Econometric Theory, **0**, 2026, 1–47.
doi:10.1017/S0266466626100668

TAIL EXPECTILE ESTIMATION IN THE SEMIPARAMETRIC GENERALIZED PARETO MODEL

YASSER ABBAS
Toulouse School of Economics, Université de Toulouse Capitole

ABDELAATI DAOUIA
Toulouse School of Economics, Université de Toulouse Capitole

BOUTHEINA NEMOUCHI
Univ Rennes, Ensai, CNRS

GILLES STUPFLER
Univ Angers, CNRS, LAREMA, SFR MATHSTIC

Expectiles have received increasing attention as a market risk measure that is both coherent and elicitable. They are defined as a least squares analog to quantiles. Their estimation from heavy-tailed loss data in an extreme value framework is an important and difficult problem, especially when the target tail expectile is beyond the range of the data. This problem has been studied only fairly recently, using solely the Weissman extrapolation method. The competing generalized Pareto approach, which makes efficient use of tail observations by incorporating the location, scale, and shape extreme value parameters into the estimation procedure, has been left untouched, with a corresponding theory completely lacking. In this article, we challenge the dominance of the Weissman device by presenting and developing the theory of two classes of semiparametric Generalized Pareto estimators: the first class relies on direct asymmetric least squares estimation, while the second is based on extreme quantile estimation. Our estimators are found to outperform the best-known Weissman-type estimators for real-valued profit–loss distributions, while staying competitive for nonnegative loss variables. A forecast comparison exercise is also conducted on various sets of financial returns, showing the superiority of the generalized Pareto approach.

1. INTRODUCTION

The risk of a financial position X is usually quantified by a law-invariant risk measure $\xi(X)$, where positive values of X represent losses (e.g., negative log-returns). Two particularly desirable properties are coherence in the sense of Artzner et al. (1999) and elicitability, which makes it possible to evaluate and compare

Address correspondence to Abdelaati Daouia, Toulouse School of Economics, Université de Toulouse Capitole, Toulouse, France, email: abdelaati.daouia@tse-fr.eu.

© The Author(s), 2026. Published by Cambridge University Press.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

competing point forecasts through suitable scoring functions (Gneiting, 2011; Fissler and Ziegel, 2016). The ubiquitous *Value-at-Risk* at level $\tau \in (0,1)$ is the τth quantile $\mathrm{VaR}_\tau \equiv q_\tau := \inf\{x \in \mathbb{R} : F(x) \geq \tau\}$ of the distribution function F of the random variable X. Although intuitive, robust, and elicitable, VaR_τ may fail to satisfy the subadditivity property and does not reflect the magnitude of losses beyond q_τ (Daníelsson et al., 2001; Acerbi, 2002). The notion of *Expected Shortfall*, $\mathrm{ES}_\tau := (1-\tau)^{-1}\int_\tau^1 q_t \,\mathrm{d}t$, addresses these shortcomings and has become the regulatory standard for market risk under the Basel Accords III/IV (Acerbi and Tasche, 2002; Wang and Zitikis, 2021); when X is continuous, it equals $\mathbb{E}(X|X > q_\tau)$ (Rockafellar and Uryasev, 2002). Yet ES_τ is not directly elicitable, but only jointly elicitable with VaR_τ (Gneiting, 2011; Fissler and Ziegel, 2016). This trade-off motivates the search for law-invariant risk measures that combine coherence and elicitability.

The alternative concept of expectiles has axiomatic and practical properties that strike a delicate balance between those of the two aforementioned competing risk measures. First envisioned by Aigner, Amemiya, and Poirier (1976) and popularized by Newey and Powell (1987) in a regression context, the τth expectile can be defined as

$$\xi_\tau := \arg\min_{\theta\in\mathbb{R}} \mathbb{E}\left\{|\tau - \mathbb{1}\{X \leq \theta\}|(X-\theta)^2 - |\tau - \mathbb{1}\{X \leq 0\}|X^2\right\}.$$

Subtracting the (unimportant in practice) term $|\tau - \mathbb{1}\{X \leq 0\}|X^2$ in the cost function ensures that ξ_τ exists as the unique minimizer in the above optimization problem as soon as $\mathbb{E}|X| < \infty$. Expectiles factor in both the probability and magnitude of tail realizations (Kuan, Yeh, and Hsu, 2009). While the measure is categorized alongside quantiles into the same class of M-quantiles defined by Breckling and Chambers (1988), it distinguishes itself within this class by being the only M-quantile that satisfies the coherence property (Bellini et al., 2014). In fact, expectiles are the only coherent and law-invariant risk measure that is also elicitable, as shown by Ziegel (2016). They also enjoy intuitive financial interpretations in the form of their acceptance sets and their connection to the gain-loss ratio (Bellini and Di Bernardino, 2017). A proper comparison between expectiles and quantiles when they are identical for different τ levels, thus defining a market risk measure with the same meaning and properties, is provided by Daouia, Girard, and Stupfler (2018) from an extreme value perspective. Philipps (2022) has developed a range of further interpretations of expectiles from multiple analytical perspectives. Although not comonotonically additive, that is, $\xi_\tau(X+Y) \neq \xi_\tau(X) + \xi_\tau(Y)$ for comonotonic random variables X and Y (Acerbi and Székely, 2014), the unique combination of coherence, law-invariance, and elicitability has garnered mounting attention for expectiles (see, for instance, Bellini and Di Bernardino, 2017, Daouia, Girard, and Stupfler, 2021, Girard, Stupfler, and Usseglio-Carleve, 2022a, 2022b, and Zaevski and Nedeltchev, 2023). The use of this metric in measuring extreme risk has also received increasing attention in the fields of insurance and financial economics, including asset risk (Davison, Padoan, and Stupfler, 2023; Liu et al.,

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

2023) as well as systemic and cybersecurity risks (Daouia et al., 2018; Daouia, Stupfler, and Usseglio-Carleve, 2024b), to cite a few.

Estimation of tail risk at extreme levels τ is an important problem for practitioners who are concerned with the risk exposure to an infrequent catastrophic event that may wipe out an investment in terms of the size of potential losses. This statistical problem is difficult due to data scarcity at the far tails, especially when the population risk measure of interest is beyond the range of the data. There seems to be a general consensus in extreme value theory that there are two major extrapolation methods which make use of top observations for tail risk assessment: (i) the Weissman method which involves estimating an "intermediate" version of the risk measure (i.e., moderately extreme and well inside the sample) alongside the tail index, then using the latter to multiplicatively shift the former to the right place of the targeted "extreme" risk measure, and (ii) the generalized Pareto approach which approximates the upper tail of the underlying distribution by a Generalized (location-scale) Pareto distribution (GPD) tail. We refer to de Haan and Ferreira (2006) for extreme VaR_τ estimation and to Daouia, Stupfler, and Usseglio-Carleve (2025) for a unified theory of extreme ES_τ inference in the semiparametric GP model. Extreme ξ_τ estimation has been studied only fairly recently, by relying solely on the Weissman extrapolation device. The first extrapolated estimators were introduced by Daouia et al. (2018) and Daouia, Girard, and Stupfler (2019, 2020), before a series of refined and bias-corrected versions were developed in Girard et al. (2022a, 2022b), Padoan and Stupfler (2022), and Daouia et al. (2024b).

Semiparametric GP estimation of extreme expectiles has so far remained unexplored in the heavy-tailed setting, despite the long-standing use of GP tail models for quantile-based risk measurement in finance (see McNeil and Frey, 2000 and Section 6.2 of Beirlant et al., 2004, among many others). Such models are also a staple of extreme value analysis in fields such as climate science (Coles, 2001). Whereas Weissman extrapolation relies on scale and shape information only, the GP approximation additionally incorporates a location parameter. It can therefore accommodate location shifts away from the scale–shape Pareto model, potentially improving estimation and forecasting accuracy.

Apart from Daouia, Padoan, and Stupfler (2024a), who consider short-tailed distributions, this article is the first to develop semiparametric GP estimators of extreme expectiles under heavy tails. We introduce two classes of estimators: the first class of direct empirical counterparts is based on asymmetric least squares, and the second, indirect class based on quantiles, that combines extreme quantile estimation with the asymptotic proportionality between extreme quantiles and expectiles. We establish their asymptotic distributions for generic estimators of the scale and shape parameters, and then specialize the general theory to Generalized Pareto maximum likelihood (GPML) estimators (Smith, 1987; Drees, Ferreira, and de Haan, 2004) and moment estimators (Dekkers, Einmahl, and de Haan, 1989). Our Monte Carlo study shows that the proposed estimators outperform the best-known Weissman-type competitors for real-valued profit–loss distributions

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

while remaining competitive for nonnegative losses; the rolling-window forecast comparison likewise favors the GP approach for the financial series considered. Although our empirical focus is financial risk, the methodology and asymptotic theory apply more broadly to systemic risk, climate risk, cyber risk, geohazards, and disaster risk. The procedures are implemented in the freely available R package `Expectrem`.

The article is organized as follows. In Section 2, we construct the two classes of generalized Pareto-type estimators and derive their asymptotic properties. A simulation study examines their finite-sample performance in Section 3 relative to the best-known Weissman-type estimators. Forecast comparisons are also conducted in Section 4 to evaluate tail risk for the stocks of Netflix, Walmart, and American Express, as well as the exchange rate of Bitcoin. Section 5 concludes. The Appendix contains an overview of the current state of research on tail expectile estimation, and all necessary mathematical proofs.

2. SEMIPARAMETRIC GP ESTIMATION OF TAIL EXPECTILES

2.1. Statistical Model

Let the financial position of interest X be a real-valued random variable, and F be its cumulative distribution function. In this article, we consider loss returns (i.e., minus log-returns), which implies that extreme losses correspond to the upper tail of F. Denote by U the associated tail quantile function defined as $U(t) = F^{\leftarrow}(1 - 1/t)$ for all $t > 1$, where $F^{\leftarrow}(\cdot) = \inf\{x \in \mathbb{R}, F(x) \geqslant \cdot\}$ stands for the generalized inverse of F. Estimating tail quantities, including expectiles and quantiles, from the perspective of extreme value statistics relies on the fundamental domain of attraction assumption that there exist a constant $\gamma \in \mathbb{R}$ and a positive scale function $a(\cdot)$ such that, for all $x > 0$,

$$\lim_{t\to\infty} \frac{U(tx) - U(t)}{a(t)} = \int_1^x s^{\gamma-1}\,\mathrm{d}s. \tag{2.1}$$

The parameter γ is referred to as the extreme value index (EVI for short): a negative, zero, or positive EVI indicates, respectively, a distribution with short, light, or heavy right tail. In the case of heavy-tailed distributions (i.e., $\gamma > 0$) that describe quite well the tail structure of most financial data, following Corollary 1.2.10 in de Haan and Ferreira (2006), the condition (2.1) is equivalent to the Weissman extrapolation relationship

$$\lim_{t\to\infty} \frac{U(tx)}{U(t)} = x^{\gamma}, \quad \text{for } x > 0. \tag{2.2}$$

Equations (2.1) and (2.2) can, respectively, be rewritten as the GP approximation

$$\forall x > 0,\ U(tx) \approx U(t) + a(t)\frac{x^{\gamma} - 1}{\gamma} \text{ when } t \text{ is large} \tag{2.3}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

and the Weissman approximation

$$\forall x > 0,\ U(tx) \approx x^{\gamma} U(t) \text{ when } t \text{ is large.} \tag{2.4}$$

The asymptotic theory of extreme value estimators for tail quantities, including the EVI and extreme expectiles and quantiles, shows that the bias incurred in using these approximations is mainly determined by the speed of convergence in the domain of attraction condition (2.1) or (2.2). This is captured, for a general $\gamma \in \mathbb{R}$, by the so-called extended second-order regular variation (SORV) assumption:

$\mathcal{E}_2(\gamma, \rho, a, A)$: There exist $\gamma \in \mathbb{R}$, a second-order parameter $\rho \leqslant 0$, a scale function $a(\cdot) > 0$, and a constant sign function $A(\cdot)$ which converges to 0 at infinity such that

$$\forall x > 0,\ \lim_{t\to\infty} \frac{1}{A(t)} \left(\frac{U(tx) - U(t)}{a(t)} - \frac{x^{\gamma} - 1}{\gamma} \right) = \int_1^x s^{\gamma-1} \int_1^s u^{\rho-1} \mathrm{d}u\, \mathrm{d}s.$$

This assumes that the right tail of the distribution of X can be approximated by a GPD tail at a known rate. As can be seen, for instance, from Sections 2.3 and 2.4 and Chapter 3 in de Haan and Ferreira (2006), as well as Section 5.6 in Beirlant et al. (2004), this kind of condition is inevitable in order to carry out extreme value analysis when no parametric model structure is imposed on F. In the heavy-tailed case, it comes as a consequence of Theorem 3.1 in Fraga Alves et al. (2007) that this extended SORV condition reduces to the assumption:

$\mathcal{C}_2(\gamma, \tilde{\rho}, \tilde{A})$: There are $\gamma > 0$, a second-order parameter $\tilde{\rho} \leqslant 0$, and a constant sign function $\tilde{A}(\cdot)$ which converges to 0 at infinity such that

$$\forall x > 0,\ \lim_{t\to\infty} \frac{1}{\tilde{A}(t)} \left(\frac{U(tx)}{U(t)} - x^{\gamma} \right) = x^{\gamma} \int_1^x s^{\tilde{\rho}-1}\, \mathrm{d}s.$$

This condition can, for instance, be checked using goodness-of-fit statistics in the form of weighted integrals over the tail of the empirical quantile function, or over its distribution function counterpart. See Section 5.2 in de Haan and Ferreira (2006) for several examples.

