29.80
0.8	10^4	-106.79	-58.79	-38.33	31.58	37.82	51.80	-38.43
0.9	10^4	-132.80	-76.56	-55.44	204.26	208.27	153.44	-55.44
0.1	10^5	-4.81	-3.72	-2.92	4.62	93.62	-0.49	-4.73
0.2	10^5	-12.13	-8.16	-5.58	-14.72	38.99	-0.93	-8.61
0.3	10^5	-21.12	-12.72	-7.93	-25.85	6.25	-1.05	-12.01
0.4	10^5	-31.41	-17.25	-10.39	-25.06	-6.58	-0.85	-15.60
0.5	10^5	-42.82	-21.85	-12.67	-18.40	-8.70	0.33	-18.47
0.6	10^5	-55.64	-28.37	-16.49	-12.02	-6.73	2.83	-21.46
0.7	10^5	-69.84	-36.69	-22.12	-3.33	-0.69	8.95	-24.38
0.8	10^5	-83.01	-44.97	-28.86	23.97	25.24	24.25	-28.95
0.9	10^5	-98.01	-57.47	-41.57	155.70	156.34	73.70	-41.57

From Table 9, the two-sample Kolmogorov-Smirnov test cannot tell the differences between distributions in Eq. (S.51) and Eq. (S.52) with a large probability for $\alpha < 0.7$. When $\alpha \geq 0.7$, the convergence slows down and the Kolmogorov-Smirnov test starts to reject the identity of the distributions with high probabilities. The Kolmogorov-Smirnov test also suggests that, for ratios with higher orders $(h_1, h_2) = (2,3), (2,4), (3,4)$ of moments, the corresponding rates of convergence are faster than for those with orders $(h_1, h_2) = (1,2), (1,3), (1,4)$.

To examine the convergence in distribution in Theorem 9 or equivalently in Eq. (S.53) here, we perform the two-sample Kolmogorov-Smirnov test stated in Subsection 6.B independently 100 times for each $\alpha = 0, 1, 0.2, \ldots, 0.9$ and approximate the probabilities of rejecting $N_n^+/n^\alpha \xrightarrow{d} U^{-\alpha}/\Gamma(1-\alpha)$ in the column Thm9 in Table 9. It is clear that the two-sample Kolmogorov-Smirnov test cannot distinguish between the sampled distributions of N_n^+/n^α and $U^{-\alpha}/\Gamma(1-\alpha)$ for most of the sample sizes with $\alpha \leq 0.7$. For $\alpha > 0.7$, it requires sample size $n = 10^4$ before the Kolmogorov-Smirnov test does not distinguish the distribution of N_n^+/n^α from the distribution of $U^{-\alpha}/\Gamma(1-\alpha)$ most of the time.

To examine the convergence in $M_1'/(v_n^-)^{1/2} \xrightarrow{a.s.} 1$ as $n \to \infty$ in Theorem 1, we independently generate 100 random samples from $F(1,\alpha)$ with the sample size $n = 10^8$ and calculate $M_1'/(v_n^-)^{1/2}$ for each sample. Then we calculate the proportions of the samples such that $|M_1'/(v_n^-)^{1/2} - 1|$ is smaller than the tolerance levels $\{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$ in Table 12. We conclude that $M_1'/(v_n^-)^{1/2}$ converges to 1 slowly, especially for larger values of α. For example, $|M_1'/(v_n^-)^{1/2} - 1| < 10^{-6}$ only rarely for $\alpha \geq 0.5$, even with a sample size $n = 10^8$.

Table 6. MSE $(\times 10^3)$ for tail-index estimators B_1, B_2, B_3, HI.N, HI.M, HI.Opt, and MHB3 with 10^4 Monte Carlo independent samples from $F(2,\alpha)$, for sample sizes $n = 10^2, 10^3, 10^4, 10^5$.