Theorem 3.1 by Fraga Alves et al. (2007) provides explicit expressions of the second-order parameters $\tilde{\rho}$ and $\tilde{A}(\cdot)$ in terms of $\gamma, \rho, a(\cdot)$ and $A(\cdot)$. These parameters $(\tilde{\rho}, \tilde{A})$ (resp. (ρ, A) in $\mathcal{E}_2$) control the bias incurred in using the Weissman (resp. GPD) approximation (2.4) (resp. (2.3)): the more negative $\tilde{\rho}$ (resp. ρ) is, the faster $\tilde{A}(t)$ (resp. $A(t)$) converges to 0 as $t \to \infty$ and the smaller the bias incurred is. Daouia et al. (2025) and Daouia et al. (2024b) provide a list of examples of commonly used families of continuous distributions satisfying both of these conditions, and with the explicit corresponding values of $\gamma, \rho, a(\cdot)$ and $A(\cdot)$.

The crucial point is that it is not, in general, the case that $\rho = \tilde{\rho}$ and $A = \tilde{A}$. As a matter of fact, it is a corollary of Theorem 3.1(i) by Fraga Alves et al. (2007) that all standard continuous distributions that satisfy both conditions have $\rho \leq \tilde{\rho}$;

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

this means that the bias incurred in using the GPD approximation (2.3) is never qualitatively worse than the bias incurred in using the Weissman approximation (2.4). In fact, the former can be substantially smaller: it is seen from Table 4 in Daouia et al. (2024b) and Table 2 in Daouia et al. (2025) that the Fréchet distribution with $\alpha = 1$ and the Inverse- Γ distribution both have $\rho = 2\tilde{\rho}$. For pure GP observations, it is, of course, always better to rely on the GPD approach, since then condition $\mathcal{E}_2$ is in fact satisfied as an equality rather than a limit, while the GP distribution only satisfies condition $\mathcal{C}_2$ with $\tilde{\rho} = -\gamma$. The fact that the GPD approximation carries lower bias is one justification for substituting it for the Weissman approximation.

We focus here on the estimation of extreme expectiles when F is attracted to the maximum domain of Pareto-type distributions with tail index $0 < \gamma < 1$, using the GPD approximation. Together with condition $\mathbb{E}|X_-| < \infty$, with $X_- = \min(X, 0)$ the negative part of X, the assumption $\gamma < 1$ ensures that the first moment of X exists, and hence expectiles of X are well-defined (note that the first moment of X and hence its expectiles do not exist when $\gamma > 1$). Under this model assumption, Bellini and Di Bernardino (2017) have established in their Proposition 2.3 an interesting asymptotic connection between extreme expectiles and their quantile analogs, namely, the asymptotic equivalence

$$\xi_\tau \sim (\gamma^{-1} - 1)^{-\gamma} q_\tau \quad \text{as} \quad \tau \to 1. \tag{2.5}$$

It follows that tail expectiles ξ_τ are more spread than tail quantiles q_τ when $\gamma > 1/2$, whereas $\xi_\tau < q_\tau$ for all large τ when $\gamma < 1/2$. The asymptotic equivalence (2.5) is fundamental in the construction of Weissman-type estimators for extreme expectiles, which we review in Appendix A.

Our objective is, rather, to construct and study below two new classes of extreme expectile estimators: the first one in Section 2.2 is built upon asymmetric least squares extrapolation, while the second one in Section 2.3 relies on extreme quantile estimation. We build here a general theory based on generic estimators of the scale and shape parameters in conjunction with approximation (2.3); the application to GPML estimators and moment estimators of the scale and shape parameters is discussed in Section 2.4.

2.2. Extremal Asymmetric Least Squares Estimation

From now on, let $X_1, \ldots, X_n$ be a sample of observations drawn from F, with corresponding survival function $\overline{F} = 1 - F$, and let $X_{1,n} \leqslant \cdots \leqslant X_{n,n}$ be their corresponding order statistics. In the intermediate case when $\tau = \tau_n \uparrow 1$ such that $n\overline{F}(\xi_{\tau_n}) \to \infty$ as $n \to \infty$, the population expectile ξ_{τ_n} is extreme but well inside the sample so that it can be consistently estimated by its empirical version

$$\widehat{\xi}_{\tau_n} = \arg\min_{u \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} \eta_{\tau_n}(X_i - u), \tag{2.6}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

where $\eta_\tau(y) = |\tau - \mathbb{1}\{y \leqslant 0\}|y^2$, as described in Appendix A. In the more challenging extreme case when $\tau = \tau'_n > \tau_n$ such that $(1-\tau_n)/(1-\tau'_n) \to \infty$ as $n \to \infty$, the population expectile $\xi_{\tau'_n}$ lies at the far tail of the underlying distribution, possibly beyond the sample maximum $X_{n,n}$ in the data. Using (2.3) with $t = 1/\overline{F}(\xi_{\tau_n})$ and $x = \overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau'_n})$, and recalling that $\lim_{t\to\infty}(U(1/\overline{F}(t)) - t)/a(1/\overline{F}(t)) = 0$ (see, e.g., the proof of Theorem B.3.19 in de Haan and Ferreira, 2006), we obtain the expectile-specific approximation

$$\xi_{\tau'_n} \approx \xi_{\tau_n} + a(1/\overline{F}(\xi_{\tau_n}))\frac{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau'_n}))^\gamma - 1}{\gamma}$$

for n large enough. Since, in the heavy-tailed case with $0 < \gamma < 1$, we have $\overline{F}(\xi_\tau)/(1-\tau) \sim \gamma^{-1} - 1$ as $\tau \to 1$ (Bellini et al., 2014, Thm. 11), it follows that

$$\frac{\overline{F}(\xi_{\tau_n})}{\overline{F}(\xi_{\tau'_n})} \approx \frac{1-\tau_n}{1-\tau'_n}$$

as $n \to \infty$, which in turn suggests the extrapolating approximation

$$\xi_{\tau'_n} \approx \xi_{\tau_n} + a(1/\overline{F}(\xi_{\tau_n}))\frac{((1-\tau_n)/(1-\tau'_n))^\gamma - 1}{\gamma}.$$

Therefore, given estimators $\breve{\sigma}_n$ and $\breve{\gamma}_n$ of $a(1/\overline{F}(\xi_{\tau_n}))$ and γ, respectively, this motivates the extreme expectile estimator

$$\breve{\xi}^\star_{\tau'_n} = \breve{\xi}^\star_{\tau'_n}(\breve{\gamma}_n, \breve{\sigma}_n) := \widehat{\xi}_{\tau_n} + \breve{\sigma}_n\frac{((1-\tau_n)/(1-\tau'_n))^{\breve{\gamma}_n} - 1}{\breve{\gamma}_n}. \tag{2.7}$$

Various estimators of the scale function $\sigma(\cdot) \equiv a(1/\overline{F}(\cdot))$ and shape parameter γ are available in the literature, including the (pseudo-)GPML estimators and the Moment-type estimators: if $k = k_n \to \infty$ is a sequence of integers such that $k/n \to 0$, as $n \to \infty$, then

- The GPML estimators (Smith, 1987; Drees et al., 2004) of $(a(n/k_n), \gamma)$ are given by

  $$(\widehat{a}^{\mathrm{ML}}(n/k_n), \widehat{\gamma}^{\mathrm{ML}}_n) = \underset{\sigma,\gamma>0}{\arg\max}\prod_{i=1}^{k_n} h(X_{n-i+1,n} - X_{n-k_n,n}|\sigma,\gamma),$$

  where the Generalized Pareto density function is defined as $h(\cdot|\sigma,\gamma) = H'(\cdot|\sigma,\gamma)$.
- The Moment estimators (Dekkers et al., 1989) of $(a(n/k_n), \gamma)$ are defined as

  $$\left(\widehat{a}^{\mathrm{Mom}}(n/k_n), \widehat{\gamma}^{\mathrm{Mom}}_n\right) = \left(X_{n-k_n,n}M^{(1)}_{k_n}(1-\widehat{\gamma}^{(-)}_n), M^{(1)}_{k_n} + \widehat{\gamma}^{(-)}_n\right),$$

  $$\text{where } \widehat{\gamma}^{(-)}_n = 1 - \frac{1}{2}\left(1 - \frac{(M^{(1)}_{k_n})^2}{M^{(2)}_{k_n}}\right)^{-1}$$

  $$\text{and } M^{(j)}_{k_n} = \frac{1}{k_n}\sum_{i=1}^{k_n}(\log X_{n-i+1,n} - \log X_{n-k_n,n})^j, \text{ for } j = 1,2.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Such estimators of the scale parameter $a(1/\overline{F}(u_n))$ typically converge on the relative scale at the rate $1/\sqrt{n\overline{F}(u_n)}$ (see Sections 3.4 and 4.2 in de Haan and Ferreira, 2006, for the i.i.d. case, and Section 6 in Drees, 2003, for the case of dependent data). For $u_n = \xi_{\tau_n}$, estimators $\breve{\sigma}_n$ of the parameter $a(1/\overline{F}(\xi_{\tau_n}))$ are therefore expected to converge on the relative scale at the rate $1/\sqrt{n\overline{F}(\xi_{\tau_n})}$, which we have seen to be asymptotically proportional to the rate of convergence $1/\sqrt{n(1-\tau_n)}$ of the intermediate expectile estimator $\widehat{\xi}_{\tau_n}$ (see Theorem 2 below). It is also customary to use estimators $\breve{\gamma}_n$ that rely on the top $k_n = n(1-\tau_n)$ observations and converge at the same rate $1/\sqrt{n(1-\tau_n)}$ (see Sections 3.4–3.6 in de Haan and Ferreira, 2006, in the i.i.d. case, and Section 6 in Drees, 2003, in the serially dependent case). A noteworthy difference between the GPML and Moment estimators is that the former are location- and scale-invariant, while the latter are only scale-invariant. This may be of importance in practical applications where the data are standardized and de-trended.

The next result shows that among the three estimators $\widehat{\xi}_{\tau_n}$, $\breve{\gamma}_n$, and $\breve{\sigma}_n$ on which $\breve{\xi}^{\star}_{\tau'_n}$ hinges in (2.7), it is the asymptotic behavior of $\breve{\gamma}_n$ which dominates that of $\breve{\xi}^{\star}_{\tau'_n}$.

THEOREM 1. *Assume that*

1. $\mathbb{E}|X_-|^2 < \infty$ *and* $\mathcal{E}_2(\gamma,\rho,a,A)$ *holds with* $0 < \gamma < 1/2$ *and* $\rho < 0$*;*
2. $\tau_n = 1 - k_n/n \to 1$ *such that* $n(1-\tau_n) = k_n \to \infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}/q_{\tau_n} \to \lambda_2 \in \mathbb{R}$, *and* $\sqrt{n(1-\tau_n)}(a((1-\tau_n)^{-1})/q_{\tau_n} - \gamma) \to \mu \in \mathbb{R}$, $n \to \infty$*;*
3. $\tau'_n \to 1$ *such that* $(1-\tau_n)/(1-\tau'_n) \to \infty$ *and* $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau'_n)) \to \infty$, $n \to \infty$*;*
4. $\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) = \mathrm{O}_{\mathbb{P}}(1)$ *and for suitable estimators* $\breve{\gamma}_n$ *and* $\breve{\sigma}_n$*,*

$$\sqrt{n(1-\tau_n)}\left(\frac{\breve{\sigma}_n}{a(1/\overline{F}(\xi_{\tau_n}))} - 1\right) = \mathrm{O}_{\mathbb{P}}(1) \;\; \textit{and} \;\; \sqrt{n(1-\tau_n)}(\breve{\gamma}_n - \gamma) \xrightarrow{\mathrm{d}} \Gamma,$$

where Γ *is a nondegenerate limit.*

Then

$$\sqrt{n(1-\tau_n)}\,\frac{\breve{\xi}^{\star}_{\tau'_n} - \xi_{\tau'_n}}{a(1/\overline{F}(\xi_{\tau_n}))\int_1^{(1-\tau_n)/(1-\tau'_n)} s^{\gamma-1}\log(s)\mathrm{d}s} \xrightarrow{\mathrm{d}} \Gamma.$$

Under the i.i.d. model assumption, the property $\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1\right) = \mathrm{O}_{\mathbb{P}}(1)$ follows immediately from the asymptotic normality of $\widehat{\xi}_{\tau_n}$ derived in Theorem 1 of Daouia et al. (2020) when $0 < \gamma < 1/2$ and $\mathbb{E}|X_-|^2 < \infty$. Note also that by Corollary 4.3.2 on p. 135 in de Haan and Ferreira (2006), the conditions of our

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Theorem 1 imply

$$\frac{\int_1^{(1-\tau_n)/(1-\tau_n')} s^{\check{\gamma}_n-1}\log(s)\mathrm{d}s}{\int_1^{(1-\tau_n)/(1-\tau_n')} s^{\gamma-1}\log(s)\mathrm{d}s} \xrightarrow{\mathbb{P}} 1, \quad \text{as } n\to\infty,$$

which is useful for constructing an asymptotic confidence interval for $\xi_{\tau_n'}$. For this same purpose and for the calculation of $\check{\xi}^{\star}_{\tau_n'}$ in (2.7), the estimator $\check{\sigma}_n$ of $a(1/\overline{F}(\xi_{\tau_n}))$ can be obtained in three different ways by setting

$$\check{\sigma}_n^{(1)} = \widehat{a}(1/\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})), \quad \check{\sigma}_n^{(2)} = \widehat{a}((1-\tau_n)^{-1}) \times ((1-\tau_n)/\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n}))^{\check{\gamma}_n},$$
$$\text{or } \check{\sigma}_n^{(3)} = \widehat{a}((1-\tau_n)^{-1}) \times (\check{\gamma}_n^{-1}-1)^{-\check{\gamma}_n}.$$

The first estimator $\check{\sigma}_n^{(1)}$ is simply the empirical counterpart of $a(1/\overline{F}(\xi_{\tau_n}))$. The second estimator $\check{\sigma}_n^{(2)}$ is obtained in view of the approximation $a(1/\overline{F}(\xi_{\tau_n})) \approx a((1-\tau_n)^{-1}) \times ((1-\tau_n)/\overline{F}(\xi_{\tau_n}))^{\gamma}$ that follows from the regular variation property $\lim_{s\to\infty} a(sz)/a(s) = z^{\gamma}$, for any $z>0$ (see Lemma 1.2.9 on page 22 in de Haan and Ferreira, 2006), when taking $s = 1/\overline{F}(q_{\tau_n})$ and $z = \overline{F}(q_{\tau_n})/\overline{F}(\xi_{\tau_n})$. The third estimator $\check{\sigma}_n^{(3)}$ is obtained from the same approximation above by replacing $\overline{F}(\xi_{\tau_n})/(1-\tau_n)$ with its limit $\gamma^{-1}-1$. These three versions of $\check{\sigma}_n$ will be denoted in the sequel as $\check{\sigma}_n^{\mathrm{ML},i}$, $i=1,2,3$, when using the GPML method and as $\check{\sigma}_n^{\mathrm{Mom},i}$, $i=1,2,3$, when using the Moment method. This will in turn result in the associated GPML-based expectile estimators $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \check{\sigma}_n^{\mathrm{ML},i})$ and Moment-type estimators $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \check{\sigma}_n^{\mathrm{Mom},i})$. These estimators can be calculated by using our function `gpdExpect` from the R package `Expectrem` with the respective methods `"direct_ML1"`,..., `"direct_ML3"`, and `"direct_MOM1"`,..., `"direct_MOM3"`.

It is also natural and instructive to compare Theorem 1 with the results for the competing Weissman-type estimators reviewed in Appendix A. We have from the proof of Theorem 1 the following equivalent statement:

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\check{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}-1\right) \xrightarrow{\mathrm{d}} \Gamma. \tag{2.8}$$

Therefore, $\check{\xi}^{\star}_{\tau_n'}$ has the same rate of convergence as $\widetilde{\xi}^{\star}_{\tau_n'}$ in (A.6) and $\widehat{\xi}^{\star}_{\tau_n'}$ in (A.7), as well as their bias-corrected versions $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ in (B.5) and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}$ in (B.6). While the asymptotic distribution Γ of these Weissman-type estimators is typically that of the Hill estimator $\widehat{\gamma}_n^{\mathrm{H}}$ or the expectile-based estimator $\widehat{\gamma}_n^{\mathrm{E}}$ in (A.3), or their bias-reduced versions $\widehat{\gamma}_n^{\mathrm{H,BR}}$ and $\widehat{\gamma}_n^{\mathrm{E,BR}}$ in (A.10) and (A.11), it would naturally correspond in our generalized Pareto approach to the limit distribution of the GPML estimator $\widehat{\gamma}_n^{\mathrm{ML}}$ or Moment estimator $\widehat{\gamma}_n^{\mathrm{Mom}}$. Interestingly, the GPML and Moment methods often correct for bias and can be more robust in small and moderate samples or when the optimal k_n selection is unclear from Hill plots. Selection of the threshold k_n is known to be a difficult task in general. Schneider, Krajina, and Krivobokova (2021) develop two threshold selection methods under

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

the assumption that $\gamma > 0$ and compare them with eight competitors, including bootstrap techniques. Approaches that are specifically tailored to GP models are the changepoint tests of Northrop and Coleman (2014) and the simultaneous threshold stability plots of Wadsworth (2016) based on point process calculations. The latter two methods are implemented in the `R` package `mev` developed by Belzile et al. (2025). In our simulation study (Section 3), we shall examine the performance of our estimators as functions of the threshold, so as to find the best performers across a large range of threshold values; our financial real data analysis (Section 4) will use a simple selection rule based on finding the best extreme expectile forecast across time.

When the data follow a GPD, $\widehat{\gamma}_n^{\mathrm{ML}}$ is the estimator that tends to achieve the lowest mean squared error (MSE) in moderate to large samples, but can be unstable for small samples. The estimator $\widehat{\gamma}_n^{\mathrm{Mom}}$ is more robust when $\widehat{\gamma}_n^{\mathrm{ML}}$ becomes unstable, while still being competitive in terms of variance (see, e.g., Beirlant et al., 2004, and Yılmaz, Kara, and Özdemir, 2021, for a detailed comparison and review on estimation procedures for the EVI). The preference for $\breve{\xi}_{\tau_n'}^{\star}$ over the Weissman-type competitors stems not only from such practical advantages of the EVI estimators $\widehat{\gamma}_n^{\mathrm{ML}}$ and $\widehat{\gamma}_n^{\mathrm{Mom}}$ over the Hill-based alternatives, but also from its reliance on additional location parameter estimates that contribute to capturing uncertainty more effectively. Our experience with real and simulated heavy-tailed profit–loss data shows that, across a wide range of k_n values and without finite-sample correction, $\breve{\xi}_{\tau_n'}^{\star}$ stands out as the best performer in terms of smoothness and stability for moderate-sized datasets, but also in terms of relative bias and mean-squared error, with its Gaussian asymptotic confidence intervals attaining coverage closer to the nominal level (because it fits the exceedances over a threshold using a parametric model, whereas the nonparametric Weissman estimators are more data-driven and require large samples to provide accurate estimates). We will specialize the discussion of our generic Theorem 1 to both shape parameter estimators $\breve{\gamma}_n \in \{\widehat{\gamma}_n^{\mathrm{ML}}, \widehat{\gamma}_n^{\mathrm{Mom}}\}$ in Section 2.4 when the scale parameter estimator takes the form $\breve{\sigma}_n := \breve{\sigma}_n^{(3)} = \widehat{a}((1-\tau_n)^{-1})\,(\breve{\gamma}_n^{-1} - 1)^{-\breve{\gamma}_n}$, with $\widehat{a}(\cdot) \in \{\widehat{a}^{\mathrm{ML}}(\cdot), \widehat{a}^{\mathrm{Mom}}(\cdot)\}$. This choice of $\breve{\sigma}_n$ is motivated by its good performance in our numerical illustrations; one could, analogously, state a result for $\breve{\sigma}_n := \breve{\sigma}_n^{(2)}$ given the $\sqrt{n(1-\tau_n)}$-convergence of $\widehat{\widehat{F}}_n(\widehat{\xi}_{\tau_n})$ (which, in the i.i.d. setting, follows from a Gaussian approximation to the tail empirical process, see, e.g., Theorem 5.1.2 on page 159 in de Haan and Ferreira, 2006), and for $\breve{\sigma}_n := \breve{\sigma}_n^{(1)}$ given a specific, explicit form of the estimator $\widehat{a}(\cdot)$. Next, we show that its associated assumption in Theorem 1 holds under convenient conditions on both $\breve{\gamma}_n$ and $\widehat{a}((1-\tau_n)^{-1})$.

THEOREM 2. *Assume that*

1. $\mathbb{E}|X_-| < \infty$ *and* $\mathcal{E}_2(\gamma, \rho, a, A)$ *holds with* $0 < \gamma < 1$ *and* $\rho < 0$*;*
2. $\tau_n \uparrow 1$ *such that* $n(1-\tau_n) \to \infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) = \mathrm{O}(1)$, $\sqrt{n(1-\tau_n)}/q_{\tau_n} = \mathrm{O}(1)$, *and* $\sqrt{n(1-\tau_n)}(a((1-\tau_n)^{-1})/q_{\tau_n} - \gamma) = \mathrm{O}(1)$, $n \to \infty$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Take $\check{\sigma}_n = \widehat{a}((1-\tau_n)^{-1}) \times (\check{\gamma}_n^{-1} - 1)^{-\check{\gamma}_n}$ *for suitable estimators* $\widehat{a}((1-\tau_n)^{-1})$ *and* $\check{\gamma}_n$ *such that*

$$\sqrt{n(1-\tau_n)}\left(\frac{\widehat{a}((1-\tau_n)^{-1})}{a((1-\tau_n)^{-1})} - 1\right) = \mathrm{O}_{\mathbb{P}}(1) \text{ and } \sqrt{n(1-\tau_n)}(\check{\gamma}_n - \gamma) \xrightarrow{\mathrm{d}} \Gamma,$$

where Γ *is a nondegenerate limit. Then*

$$\sqrt{n(1-\tau_n)}\left(\frac{\check{\sigma}_n}{a(1/\overline{F}(\xi_{\tau_n}))} - 1\right) = \mathrm{O}_{\mathbb{P}}(1).$$

In the financial context of real-valued profit–loss heavy-tailed distributions, we provide Monte Carlo evidence in Section 3 that, at least, our GPML estimator $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \check{\sigma}_n^{\mathrm{ML},2})$ outperforms the best-known Weissman competitors among those described in Appendix A, both in terms of relative bias and mean-squared error, with its associated asymptotic confidence interval having slightly better coverage at the expense of being wider.

2.3. Extremal Quantile-Based Estimation

Using now (2.3) with $t = (1-\tau_n)^{-1}$ and $x = (1-\tau_n)/(1-\tau_n')$ leads to the well-known Generalized Pareto quantile approximation

$$q_{\tau_n'} \approx q_{\tau_n} + a((1-\tau_n)^{-1})\frac{((1-\tau_n)/(1-\tau_n'))^{\gamma} - 1}{\gamma}$$

for all n large enough, which results in the extrapolated extreme quantile estimator

$$\check{q}^{\star}_{\tau_n'} = \widehat{q}_{\tau_n} + \frac{\widehat{\sigma}_n}{\widehat{\gamma}_n}\left(\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_n} - 1\right)$$

for suitable estimators $\widehat{\sigma}_n$ and $\widehat{\gamma}_n$ of the scale and shape parameters $a((1-\tau_n)^{-1})$ and γ, respectively. In our heavy-tailed model ($\gamma > 0$), the asymptotic connection $\xi_{\tau_n'} \sim (\gamma^{-1} - 1)^{-\gamma} q_{\tau_n'}$ from (2.5) motivates then the quantile-based estimator

$$\begin{aligned}\check{\xi}^{\star}_{\tau_n'} = \check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n, \widehat{\sigma}_n) &:= (\widehat{\gamma}_n^{-1} - 1)^{-\widehat{\gamma}_n}\check{q}^{\star}_{\tau_n'} \\ &= (\widehat{\gamma}_n^{-1} - 1)^{-\widehat{\gamma}_n}\left(\widehat{q}_{\tau_n} + \frac{\widehat{\sigma}_n}{\widehat{\gamma}_n}\left[\left(\frac{1-\tau_n'}{1-\tau_n}\right)^{-\widehat{\gamma}_n} - 1\right]\right)\end{aligned} \tag{2.9}$$

for $\xi_{\tau_n'}$, which extrapolates the intermediate expectile estimator $\check{\xi}_{\tau_n} := (\widehat{\gamma}_n^{-1} - 1)^{-\widehat{\gamma}_n}\widehat{q}_{\tau_n}$ to the far tail at τ_n'. Next, we establish the asymptotic distribution of this new GPD-type estimator for generic estimators $\widehat{\gamma}_n$ and $\widehat{\sigma}_n$ of the shape and scale parameters.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

THEOREM 3. *Assume that*

1. $\mathbb{E}|X_-| < \infty$ *and* $\mathcal{E}_2(\gamma,\rho,a,A)$ *holds with* $0 < \gamma < 1$ *and* $\rho < 0$*;*
2. $\tau_n = 1 - k_n/n \to 1$ *such that* $n(1-\tau_n) = k_n \to \infty$, $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1}) \to \lambda_1 \in \mathbb{R}$, $\sqrt{n(1-\tau_n)}/q_{\tau_n} \to \lambda_2 \in \mathbb{R}$, *and* $\sqrt{n(1-\tau_n)}(a((1-\tau_n)^{-1})/q_{\tau_n} - \gamma) \to \mu \in \mathbb{R}$, $n \to \infty$*;*
3. $\tau_n' \to 1$ *such that* $(1-\tau_n)/(1-\tau_n') \to \infty$ *and* $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')) \to \infty$, $n \to \infty$*;*
4. *for suitable estimators* $\widehat{\gamma}_n$ *of* γ *and* $\widehat{\sigma}_n$ *of* $\sigma(q_{\tau_n}) = a((1-\tau_n)^{-1})$,

$$\sqrt{n(1-\tau_n)}\left(\widehat{\gamma}_n - \gamma, \frac{\widehat{\sigma}_n}{\sigma(q_{\tau_n})} - 1, \frac{\widehat{q}_{\tau_n} - q_{\tau_n}}{\sigma(q_{\tau_n})}\right) \xrightarrow{\mathrm{d}} (\Gamma,\Lambda,B), \tag{2.10}$$

where (Γ,Λ,B) *is a nontrivial trivariate weak limit.*

Then

$$\sqrt{n(1-\tau_n)}\frac{\check{\xi}^{\star}_{\tau_n'} - \xi_{\tau_n'}}{a((1-\tau_n)^{-1})\int_1^{(1-\tau_n)/(1-\tau_n')} s^{\gamma-1}\log(s)\mathrm{d}s} \xrightarrow{\mathrm{d}} \frac{\Gamma}{(\gamma^{-1}-1)^{\gamma}}.$$