α	n	B1	B2	B3	HI	HI.M	HI.Opt	MHB3
0.1	10^2	0.44	0.47	0.42	6.69	129.57	0.47	0.46
0.2	10^2	1.70	1.79	1.50	11.42	89.75	1.93	1.69
0.3	10^2	3.96	3.79	2.98	15.82	64.35	4.61	3.51
0.4	10^2	7.48	6.05	4.49	21.89	50.79	8.50	5.34
0.5	10^2	13.31	8.83	6.08	29.28	46.20	14.68	7.25
0.6	10^2	22.66	12.26	7.62	39.71	52.45	25.28	8.71
0.7	10^2	39.44	16.83	9.42	57.73	75.42	50.26	9.99
0.8	10^2	82.77	25.51	13.22	116.70	160.64	131.64	13.33
0.9	10^2	150.05	56.92	26.77	556.50	708.05	646.60	26.77
0.1	10^3	0.23	0.24	0.20	4.11	45.76	0.10	0.20
0.2	10^3	0.92	0.95	0.67	6.36	24.47	0.41	0.70
0.3	10^3	2.07	1.99	1.27	9.85	15.10	0.88	1.37
0.4	10^3	3.87	3.26	1.97	13.37	12.23	1.59	2.16
0.5	10^3	6.71	4.84	2.88	16.97	12.84	2.60	3.18
0.6	10^3	10.29	6.15	3.54	18.33	14.29	4.32	3.90
0.7	10^3	15.50	7.71	4.20	19.80	17.00	8.28	4.40
0.8	10^3	24.17	9.95	5.17	28.47	28.39	21.43	5.21
0.9	10^3	48.59	14.40	7.64	144.27	153.30	119.16	7.64
0.1	10^4	0.14	0.15	0.11	2.61	20.05	0.02	0.10
0.2	10^4	0.58	0.59	0.36	4.73	9.13	0.09	0.34
0.3	10^4	1.34	1.29	0.72	7.33	6.31	0.19	0.68
0.4	10^4	2.46	2.11	1.14	9.15	6.39	0.34	1.10
0.5	10^4	3.93	2.95	1.63	9.12	6.76	0.54	1.59
0.6	10^4	6.10	3.89	2.19	7.93	6.13	0.85	2.19
0.7	10^4	8.84	4.77	2.66	6.46	5.30	1.52	2.71
0.8	10^4	12.32	5.68	3.07	6.97	6.67	4.35	3.08
0.9	10^4	18.34	7.16	3.94	57.51	58.26	26.36	3.94
0.1	10^5	0.10	0.10	0.07	1.91	10.16	0.01	0.06
0.2	10^5	0.41	0.41	0.23	3.88	4.49	0.02	0.20
0.3	10^5	0.90	0.86	0.42	5.51	4.00	0.04	0.38
0.4	10^5	1.66	1.43	0.73	5.66	4.03	0.07	0.66
0.5	10^5	2.64	2.00	1.08	4.38	3.39	0.12	0.97
0.6	10^5	4.00	2.63	1.49	2.86	2.30	0.18	1.40
0.7	10^5	5.80	3.33	1.91	1.88	1.63	0.32	1.89
0.8	10^5	7.62	3.70	2.09	1.85	1.78	0.92	2.09
0.9	10^5	10.04	4.28	2.43	28.91	28.99	5.90	2.43

Table 7. Bias $(\times 10^3)$ (average of [R_L minus 2]) sample size $n = 10^2, 10^3, 10^4, 10^5$ from $F(1,\alpha)$ for $\alpha = 0.1, 0.2, \ldots, 0.9$.

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
10^5	0.00	0.00	-0.02	-0.10	-0.52	-2.61	-13.43	-80.07	-680.29
10^4	0.00	-0.02	-0.11	-0.49	-2.10	-8.37	-35.09	-167.29	-1185.48
10^3	-0.02	-0.17	-0.74	-2.65	-9.00	-29.96	-103.42	-404.20	-2430.44
10^2	-0.23	-1.48	-5.62	-17.01	-47.47	-129.41	-387.66	-1393.54	-8169.25

Table 8. MSE $(\times 10^3)$ (mean squared [R_L minus 2]) with sample size $n = 10^2, 10^3, 10^4, 10^5$ from $F(1,\alpha)$ for $\alpha = 0.1, 0.2, \ldots, 0.9$.