As can be seen from the proof of Theorem 3, an equivalent statement of this convergence is

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\check{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} \Gamma, \tag{2.11}$$

which yields exactly the same rate of convergence as the competing Weissman-type estimators seen in Appendix A. Similarly to these estimators, it is the asymptotic distribution Γ of the chosen tail index estimator $\widehat{\gamma}_n$ which dominates the limit distribution of our extreme expectile estimator $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n,\widehat{\sigma}_n)$. In Section 2.4 below, we shall specialize the discussion of Theorem 3 to the GPML and Moment-type estimators $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}},\widehat{\sigma}_n^{\mathrm{ML}})$ and $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}},\widehat{\sigma}_n^{\mathrm{Mom}})$ obtained by taking the estimators $\widehat{\gamma}_n$ and $\widehat{\sigma}_n$ in (2.9) for the shape and scale parameters γ and $\sigma(q_{\tau_n}) \equiv a(n/k_n)$ to be either $\widehat{\gamma}_n^{\mathrm{ML}}$ and $\widehat{\sigma}_n^{\mathrm{ML}}$ or $\widehat{\gamma}_n^{\mathrm{Mom}}$ and $\widehat{\sigma}_n^{\mathrm{Mom}}$, respectively, where

$$\widehat{\sigma}_n^{\mathrm{ML}} := \widehat{a}^{\mathrm{ML}}(n/k_n) \quad \text{and} \quad \widehat{\sigma}_n^{\mathrm{Mom}} := \widehat{a}^{\mathrm{Mom}}(n/k_n).$$

The two extrapolated estimators $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}},\widehat{\sigma}_n^{\mathrm{ML}})$ and $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}},\widehat{\sigma}_n^{\mathrm{Mom}})$ can be calculated by using the function `gpdExpect` from the R package `Expectrem`, with `method= "indirect_ML"` and `method= "indirect_MOM"`, respectively. Our simulation study in Section 3 indicates that their finite-sample performance is very similar to that of the third direct GPML and Moment estimators $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}},\check{\sigma}_n^{\mathrm{ML},3})$ and $\check{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}},\check{\sigma}_n^{\mathrm{Mom},3})$, respectively (as the only difference lies, by construction, between their related intermediate direct and indirect expectile

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

estimators). Moreover, in the case of nonnegative loss distributions, the Moment-type estimator $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n)$ shows competitive relative mean-squared errors with remarkably lower relative bias estimates than those of the bias-corrected Weissman estimators. In the case of real-valued profit–loss distributions, this indirect Moment estimator still posts respectable results for low values of γ, but the direct GPML estimator $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \check{\sigma}^{\mathrm{ML},2}_n)$ stands out in terms of relative mean-squared errors and coverage probabilities of its associated confidence interval, exhibiting consistent results over all values of γ.

2.4. GPML and Moment-Type Estimators

In this section, we focus the discussion of the asymptotic theory on the GPML and Moment estimators $(\widehat{a}^{\mathrm{ML}}(n/k_n), \widehat{\gamma}^{\mathrm{ML}}_n)$ and $\left(\widehat{a}^{\mathrm{Mom}}(n/k_n), \widehat{\gamma}^{\mathrm{Mom}}_n\right)$ of the scale and shape parameters $(a(n/k_n), \gamma)$, for the discretized intermediate level $\tau_n = 1 - k_n/n$ with $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. Although our generic Theorems 1–3 are valid in a general setup allowing for serial dependency in the data, we impose an i.i.d. assumption on the underlying sample.

We first consider the extrapolated asymmetric least squares estimators $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \check{\sigma}^{\mathrm{ML}}_n)$ and $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \check{\sigma}^{\mathrm{Mom}}_n)$ defined in (2.7) by taking therein the estimators $\check{\gamma}_n$ and $\check{\sigma}_n$ of the shape and scale parameters γ and $\sigma(\xi_{\tau_n}) \equiv a(1/\overline{F}(\xi_{\tau_n}))$ to be either $\widehat{\gamma}^{\mathrm{ML}}_n$ and $\check{\sigma}^{\mathrm{ML}}_n$ or $\widehat{\gamma}^{\mathrm{Mom}}_n$ and $\check{\sigma}^{\mathrm{Mom}}_n$, respectively, with

$$\check{\sigma}^{\mathrm{ML}}_n := \widehat{a}^{\mathrm{ML}}(n/k_n) \times ((1/\widehat{\gamma}^{\mathrm{ML}}_n) - 1)^{-\widehat{\gamma}^{\mathrm{ML}}_n} = \widehat{\sigma}^{\mathrm{ML}}_n \times ((1/\widehat{\gamma}^{\mathrm{ML}}_n) - 1)^{-\widehat{\gamma}^{\mathrm{ML}}_n}$$

$$\text{and } \check{\sigma}^{\mathrm{Mom}}_n := \widehat{a}^{\mathrm{Mom}}(n/k_n) \times ((1/\widehat{\gamma}^{\mathrm{Mom}}_n) - 1)^{-\widehat{\gamma}^{\mathrm{Mom}}_n} = \widehat{\sigma}^{\mathrm{Mom}}_n \times ((1/\widehat{\gamma}^{\mathrm{Mom}}_n) - 1)^{-\widehat{\gamma}^{\mathrm{Mom}}_n}.$$

Following Theorem 3.4.2 on p. 92 in de Haan and Ferreira (2006) and the asymptotic independence between functions of the stochastic process $(X_{n-\lfloor k_n s \rfloor, n} - X_{n-k_n,n})_{0<s\le 1}$ and the random variable $X_{n-k_n,n}$ (see, e.g., Corollary 2.4.6 on page 52 in de Haan and Ferreira, 2006), the GPML estimators $\widehat{\gamma}^{\mathrm{ML}}_n$ and $\widehat{\sigma}^{\mathrm{ML}}_n$ satisfy, under the second-order condition $\mathcal{E}_2(\gamma, \rho, a, A)$ with $\gamma > 0$ and the bias condition $\sqrt{k_n} A(n/k_n) \to \lambda_1 \in \mathbb{R}$, the joint convergence

$$\sqrt{k_n}\left(\widehat{\gamma}^{\mathrm{ML}}_n - \gamma, \frac{\widehat{\sigma}^{\mathrm{ML}}_n}{a(n/k_n)} - 1, \frac{X_{n-k_n,n} - U(n/k_n)}{a(n/k_n)}\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\lambda_1 b^{\mathrm{ML}}, \Sigma^{\mathrm{ML}}\right), \quad \textbf{(2.12)}$$

where $b^{\mathrm{ML}} = \left(\frac{\gamma+1}{(1-\rho)(1+\gamma-\rho)}, \frac{-\rho}{(1-\rho)(1+\gamma-\rho)}, 0\right)$ and the matrix Σ^{ML} is given by

$$\Sigma^{\mathrm{ML}} = \begin{pmatrix} (1+\gamma)^2 & -(1+\gamma) & 0 \\ -(1+\gamma) & 1+(1+\gamma)^2 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Under the extra condition $\sqrt{n(1-\tau_n)}(a((1-\tau_n)^{-1})/q_{\tau_n} - \gamma) \to \mu \in \mathbb{R}$, it can be seen from Corollary 4.2.2 on p. 133 in de Haan and Ferreira (2006) that the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Moment estimators $\widehat{\gamma}_n^{\mathrm{Mom}}$ and $\widehat{\sigma}_n^{\mathrm{Mom}}$ satisfy the joint convergence

$$\sqrt{k_n}\left(\widehat{\gamma}_n^{\mathrm{Mom}}-\gamma,\frac{\widehat{\sigma}_n^{\mathrm{Mom}}}{a(n/k_n)}-1,\frac{X_{n-k_n,n}-U(n/k_n)}{a(n/k_n)}\right)\xrightarrow{\mathrm{d}}\mathcal{N}\left(\lambda^{\mathrm{Mom}}\left(b_{\gamma,\rho}^{\widehat{\gamma}_n^{\mathrm{Mom}}},b_{\gamma,\rho}^{\widehat{\sigma}_n^{\mathrm{Mom}}},0\right),\Sigma^{\mathrm{Mom}}\right),\qquad\textbf{(2.13)}$$

where

$$\lambda^{\mathrm{Mom}}:=\begin{cases}-\mu & \text{if } \gamma=-\rho \text{ or } (0<\gamma<-\rho \text{ and } l\neq 0),\\ \dfrac{\lambda_1\rho}{\gamma+\rho} & \text{if } 0<-\rho<\gamma \text{ or } (0<\gamma<-\rho \text{ and } l=0),\end{cases}$$

$$b_{\gamma,\rho}^{\widehat{\gamma}_n^{\mathrm{Mom}}}:=\begin{cases}-\dfrac{\gamma}{(1+\gamma)^2} & \text{if } 0<\gamma<-\rho \text{ and } l\neq 0,\\ \dfrac{\gamma-\gamma\rho+\rho}{\rho(1-\rho)^2} & \text{if } 0<-\rho\leq\gamma \text{ or } (0<\gamma<-\rho \text{ and } l=0),\end{cases}$$

$$b_{\gamma,\rho}^{\widehat{\sigma}_n^{\mathrm{Mom}}}:=\begin{cases}\dfrac{\gamma}{(1+\gamma)^2} & \text{if } 0<\gamma<-\rho \text{ and } l\neq 0,\\ \dfrac{-\rho}{(1-\rho)^2} & \text{if } 0<-\rho\leq\gamma \text{ or } (0<\gamma<-\rho \text{ and } l=0),\end{cases}$$

$$\text{and } \Sigma^{\mathrm{Mom}}:=\begin{pmatrix}\gamma^2+1 & \gamma-1 & 0\\ \gamma-1 & \gamma^2+2 & \gamma\\ 0 & \gamma & 1\end{pmatrix}.$$

Note that the asymptotic variance of $\widehat{\gamma}_n^{\mathrm{Mom}}$ is lower than that of $\widehat{\gamma}_n^{\mathrm{ML}}$, but the analysis of the respective biases of these two estimators, and therefore a definitive comparison of their finite-sample performance, is more complex. With the convergence results (2.12) and (2.13) at hand, we show that the two direct extreme expectile estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}},\breve{\sigma}_n^{\mathrm{ML}})$ and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}},\breve{\sigma}_n^{\mathrm{Mom}})$ are asymptotically normal.

COROLLARY 1. *Under the first three conditions of Theorem 1, and provided the data are i.i.d., we have*

$$\sqrt{k_n}\frac{\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}},\breve{\sigma}_n^{\mathrm{ML}})-\xi_{\tau_n'}}{\breve{\sigma}_n^{\mathrm{ML}}\int_1^{k_n/(n(1-\tau_n'))}s^{\widehat{\gamma}_n^{\mathrm{ML}}-1}\log(s)\mathrm{d}s}\xrightarrow{\mathrm{d}}\mathcal{N}\left(\frac{\lambda_1(1+\gamma)}{(1-\rho)(1+\gamma-\rho)},(1+\gamma)^2\right)$$

$$\textit{and }\sqrt{k_n}\frac{\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}},\breve{\sigma}_n^{\mathrm{Mom}})-\xi_{\tau_n'}}{\breve{\sigma}_n^{\mathrm{Mom}}\int_1^{k_n/(n(1-\tau_n'))}s^{\widehat{\gamma}_n^{\mathrm{Mom}}-1}\log(s)\mathrm{d}s}\xrightarrow{\mathrm{d}}\mathcal{N}\left(\lambda^{\mathrm{Mom}}b_{\gamma,\rho}^{\widehat{\gamma}_n^{\mathrm{Mom}}},\gamma^2+1\right).$$

Turning to the indirect quantile-based estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}},\widehat{\sigma}_n^{\mathrm{ML}})$ and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}},\widehat{\sigma}_n^{\mathrm{Mom}})$ defined in (2.9), we obtain the following asymptotic normality results.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

COROLLARY 2. *Under the first three conditions of Theorem 3, and provided the data are i.i.d., we have*

$$\sqrt{k_n}\frac{\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n,\widehat{\sigma}^{\mathrm{ML}}_n)-\xi_{\tau'_n}}{\widehat{\sigma}^{\mathrm{ML}}_n\int_1^{k_n/(n(1-\tau'_n))}s^{\widehat{\gamma}^{\mathrm{ML}}_n-1}\log(s)\mathrm{d}s}\xrightarrow{\mathrm{d}}\mathcal{N}\left(\frac{\lambda_1(\gamma+1)(\gamma^{-1}-1)^{-\gamma}}{(1-\rho)(1+\gamma-\rho)},(1+\gamma)^2(\gamma^{-1}-1)^{-2\gamma}\right)$$

and

$$\sqrt{k_n}\frac{\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n,\widehat{\sigma}^{\mathrm{Mom}}_n)-\xi_{\tau'_n}}{\widehat{\sigma}^{\mathrm{Mom}}_n\int_1^{k_n/(n(1-\tau'_n))}s^{\widehat{\gamma}^{\mathrm{Mom}}_n-1}\log(s)\mathrm{d}s}\xrightarrow{\mathrm{d}}\mathcal{N}\left(\lambda^{\mathrm{Mom}}b^{\widehat{\gamma}^{\mathrm{Mom}}_n}_{\gamma,\rho}(\gamma^{-1}-1)^{-\gamma},(\gamma^2+1)(\gamma^{-1}-1)^{-2\gamma}\right).$$

In order to keep the article to manageable proportions, we do not discuss the application of our theorems and all the associated ramifications in the presence of dependence in the data. Similarly to extreme quantile estimation under mixing conditions (Drees, 2003) or ARMA-GARCH models (He et al., 2022), our methods may work under serial dependence with an increased asymptotic variance. Theoretical results along these lines are left for future research. In our case study of financial time series, we consider a practical solution to eliminate serial dependence by filtering the time series with an ARMA-GARCH model. The residuals from the fitted model are then treated as i.i.d. after performing a series of Ljung–Box independence tests on residuals and their squares.

3. SIMULATION STUDY

We evaluate the performance of our extreme expectile estimators through an extensive simulation study. We consider two nonnegative loss distributions and two real-valued profit–loss distributions for the distribution of X, all with tail index $\gamma > 0$:

- Fréchet distribution with survival function $\overline{F}(x) = 1-\exp\left(-x^{-1/\gamma}\right)$ for $x > 0$;
- Burr distribution with survival function $\overline{F}(x) = 1/(1+x^{1/\gamma})$ for $x > 0$;
- Symmetric Burr distribution with density function $f(x) = \frac{1}{2\gamma}\frac{|x|^{1/\gamma-1}}{(1+|x|^{1/\gamma})^2}$ for $x \in \mathbb{R}$;
- Student distribution with degrees of freedom $1/\gamma$.

The first two nonnegative distributions are commonly seen in insurance applications, while the latter two real-valued distributions are more suitable for financial returns. For each distribution, we examine the cases $\gamma = 0.25$, 0.3, 0.35, and 0.4 for all competing estimators

$$\overline{\xi}_{\tau'_n}\in\left\{\widehat{\xi}_{\tau'_n},\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{E,BR}}_n),\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H,BR}}_n),\widehat{\xi}^{\star,\mathrm{PS}}_{\tau'_n},\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau'_n},\right.$$
$$\left.\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n,\breve{\sigma}^{\mathrm{ML},i}_n),\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n,\breve{\sigma}^{\mathrm{Mom},i}_n),\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n,\widehat{\sigma}^{\mathrm{ML}}_n),\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n,\widehat{\sigma}^{\mathrm{Mom}}_n)\right\},$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

where $i \in \{1,2,3\}$ and $\widehat{\xi}_{\tau_n'}$ is the naive, non-extrapolated empirical expectile estimator described in (2.6) at level τ_n'. The choice of γ values between 0.25 and 0.4 describes quite well the tail heaviness of most financial datasets (see, e.g., the R package `CASdatasets`). We generate $B = 5,000$ samples $\{X_1^{(b)}, \ldots, X_n^{(b)}\}$ for $b = 1, \ldots, B$, with $n = 1,500$ in each simulation case, and estimate the extreme expectile $\xi_{\tau_n'}$ at the level $\tau_n' = 1 - 1/n$, where $\xi_{\tau_n'}$ is of the same order of magnitude as the sample maximum $X_{n,n}$. We compare the finite-sample performance of the different estimators based on the following criteria:

1. Relative bias and mean-squared errors

$$\text{RBias}(\overline{\xi}_{\tau_n'}) = \frac{1}{B}\sum_{b=1}^{B}\left(\frac{\overline{\xi}_{\tau_n'}^{(b)}}{\xi_{\tau_n'}} - 1\right) \quad \text{and} \quad \text{RMSE}(\overline{\xi}_{\tau_n'}) = \frac{1}{B}\sum_{b=1}^{B}\left(\frac{\overline{\xi}_{\tau_n'}^{(b)}}{\xi_{\tau_n'}} - 1\right)^2,$$

   where $\overline{\xi}_{\tau_n'}^{(b)}$ is the expectile estimator generated from the bth sample.
2. Average length and coverage probability of the 95% asymptotic Gaussian confidence intervals, where the former is defined as the average width of the confidence intervals generated around each estimator, and the latter is the frequency of times said intervals included the true value of the extreme expectile.

We plot the estimates of $\text{RBias}(\overline{\xi}_{\tau_n'})$ and $\text{RMSE}(\overline{\xi}_{\tau_n'})$ against k for the positive-valued distributions in Figure 1, and those for the real-valued distributions in Figure 2. Figure 1 reveals that the BR (Bias-reduced) and PS (Padoan and Stupfler, 2022) estimators $\widehat{\xi}_{\tau_n'}^{\star,\text{BR}}(\widehat{\gamma}_n^{\text{E,BR}})$, $\widetilde{\xi}_{\tau_n'}^{\star,\text{BR}}(\widehat{\gamma}_n^{\text{H,BR}})$, $\widehat{\xi}_{\tau_n'}^{\star,\text{PS}}$, and $\widetilde{\xi}_{\tau_n'}^{\star,\text{PS}}$, seen in Appendix B and represented as solid lines in the figure, perform well in the positive-valued cases of the Fréchet and Burr distributions. However, the Burr distribution also sees our two direct GPML estimators $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{ML}}, \breve{\sigma}_n^{\text{ML},1})$ and $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{ML}}, \breve{\sigma}_n^{\text{ML},2})$, represented, respectively, by the dashed turquoise and purple lines, reach similar RMSE levels to their BR and PS competitors, but for narrower ranges of k values that are typically higher than the optimal k values for the BR and PS families of estimators. Our indirect and third direct Moment estimators $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{Mom}}, \widehat{\sigma}_n^{\text{Mom}})$ and $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{Mom}}, \breve{\sigma}_n^{\text{Mom},3})$, shown in dash-dotted blue and dashed pink, respectively, shore up this shortcoming by maintaining competitive RMSE measures over long ranges of k values for both the Fréchet and Burr distributions and low RBias estimates for both distributions and all values of γ.

The direct GPML estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{ML}}, \breve{\sigma}_n^{\text{ML},2})$, again in dashed purple in Figure 2, however, stands out as the overall best performer in terms of RMSE and RBias when the underlying distribution is real-valued, exhibiting consistent results over all values of γ. The competing GPML estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{ML}}, \breve{\sigma}_n^{\text{ML},1})$, in dashed turquoise, closely matches its top performing direct GPML counterpart only for the Student distribution with γ values 0.35 and 0.4, while the indirect and third direct GPML estimators $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{ML}}, \widehat{\sigma}_n^{\text{ML}})$ and $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\text{ML}}, \breve{\sigma}_n^{\text{ML},3})$, in dash-dotted green and dashed magenta, respectively, seem to be best for the Student distribution with

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

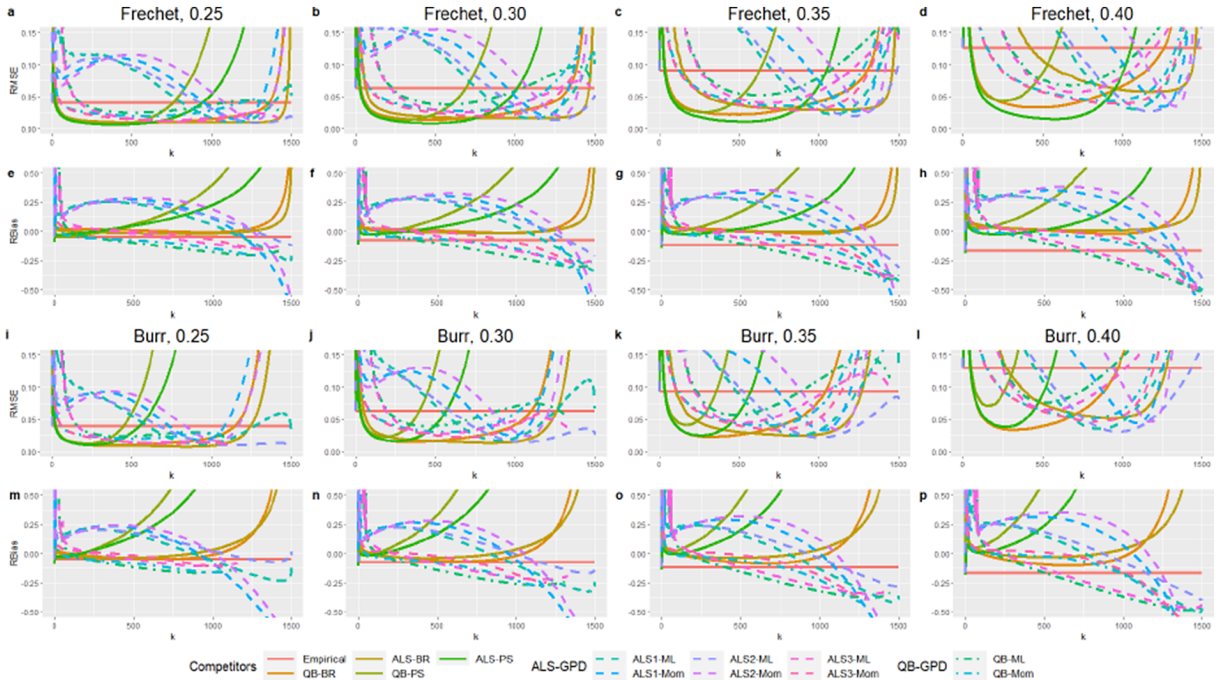


FIGURE 1. Simulation results for positive-valued distributions. Panels (a)–(d) and (e)–(h) show the Monte-Carlo estimates of RMSE and RBias, respectively, for the Fréchet distribution, while panels (i)–(l) and (m)–(p) show those for the Burr distribution. Each group of four panels is ordered by tail index value $\gamma \in \{0.25, 0.3, 0.35, 0.4\}$. For $n = 1{,}500$ and $\tau_n' = 1 - 1/n$, we consider the naive empirical estimator $\widehat{\xi}_{\tau_n'}$ (Empirical), the direct BR and PS estimators $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{E,BR}})$ (ALS-BR) and $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (ALS-PS), their indirect counterparts $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{H,BR}})$ (QB-BR) and $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (QB-PS), our direct GPML estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \breve{\sigma}_n^{\mathrm{ML},1})$ (ALS1-ML), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \breve{\sigma}_n^{\mathrm{ML},2})$ (ALS2-ML), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \breve{\sigma}_n^{\mathrm{ML},3})$ (ALS3-ML), the direct Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},1})$ (ALS1-Mom), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},2})$ (ALS2-Mom), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},3})$ (ALS3-Mom), and the indirect GPML and Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \widehat{\sigma}_n^{\mathrm{ML}})$ (QB-ML) and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \widehat{\sigma}_n^{\mathrm{Mom}})$ (QB-Mom).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

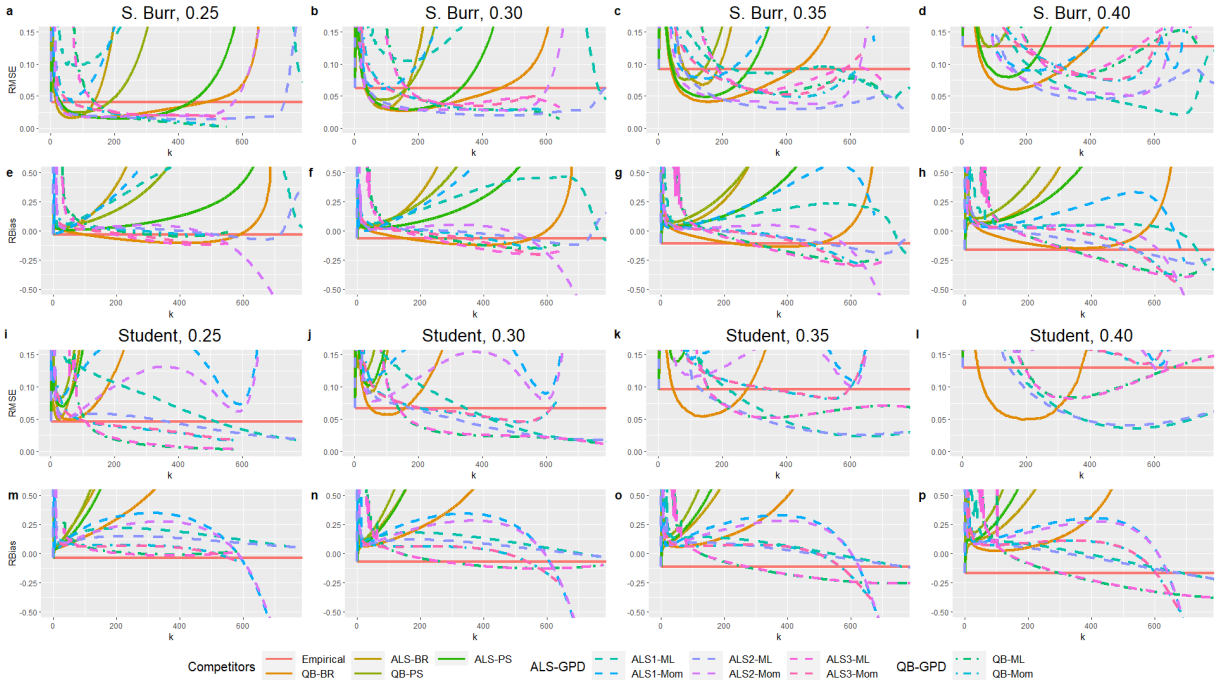


FIGURE 2. Simulation results for real-valued distributions. Panels (a)–(d) and (e)–(h) show the Monte-Carlo estimates of RMSE and RBias, respectively, for the symmetric Burr distribution, while panels (i)–(l) and (m)–(p) show those for the Student distribution. Each group of four panels is ordered by tail index value $\gamma \in \{0.25, 0.3, 0.35, 0.4\}$. For $n = 1{,}500$ and $\tau_n' = 1 - 1/n$, we consider the naive empirical estimator $\widehat{\xi}_{\tau_n'}$ (Empirical), the direct BR and PS estimators $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ (ALS-BR) and $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (ALS-PS), their indirect counterparts $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$ (QB-BR) and $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (QB-PS), our direct GPML estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},1}_n)$ (ALS1-ML), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},2}_n)$ (ALS2-ML), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},3}_n)$ (ALS3-ML), the direct Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},1}_n)$ (ALS1-Mom), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},2}_n)$ (ALS2-Mom), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},3}_n)$ (ALS3-Mom), and the indirect GPML and Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \widehat{\sigma}^{\mathrm{ML}}_n)$ (QB-ML) and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n)$ (QB-Mom).