α	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$n = 10^5$	0.00	0.00	0.00	0.00	0.00	0.01	0.27	8.90	591.68
$n = 10^4$	0.00	0.00	0.00	0.00	0.01	0.12	1.94	40.75	1900.70
$n = 10^3$	0.00	0.00	0.00	0.01	0.16	1.64	18.14	260.27	9088.88
$n = 10^2$	0.00	0.01	0.08	0.74	5.46	37.33	343.31	5262.76	13123327.23

G. References

1. Brown, Mark, Cohen, Joel E., and de la Peña, Victor H. (2017). Taylor's law, via ratios, for some distributions with infinite mean. *Journal of Applied Probability* **54**, 657–669.
2. Einmahl, J.H.J. (1990). The empirical distribution function as a tail estimator. *Statistica Neerlandica* **44**, 79–82.
3. Albrecher, Hansjoerg, Ladoucette, Sophie A. and Teugels, Jef L. (2010). Asymptotics of the sample coefficient of variation and the sample dispersion. *Journal of Statistical Planning and Inference* **140**, 358–368.
4. Arnold, S. F. (1990). *Mathematical Statistics.* Prentice-Hall, Englewood Cliffs, NJ.
5. Feller, William (1971). *An introduction to probability theory and its applications. Vol. II.* John Wiley & Sons, Inc., New York-London-Sydney.
6. Grimmett, Geoffrey and Stirzacker, David (2001). *Probability and Random Processes (3rd ed.).* Oxford University Press, Oxford-New York.
7. Samorodnitsky, Gennady and Taqqu, Murad S. (1994). *Stable non-gaussian random processes.*. Chapman & Hall, New York.
8. Rockafellar, R Tyrrell, Uryasev, Stan and Zabarankin, Michael (2006). Generalized deviations in risk analysis. *Finance and Stochastics* **10**, 51–74.
9. LePage, Raoul, Woodroofe, Michael and Zinn, Joel (1981). Convergence to a stable distribution via order statistics. *Annals of Probability*, 624–632.
10. Cohen, Joel E., Davis, Richard A. and Samorodnitsky, Gennady (2020). Heavy-tailed distributions, correlations, kurtosis and Taylor's Law of fluctuation scaling. *Proceedings of the Royal Society A* **476**, 20200610.

Table 9. Probability of rejection of the convergence in distribution in Theorems 3, 6, and 9, according to the two-sample Kolmogorov-Smirnov test with margins of error = 0.05, for sample size $n = 10^2, 10^3, 10^4, 10^5$ from $F(1, \alpha)$ for $\alpha = 0.1, 0.2, \ldots, 0.9$, and $m = 1000$ in $\mathrm{Eq.}$ (S.55).

α	n	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)	Thm9
0.1	10^2	0.02	0.02	0.02	0.04	0.04	0.05	0.02
	10^3	0.05	0.04	0.03	0.04	0.07	0.07	0.01
	10^4	0.05	0.06	0.06	0.04	0.04	0.05	0.00
	10^5	0.05	0.05	0.07	0.07	0.07	0.07	0.00
0.2	10^2	0.04	0.05	0.05	0.04	0.03	0.03	0.01
	10^3	0.04	0.04	0.03	0.04	0.04	0.03	0.02
	10^4	0.06	0.05	0.07	0.06	0.06	0.06	0.00
	10^5	0.05	0.07	0.07	0.06	0.08	0.08	0.00
0.3	10^2	0.05	0.05	0.03	0.03	0.04	0.03	0.00
	10^3	0.03	0.04	0.03	0.02	0.02	0.02	0.02
	10^4	0.04	0.04	0.04	0.04	0.03	0.04	0.00
	10^5	0.07	0.06	0.04	0.04	0.05	0.07	0.00
0.4	10^2	0.07	0.06	0.05	0.02	0.04	0.02	0.00
	10^3	0.03	0.05	0.05	0.04	0.03	0.04	0.02
	10^4	0.08	0.06	0.05	0.04	0.04	0.05	0.00
	10^5	0.06	0.05	0.06	0.07	0.04	0.01	0.00
0.5	10^2	0.05	0.04	0.04	0.03	0.02	0.04	0.00
	10^3	0.03	0.02	0.02	0.04	0.04	0.03	0.03
	10^4	0.02	0.02	0.02	0.03	0.03	0.05	0.00
	10^5	0.08	0.07	0.07	0.07	0.07	0.04	0.00
0.6	10^2	0.03	0.04	0.03	0.03	0.06	0.07	0.02
	10^3	0.04	0.04	0.04	0.02	0.04	0.04	0.03
	10^4	0.08	0.06	0.05	0.04	0.04	0.05	0.00
	10^5	0.10	0.08	0.09	0.05	0.06	0.05	0.00
0.7	10^2	0.43	0.44	0.46	0.54	0.41	0.16	0.07
	10^3	0.98	0.88	0.77	0.05	0.04	0.02	0.07
	10^4	1.00	0.92	0.87	0.02	0.03	0.04	0.00
	10^5	0.98	0.85	0.79	0.06	0.06	0.05	0.00
0.8	10^2	1.00	1.00	1.00	1.00	1.00	0.95	0.28
	10^3	1.00	1.00	1.00	0.11	0.07	0.04	0.20
	10^4	1.00	1.00	1.00	0.03	0.04	0.04	0.00
	10^5	1.00	1.00	1.00	0.05	0.04	0.03	0.00
0.9	10^2	1.00	1.00	1.00	1.00	1.00	1.00	0.78
	10^3	1.00	1.00	1.00	1.00	1.00	0.18	0.76
	10^4	1.00	1.00	1.00	0.08	0.08	0.06	0.00
	10^5	1.00	1.00	1.00	0.03	0.04	0.03	0.00