γ values 0.25 and 0.3. The indirect Moment estimator $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n)$ and third direct Moment estimator $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \check{\sigma}^{\mathrm{Mom},3}_n)$, again in dash-dotted blue and dashed pink, respectively, post respectable results for both distributions, but only for $\gamma = 0.25$ and 0.3. These results imply that our extrapolated GPD-based estimators would be well-suited to the applications to financial risk analysis which feature real-valued profit–loss distributions with comparable tail heaviness to what we have examined in these simulation experiments.

When ignoring the asymptotic bias of our direct asymmetric least squares estimators $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \check{\sigma}^{\mathrm{ML},i}_n)$ and $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \check{\sigma}^{\mathrm{Mom},i}_n)$ for $i = 1, 2, 3$, in Corollary 1, their associated Gaussian $100(1-\alpha)\%$ asymptotic confidence intervals for $\xi_{\tau'_n}$ are

$$\check{I}^{\mathrm{ML},i}_{\tau'_n}(\alpha) = \left[\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \check{\sigma}^{\mathrm{ML},i}_n) \pm z_{1-\alpha/2} \frac{\check{\sigma}^{\mathrm{ML},i}_n \left(1+\widehat{\gamma}^{\mathrm{ML}}_n\right)}{\sqrt{k_n}} \phi_{\widehat{\gamma}^{\mathrm{ML}}_n}(d_n) \right]$$

$$\check{I}^{\mathrm{Mom},i}_{\tau'_n}(\alpha) = \left[\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \check{\sigma}^{\mathrm{Mom},i}_n) \pm z_{1-\alpha/2} \frac{\check{\sigma}^{\mathrm{Mom},i}_n \sqrt{\left(\widehat{\gamma}^{\mathrm{Mom}}_n\right)^2 + 1}}{\sqrt{k_n}} \phi_{\widehat{\gamma}^{\mathrm{Mom}}_n}(d_n) \right],$$

$$\text{with } d_n := k_n/(n(1-\tau'_n)) \text{ and } \phi_\gamma(t) := \int_1^t s^{\gamma-1} \log(s) \mathrm{d}s.$$

These confidence intervals can be calculated by using the function `gpdExpect` from the R package `Expectrem` with the respective methods `"direct_ML1"` ,..., `"direct_ML3"`, and `"direct_MOM1"` ,..., `"direct_MOM3"`. Likewise, under the assumption $\lambda_1 = \lambda^{\mathrm{Mom}} = 0$ in Corollary 2, the resulting Gaussian $100(1-\alpha)\%$ asymptotic confidence intervals for $\xi_{\tau'_n}$ based upon the indirect quantile-based estimators $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \widehat{\sigma}^{\mathrm{ML}}_n)$ and $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n)$ are

$$\check{I}^{\mathrm{ML}}_{\tau'_n}(\alpha) = \left[\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \widehat{\sigma}^{\mathrm{ML}}_n) \pm z_{1-\alpha/2} \frac{\widehat{a}^{\mathrm{ML}}(n/k_n) \left(1+\widehat{\gamma}^{\mathrm{ML}}_n\right)}{\sqrt{k_n} \left(\left(\widehat{\gamma}^{\mathrm{ML}}_n\right)^{-1} - 1 \right)^{\widehat{\gamma}^{\mathrm{ML}}_n}} \phi_{\widehat{\gamma}^{\mathrm{ML}}_n}(d_n) \right]$$

$$\text{and } \check{I}^{\mathrm{Mom}}_{\tau'_n}(\alpha) = \left[\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n) \pm z_{1-\alpha/2} \frac{\widehat{a}^{\mathrm{Mom}}(n/k_n) \sqrt{\left(\widehat{\gamma}^{\mathrm{Mom}}_n\right)^2 + 1}}{\sqrt{k_n} \left(\left(\widehat{\gamma}^{\mathrm{Mom}}_n\right)^{-1} - 1 \right)^{\widehat{\gamma}^{\mathrm{Mom}}_n}} \phi_{\widehat{\gamma}^{\mathrm{Mom}}_n}(d_n) \right].$$

These confidence intervals can also be calculated by using the function `gpdExpect` with `method= "indirect_ML"` and `method= "indirect_MOM"`, respectively.

We plot the coverage probabilities and average lengths of the 95% confidence intervals against k for the positive and real-valued distributions in Figures 3 and 4, respectively. While most confidence intervals associated with the 12 competing extrapolated estimators are able to get close to the intended 95% coverage as k varies, very few can achieve that for a wide range of k values. Our direct GPML confidence intervals $\check{I}^{\mathrm{ML},i}_{\tau'_n}$, and particularly those associated with the

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

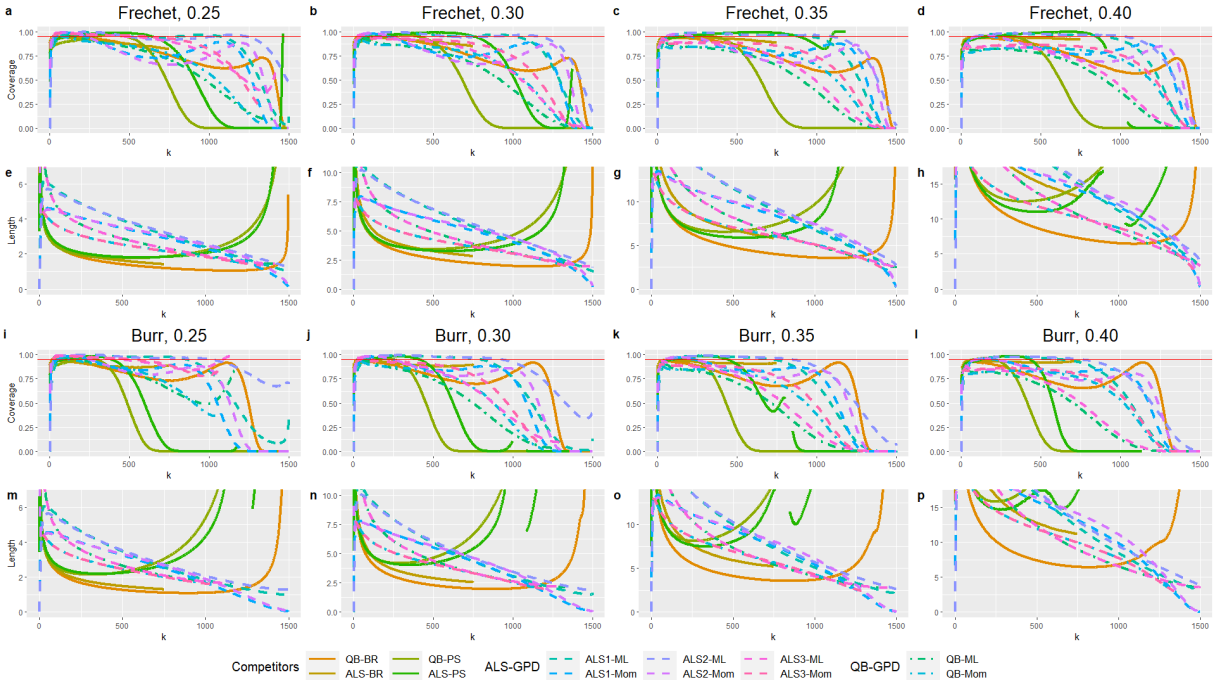


FIGURE 3. Simulation results for positive-valued distributions. Panels (a)–(d) and (e)–(h) show the coverage probabilities and average lengths, respectively, for the Fréchet distribution, while panels (i)–(l) and (m)–(p) show those for the Burr distribution. Each group of four panels is ordered by tail index value $\gamma \in \{0.25, 0.3, 0.35, 0.4\}$. For $n = 1,500$ and $\tau_n' = 1 - 1/n$, we consider the Gaussian 95% asymptotic confidence intervals associated with the direct BR and PS estimators $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ (ALS-BR) and $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (ALS-PS), their indirect counterparts $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$ (QB-BR) and $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (QB-PS), our direct GPML estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},1}_n)$ (ALS1-ML), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},2}_n)$ (ALS2-ML), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},3}_n)$ (ALS3-ML), the direct Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},1}_n)$ (ALS1-Mom), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},2}_n)$ (ALS2-Mom), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},3}_n)$ (ALS3-Mom), and the indirect GPML and Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \widehat{\sigma}^{\mathrm{ML}}_n)$ (QB-ML) and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n)$ (QB-Mom).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

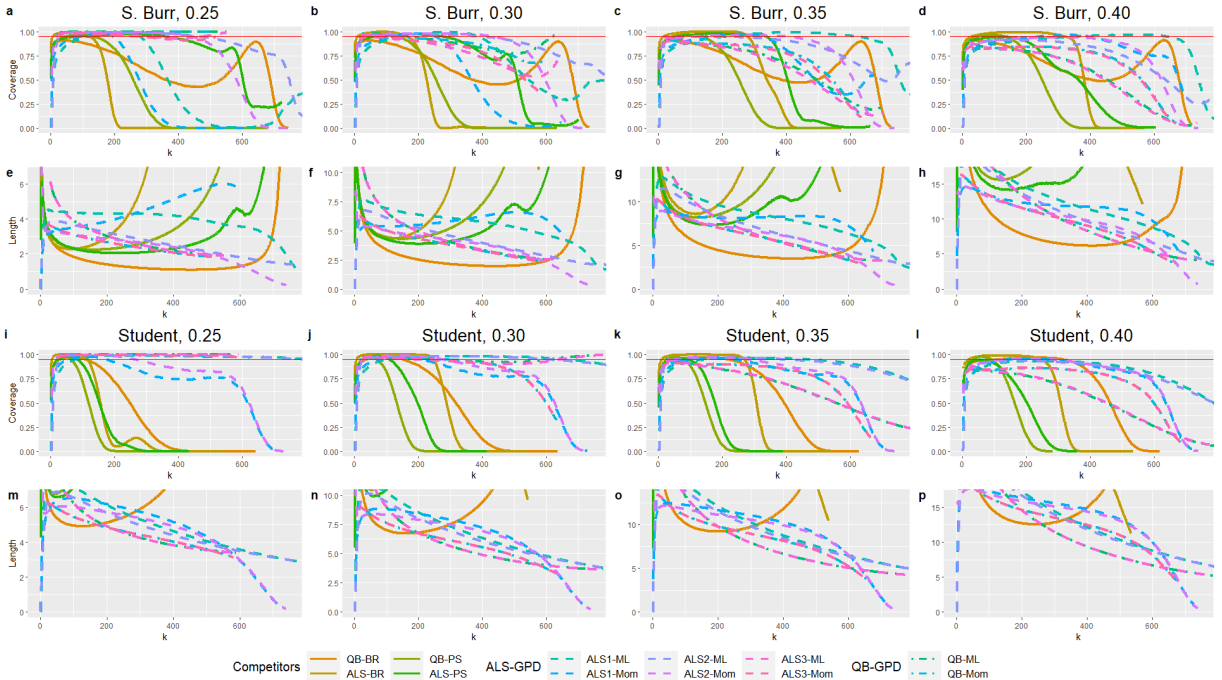


FIGURE 4. Simulation results for real-valued distributions. Panels (a)–(d) and (e)–(h) show the coverage probabilities and average lengths, respectively, for the symmetric Burr distribution, while panels (i)–(l) and (m)–(p) show those for the Student distribution. Each group of four panels is ordered by tail index value $\gamma \in \{0.25, 0.3, 0.35, 0.4\}$. For $n = 1,500$ and $\tau_n' = 1 - 1/n$, we consider the Gaussian 95% asymptotic confidence intervals associated with the direct BR and PS estimators $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ (ALS-BR) and $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (ALS-PS), their indirect counterparts $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$ (QB-BR) and $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (QB-PS), our direct GPML estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},1}_n)$ (ALS1-ML), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},2}_n)$ (ALS2-ML), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \breve{\sigma}^{\mathrm{ML},3}_n)$ (ALS3-ML), the direct Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},1}_n)$ (ALS1-Mom), $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},2}_n)$ (ALS2-Mom), and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \breve{\sigma}^{\mathrm{Mom},3}_n)$ (ALS3-Mom), and the indirect GPML and Moment estimators $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{ML}}_n, \widehat{\sigma}^{\mathrm{ML}}_n)$ (QB-ML) and $\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{Mom}}_n, \widehat{\sigma}^{\mathrm{Mom}}_n)$ (QB-Mom).

estimators $\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \check{\sigma}^{\mathrm{ML},1}_n)$ and $\check{\xi}^{\star}_{\tau_n}(\widehat{\gamma}^{\mathrm{ML}}_n, \check{\sigma}^{\mathrm{ML},2}_n)$ with coverage probabilities in dashed turquoise and dashed purple, respectively, achieve coverage levels that are reasonably close to the nominal level for both the positive and real-valued distributions, and for all studied values of γ; it should here be taken into account that, unlike the estimators $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$, the proposed GPML intervals do not benefit from extensive finite-sample corrections. And while these estimators fall short of the indirect bias-reduced $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$ estimator's tight interval lengths, they still post respectable results, especially around the higher k values where they record their most competitive RMSE and RBias in Figures 1 and 2. The larger average lengths of our confidence intervals can be attributed to the additional uncertainty that our estimation and inference procedures must account for, as they incorporate estimates of the location, scale, and shape parameters of the GP distribution, whereas the Weissman-type competitors are constructed on scale and shape parameter estimates only. Another notable outlier is the second direct Moment estimator $\check{\xi}^{\star}_{\tau_n}(\widehat{\gamma}^{\mathrm{Mom}}_n, \check{\sigma}^{\mathrm{Mom},2}_n)$ as it shows encouraging coverage, in dashed mauve, for all cases of the symmetric Burr distribution, as well as Student cases with $\gamma = 0.35$ and 0.4.

Finally, it should be noted that our experience with simulated data indicates that the chosen asymptotic confidence intervals $\check{I}^{\mathrm{ML},i}_{\tau'_n}$, $\check{I}^{\mathrm{Mom},i}_{\tau'_n}$, $\check{I}^{\mathrm{ML}}_{\tau'_n}$, and $\check{I}^{\mathrm{Mom}}_{\tau'_n}$, obtained from the asymptotics in Theorems 1 and 3, afford better results (for both nonnegative and real-valued distributions) than their respective counterparts that are derived from the equivalent convergences in (2.8) and (2.11), mainly due to the reliance of the former confidence intervals on additional location parameter estimates. This is illustrated for the positive and real-valued distributions in Figures 5 and 6, respectively, where we visualize the dominance of the chosen direct GPML confidence intervals $\check{I}^{\mathrm{ML},1}_{\tau'_n}$ and $\check{I}^{\mathrm{ML},2}_{\tau'_n}$, in terms of both coverage probabilities and average lengths, over their alternate confidence intervals that are constructed from the convergence (2.8).

4. APPLICATION TO DYNAMIC TAIL RISK FORECASTING

We apply our extreme expectile estimation methodologies to evaluate tail risk for the stocks of Netflix, Walmart, and American Express, as well as the exchange rate of Bitcoin (BTC-USD). Our goal is to forecast the expectile risk measure of tomorrow given the knowledge of today. To this end, we adopt a rolling window approach to our analysis, with each window size set to $n = 1,500$, and utilize data on the daily loss returns (i.e., negative log-returns) $Y_1, \ldots, Y_n$ of the aforementioned stocks and exchange rate. Using this setup and a careful selection of observation dates, we arrive at the following data specifications:

- Netflix: 498 estimation windows between 2008/12/03 and 2016/11/08.
- Walmart: 482 windows between 2013/11/13 and 2021/09/28.
- American Express: 766 windows between 2011/11/16 and 2020/11/17.
- Bitcoin-USD: 217 windows between 2016/04/28 and 2021/01/08.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

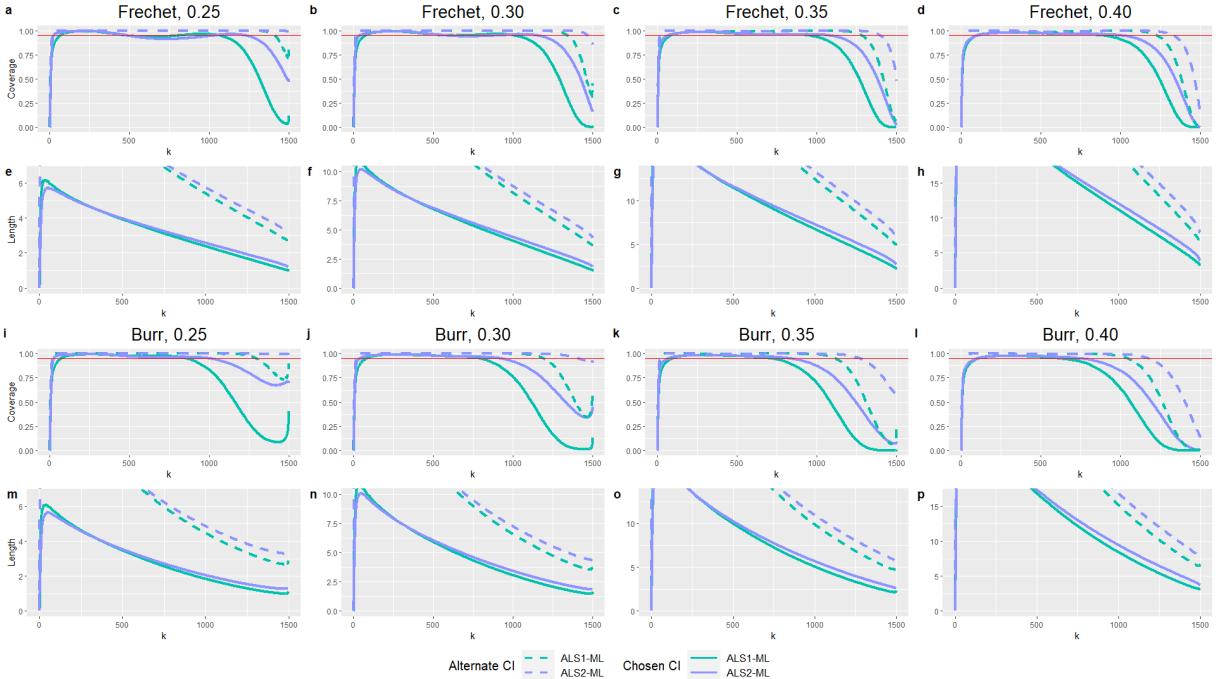


FIGURE 5. Simulation results for positive-valued distributions. Panels (a)–(d) and (e)–(h) show the coverage probabilities and average lengths, respectively, for the Fréchet distribution, while panels (i)–(l) and (m)–(p) show those for the Burr distribution. Each group of four panels is ordered by tail index value $\gamma \in \{0.25, 0.3, 0.35, 0.4\}$. For $n = 1,500$ and $\tau_n' = 1 - 1/n$, we consider the chosen direct GPML confidence intervals $\check{I}_{\tau_n'}^{\mathrm{ML},1}$ and $\check{I}_{\tau_n'}^{\mathrm{ML},2}$ (solid lines), constructed from the convergence in Theorem 1, and their alternate counterparts (dashed lines) that are constructed from the equivalent convergence (2.8).

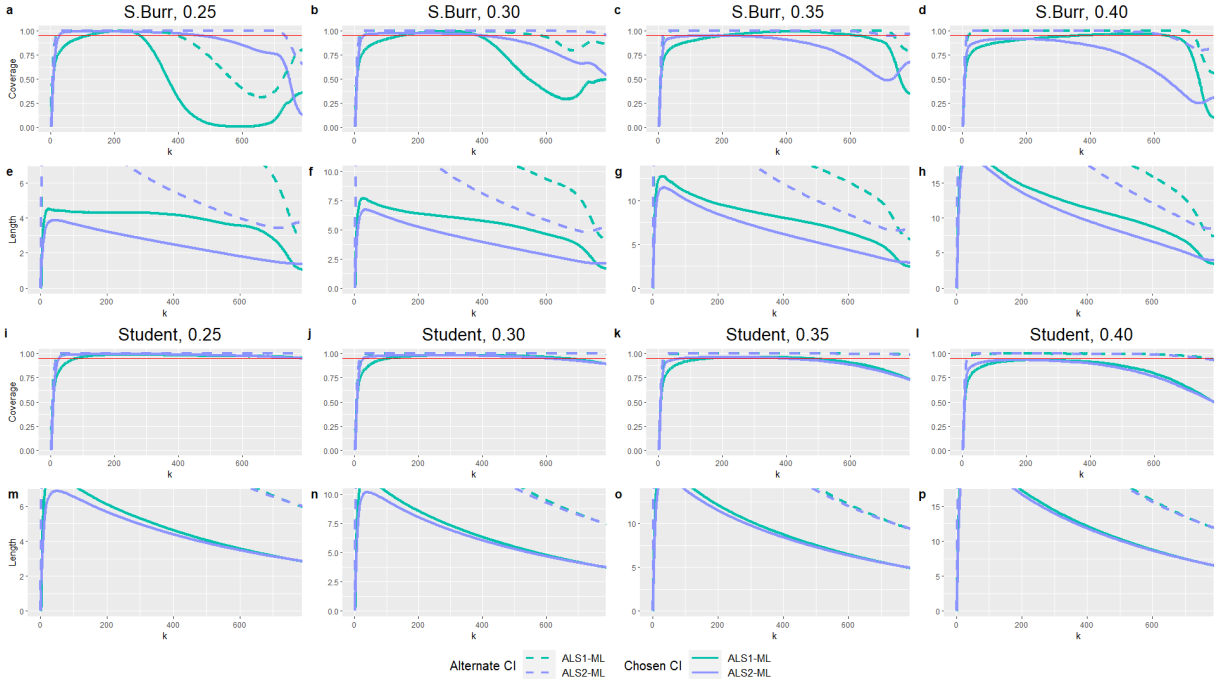


FIGURE 6. Simulation results for real-valued distributions. Panels (a)–(d) and (e)–(h) show the coverage probabilities and average lengths, respectively, for the symmetric Burr distribution, while panels (i)–(l) and (m)–(p) show those for the Student distribution. Each group of four panels is ordered by tail index value $\gamma \in \{0.25, 0.3, 0.35, 0.4\}$. For $n = 1,500$ and $\tau_n' = 1 - 1/n$, we consider the chosen direct GPML confidence intervals $\check{I}_{\tau_n'}^{\mathrm{ML},1}$ and $\check{I}_{\tau_n'}^{\mathrm{ML},2}$ (solid lines), constructed from the convergence in Theorem 1, and their alternate counterparts (dashed lines) that are constructed from the equivalent convergence (2.8).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

We follow the strategy laid out in Girard et al. (2022b) of filtering our time series through the lens of an ARMA(1,1)-GARCH(1,1) model. For a rolling window $Y_1, \ldots, Y_n$, we set up the model $Y_t = \mu + \phi Y_{t-1} + u_t + \theta u_{t-1}$, where $u_t = \sigma_t \varepsilon_t$ such that $\sigma_t^2 = c + a u_{t-1}^2 + b\sigma_{t-1}^2$ and (ε_t) is an unobserved independent white noise sequence, while the constants $\mu, \phi, \theta, a, b,$ and c are model parameters. Using the positive homogeneity and location equivariance of expectiles, we can then write the conditional τth expectile for the next day as

$$\xi_\tau(Y_{n+1}|\mathcal{F}_n) = \mu + \phi Y_n + \sigma_{n+1}\xi_\tau(\varepsilon) + \theta u_n,$$

where $\mathcal{F}_n$ is the σ-algebra generated by the ARMA-GARCH process up to time n.