Table 10. Probability of rejection of the convergence in distribution in Theorems 3, 6, and 9, according to the two-sample Kolmogorov-Smirnov test with margins of error = 0.05, for sample size $n = 10^2, 10^3, 10^4, 10^5$ from $F(1, \alpha)$ for $\alpha = 0.1, 0.2, \ldots, 0.9$, and $m = 500$ in Eq. (S.55).

α	n	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)	Thm9
0.1	10^2	0.05	0.06	0.07	0.07	0.08	0.08	0.00
	10^3	0.03	0.02	0.02	0.03	0.01	0.01	0.01
	10^4	0.04	0.03	0.04	0.05	0.04	0.04	0.00
	10^5	0.03	0.04	0.05	0.04	0.04	0.04	0.00
0.2	10^2	0.05	0.06	0.05	0.07	0.06	0.05	0.02
	10^3	0.00	0.01	0.02	0.01	0.01	0.02	0.01
	10^4	0.07	0.06	0.07	0.06	0.07	0.05	0.00
	10^5	0.02	0.02	0.02	0.02	0.01	0.02	0.00
0.3	10^2	0.08	0.08	0.06	0.07	0.08	0.09	0.00
	10^3	0.02	0.02	0.02	0.02	0.02	0.05	0.03
	10^4	0.06	0.08	0.08	0.06	0.05	0.04	0.00
	10^5	0.01	0.00	0.01	0.03	0.03	0.02	0.00
0.4	10^2	0.08	0.07	0.08	0.07	0.07	0.07	0.01
	10^3	0.04	0.02	0.01	0.02	0.02	0.02	0.01
	10^4	0.05	0.06	0.04	0.06	0.05	0.03	0.00
	10^5	0.01	0.02	0.02	0.02	0.02	0.03	0.00
0.5	10^2	0.09	0.09	0.07	0.06	0.07	0.08	0.01
	10^3	0.01	0.01	0.01	0.02	0.02	0.01	0.05
	10^4	0.07	0.09	0.09	0.05	0.06	0.08	0.00
	10^5	0.03	0.02	0.02	0.02	0.03	0.05	0.00
0.6	10^2	0.06	0.05	0.06	0.08	0.07	0.07	0.03
	10^3	0.03	0.01	0.01	0.02	0.02	0.01	0.08
	10^4	0.11	0.09	0.09	0.05	0.06	0.09	0.00
	10^5	0.08	0.05	0.06	0.02	0.01	0.02	0.00
0.7	10^2	0.93	0.90	0.85	0.50	0.33	0.16	0.05
	10^3	1.00	0.97	0.96	0.01	0.01	0.01	0.10
	10^4	1.00	0.99	0.97	0.08	0.07	0.08	0.00
	10^5	1.00	1.00	0.99	0.03	0.04	0.03	0.00
0.8	10^2	1.00	1.00	1.00	1.00	1.00	0.98	0.30
	10^3	1.00	1.00	1.00	0.04	0.03	0.01	0.23
	10^4	1.00	1.00	1.00	0.07	0.09	0.09	0.00
	10^5	1.00	1.00	1.00	0.03	0.03	0.02	0.00
0.9	10^2	1.00	1.00	1.00	1.00	1.00	1.00	0.78
	10^3	1.00	1.00	1.00	1.00	1.00	0.08	0.70
	10^4	1.00	1.00	1.00	0.09	0.07	0.10	0.00
	10^5	1.00	1.00	1.00	0.03	0.04	0.03	0.00

Table 11. Probability of rejection of the convergence in distribution in Theorems 3, 6, and 9, according to the two-sample Kolmogorov-Smirnov test with margins of error = 0.05, for sample size $n = 10^2, 10^3, 10^4, 10^5$ from $F(1, \alpha)$ for $\alpha = 0.1, 0.2, \ldots, 0.9$, and $m = 2000$ in Eq. (S.55).

α	n	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)	Thm9
0.1	10^2	0.06	0.06	0.06	0.05	0.06	0.06	0.00
	10^3	0.04	0.03	0.03	0.04	0.03	0.03	0.01
	10^4	0.01	0.01	0.00	0.00	0.00	0.01	0.00
	10^5	0.05	0.05	0.05	0.05	0.05	0.04	0.00
0.2	10^2	0.06	0.05	0.04	0.05	0.06	0.07	0.02
	10^3	0.05	0.04	0.04	0.02	0.03	0.02	0.01
	10^4	0.03	0.02	0.04	0.04	0.04	0.06	0.00
	10^5	0.03	0.04	0.04	0.05	0.05	0.04	0.00
0.3	10^2	0.09	0.08	0.08	0.06	0.05	0.05	0.01
	10^3	0.04	0.04	0.04	0.01	0.01	0.02	0.01
	10^4	0.03	0.06	0.05	0.02	0.03	0.02	0.00
	10^5	0.03	0.04	0.05	0.03	0.03	0.02	0.00
0.4	10^2	0.12	0.09	0.09	0.06	0.05	0.06	0.00
	10^3	0.03	0.05	0.05	0.06	0.04	0.04	0.01
	10^4	0.02	0.02	0.03	0.03	0.02	0.01	0.00
	10^5	0.04	0.05	0.05	0.04	0.05	0.06	0.00
0.5	10^2	0.11	0.08	0.08	0.04	0.04	0.05	0.01
	10^3	0.01	0.02	0.01	0.04	0.01	0.02	0.02
	10^4	0.05	0.04	0.06	0.01	0.00	0.00	0.00
	10^5	0.03	0.03	0.04	0.04	0.05	0.05	0.00
0.6	10^2	0.10	0.07	0.06	0.09	0.08	0.07	0.03
	10^3	0.01	0.03	0.02	0.02	0.03	0.03	0.02
	10^4	0.06	0.07	0.06	0.03	0.02	0.02	0.00
	10^5	0.07	0.05	0.04	0.04	0.04	0.04	0.00
0.7	10^2	0.15	0.18	0.17	0.57	0.40	0.13	0.05
	10^3	0.76	0.51	0.43	0.02	0.02	0.02	0.05
	10^4	0.79	0.55	0.48	0.02	0.04	0.01	0.00
	10^5	0.83	0.61	0.43	0.02	0.03	0.03	0.00
0.8	10^2	1.00	1.00	1.00	1.00	1.00	0.96	0.32
	10^3	1.00	1.00	1.00	0.10	0.06	0.04	0.17
	10^4	1.00	1.00	1.00	0.05	0.04	0.01	0.00
	10^5	1.00	1.00	1.00	0.03	0.04	0.03	0.00
0.9	10^2	1.00	1.00	1.00	1.00	1.00	1.00	0.81
	10^3	1.00	1.00	1.00	1.00	0.99	0.18	0.72
	10^4	1.00	1.00	1.00	0.09	0.07	0.10	0.00
	10^5	1.00	1.00	1.00	0.09	0.09	0.04	0.00

Table 12. Proportions of the ratios $M_1'/(v_n^-)^{1/2}$ satisfying $|M_1'/(v_n^-)^{1/2} - 1| < \varepsilon$ for tolerances $\varepsilon = 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}$ in 100 random samples of size $n = 10^8$ from $F(1, \alpha)$ for $\alpha = 0.1, 0.2, \ldots, 0.9$.

α	10^{-3}	10^{-4}	10^{-5}	10^{-6}
0.1	1.00	1.00	1.00	1.00
0.2	1.00	1.00	1.00	0.98
0.3	1.00	1.00	1.00	0.49
0.4	1.00	1.00	0.46	0.09
0.5	1.00	0.66	0.05	0.00
0.6	0.93	0.06	0.00	0.00
0.7	0.11	0.00	0.00	0.00
0.8	0.00	0.00	0.00	0.00
0.9	0.00	0.00	0.00	0.00