We first estimate the ARMA-GARCH model parameters using the function `garchFit` in the `R` package `fGarch`, which allows us to retrieve the raw residuals $\widehat{u}_i$ through the `@residuals` option. Following the ideas of Girard et al. (2022b), we then fit a GARCH(1,1) model to $\widehat{u}_i$ and retrieve the predictions $\widehat{\varepsilon}_i$ of the innovations from the GARCH residuals. These filtered residuals can then be considered as independent and identically distributed copies of ε, which facilitates the estimation procedures laid out in this work. The extrapolated extreme expectile estimates

$$\begin{aligned}\overline{\xi}_{\tau_n'} \in \Big\{ &\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{E,BR}}), \widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{H,BR}}), \widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}, \widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}, \\ &\breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \breve{\sigma}_n^{\mathrm{ML},i}), \breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},i}), \breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{ML}}, \widehat{\sigma}_n^{\mathrm{ML}}), \breve{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}_n^{\mathrm{Mom}}, \widehat{\sigma}_n^{\mathrm{Mom}}) \Big\},\end{aligned} \tag{4.1}$$

for $i \in \{1,2,3\}$, paired with estimates of the ARMA-GARCH model parameters $\widehat{\mu}$, $\widehat{\phi}$, $\widehat{\theta}$, $\widehat{a}$, $\widehat{b}$, and $\widehat{c}$, can be combined to construct the dynamic predictions of the extreme expectiles of Y_{n+1} given Y_n as

$$\overline{\xi}_{\tau_n'}(Y_{n+1}|\mathcal{F}_n) = \widehat{\mu} + \widehat{\phi} Y_n + \widehat{\sigma}_{n+1}\,\overline{\xi}_{\tau_n'}(\varepsilon) + \widehat{\theta}\widehat{u}_n. \tag{4.2}$$

Recall that a functional T is elicitable relative to a class of distributions $\mathcal{F}$ if there exists a scoring function $S : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ such that $\mathbb{E}_F\{S(t,Y)\} \le \mathbb{E}_F\{S(x,Y)\}$ for every $F \in \mathcal{F}$, every $t \in T(F)$, and every $x \in \mathbb{R}$, with equality if and only if $x \in T(F)$ (Gneiting, 2011, Def. 2). This property provides a natural basis for evaluating and comparing the 12 competing forecasters. For the expectile at level τ_n', we use the asymmetric least squares scoring function $L_{\tau_n'} : \mathbb{R}^2 \to [0,\infty)$, $(\xi, y) \mapsto |\tau_n' - \mathbb{1}\{y \le \xi\}|(y-\xi)^2$, which gives the loss $L_{\tau_n'}(\xi, y)$ when the point forecast ξ is issued and the realization y of the future observation materializes. Following Ziegel (2016), given T forecast cases with point forecasts $(\xi_1^{(m)}, \ldots, \xi_T^{(m)})$ and realizing observations $(y_1, \ldots, y_T)$, where the index m numbers the competing forecasters, the latter can be ranked in terms of their average realized loss (the lower the better)

$$\mathcal{L}^{(m)}_{\tau_n'} = \frac{1}{T}\sum_{t=1}^{T} L_{\tau_n'}\left(\xi_t^{(m)}, y_t\right), \quad m = 1, \ldots, 12, \tag{4.3}$$

where each forecast case $t \in \{1, \ldots, T\}$ corresponds in our estimation context to a rolling window of loss returns $Y_1, \ldots, Y_n$, with y_t representing the realization of the future observation Y_{n+1}, and $\xi_t^{(m)}$ being its point forecast $\overline{\xi}_{\tau_n'}(Y_{n+1}|\mathcal{F}_n)$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

obtained in (4.2) from the 12 competitors $\overline{\xi}_{\tau_n'}(\varepsilon)$ in (4.1). In our assessment of tail risk, we compare the average scores $\mathcal{L}_{\tau_n'}^{(m)}$ for a fine grid of challenging levels $\tau_n' \in \{0.99, \ldots, 1 - 1/n\}$, where data are sparse and inference is difficult to handle. To guide our choice of the intermediate sequence k_n, which plays the role of the tuning parameter in our setup, we calculate the average realized loss for every value k in the range of all possible values of k_n, then select the value that yields the lowest loss for each estimator and each value of τ_n'. This approach simplifies the optimal k_n selection process while allowing us to compare the performance of each forecaster fairly. Figures 7–10 show the results of our analysis of daily loss returns of Netflix, Walmart, American Express, and Bitcoin-USD, respectively.

In each figure, Panel (A) showcases the observed data and wraps the realizations y_t between two dashed red lines. Panel (B) shows the results of the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) test for each rolling window, illustrating the stationarity of our loss returns over all chosen time periods, which is required for ARMA-GARCH modeling. Panel (C) then examines the independence of the residuals $\widehat{\varepsilon}_i$ using a Ljung–Box test on these residuals and their squares. This independence is a prerequisite for conducting inference on extreme expectiles from our Corollaries 1 and 2 and the asymptotic theory of the Weissman-type competitors. The residuals' tail heaviness is evaluated in Panel (D) through a study of the GPML ($\widehat{\gamma}_n^{\mathrm{ML}}$) and Moment ($\widehat{\gamma}_n^{\mathrm{Mom}}$) estimates of the EVI against an appropriate range of possible values k of the intermediate level k_n. The results from each rolling window are drawn as an individual curve against k and indicate that the residuals exhibit consistent tail heaviness over all estimation windows for a wide range of k values. Panel (E) then displays the average realized loss $\mathcal{L}_{\tau_n'}^{(m)}$, defined in (4.3), for each estimator against τ_n', using the optimal k_n value that minimizes said loss. The good news for practitioners concerned with the accuracy of daily forecasts is that our GPD-based forecasters outperform all Weissman-type competitors, for each studied financial time series. More specifically, the direct Moment estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},1})$ establishes a clear lead for Netflix, Walmart, and Bitcoin-USD data, with estimators like the second direct Moment estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},2})$, the third direct Moment estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},3})$, and the indirect Moment estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \widehat{\sigma}_n^{\mathrm{Mom}})$ trailing close behind. The American Express data see the indirect Moment estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \widehat{\sigma}_n^{\mathrm{Mom}})$, respectively, the direct GPML estimator $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{ML}}, \breve{\sigma}_n^{\mathrm{ML},2})$, outperform all of the Weissman competitors for $\tau_n' < 0.995$, respectively, $\tau_n' \geq 0.995$. The results imply that our Moment- and GPML-based estimators are suitable for financial data, which corroborates our simulation findings in Section 3, where said estimators scored ahead of the competition in the case of real-valued distributions. The final panel (F) wraps up by showing next-day forecasts of $\xi_{\tau_n'}$ at level $\tau_n' = 0.99$, for the daily loss returns over the observation period, using the dynamic predictions (4.2) of the top-ranked forecaster $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},1})$ for Netflix, Walmart, and Bitcoin-USD, and $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \widehat{\sigma}_n^{\mathrm{Mom}})$ for American Express, along with their 95%

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

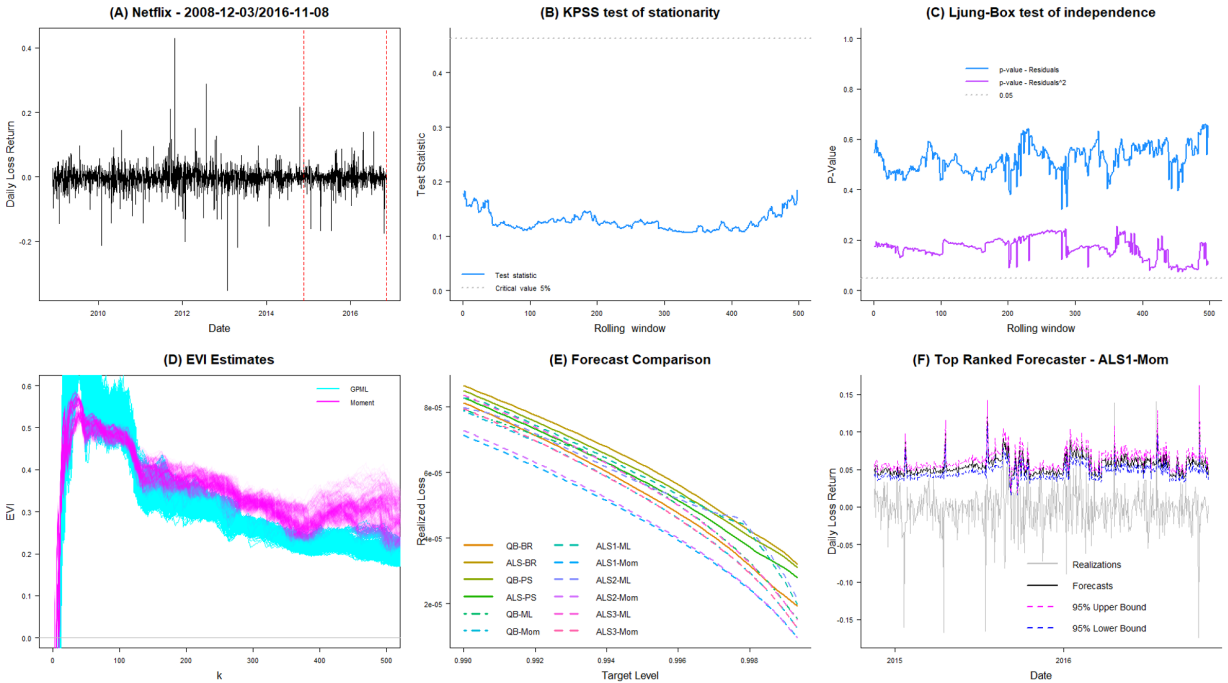


FIGURE 7. Netflix data. (A) Observed negative log returns, with realizations wrapped between dashed red lines. (B) KPSS test statistic of stationarity (in blue) with the 5% critical value in dotted gray, both against the rolling window index. (C) p-values of the Ljung–Box test of independence, with 30 lags, for the residuals $\widehat{\varepsilon}_i$ and their squares $\widehat{\varepsilon}_i^2$ against the rolling window index. (D) Collage of $\widehat{\gamma}_n^{\mathrm{ML}}$ and $\widehat{\gamma}_n^{\mathrm{Mom}}$ against k; each curve corresponds to a distinct rolling window. (E) Comparison of the competing estimators on the basis of realized loss $\mathcal{L}_{\tau_n'}^{(m)}$ for $\tau_n' \in \{0.99, \ldots, 1 - 1/n\}$ and $m = 1, \ldots, 12$. (F) Next-day forecasts (black curve) $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},1})$ of $\xi_{\tau_n'}$ at level $\tau_n' = 0.99$, for the daily loss returns computed sequentially over each rolling window, with 95% lower (blue curve) and upper (magenta curve) asymptotic confidence bounds, along with the realization of the future observation (gray curve).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

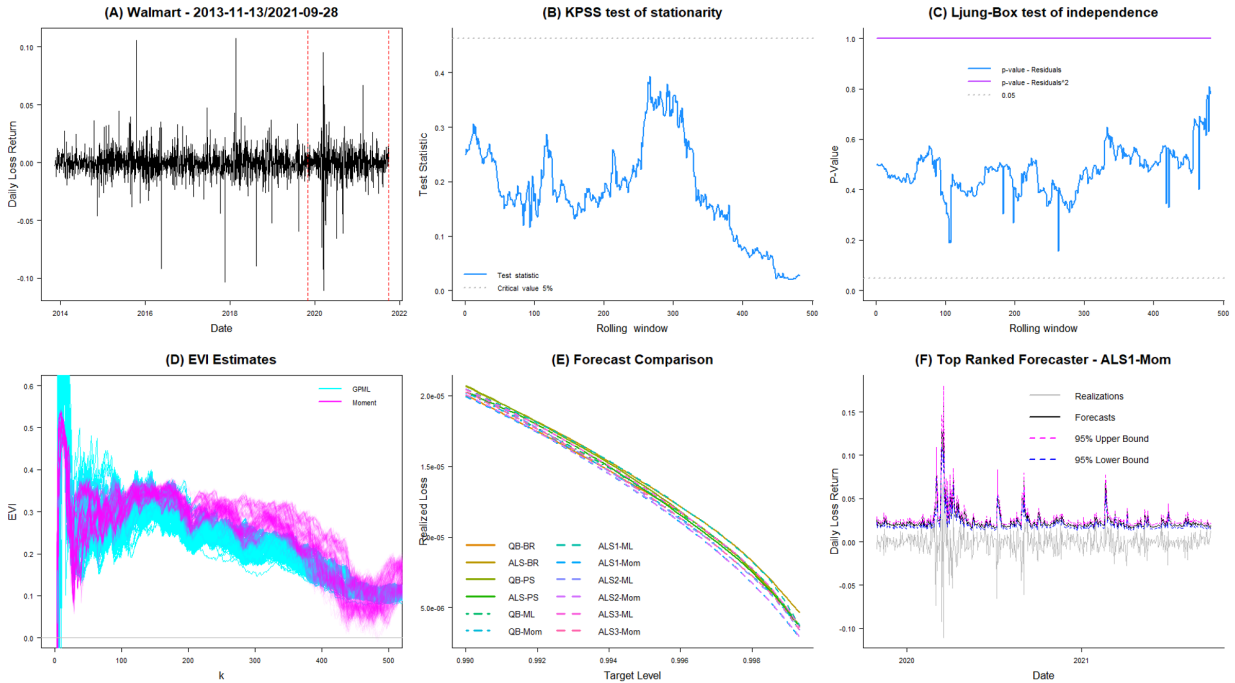


FIGURE 8. Walmart data. (A) Observed negative log returns, with realizations wrapped between dashed red lines. (B) KPSS test statistic of stationarity (in blue) with the 5% critical value in dotted gray, both against the rolling window index. (C) p-values of the Ljung–Box test of independence, with 30 lags, for the residuals $\widehat{\varepsilon}_i$ and their squares $\widehat{\varepsilon}_i^2$ against the rolling window index. (D) Collage of $\widehat{\gamma}_n^{\mathrm{ML}}$ and $\widehat{\gamma}_n^{\mathrm{Mom}}$ against k; each curve corresponds to a distinct rolling window. (E) Comparison of the competing estimators on the basis of realized loss $\mathcal{L}_{\tau_n'}^{(m)}$ for $\tau_n' \in \{0.99, \ldots, 1 - 1/n\}$ and $m = 1, \ldots, 12$. (F) Next-day forecasts (black curve) $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},1})$ of $\xi_{\tau_n'}$ at level $\tau_n' = 0.99$, for the daily loss returns computed sequentially over each rolling window, with 95% lower (blue curve) and upper (magenta curve) asymptotic confidence bounds, along with the realization of the future observation (gray curve).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

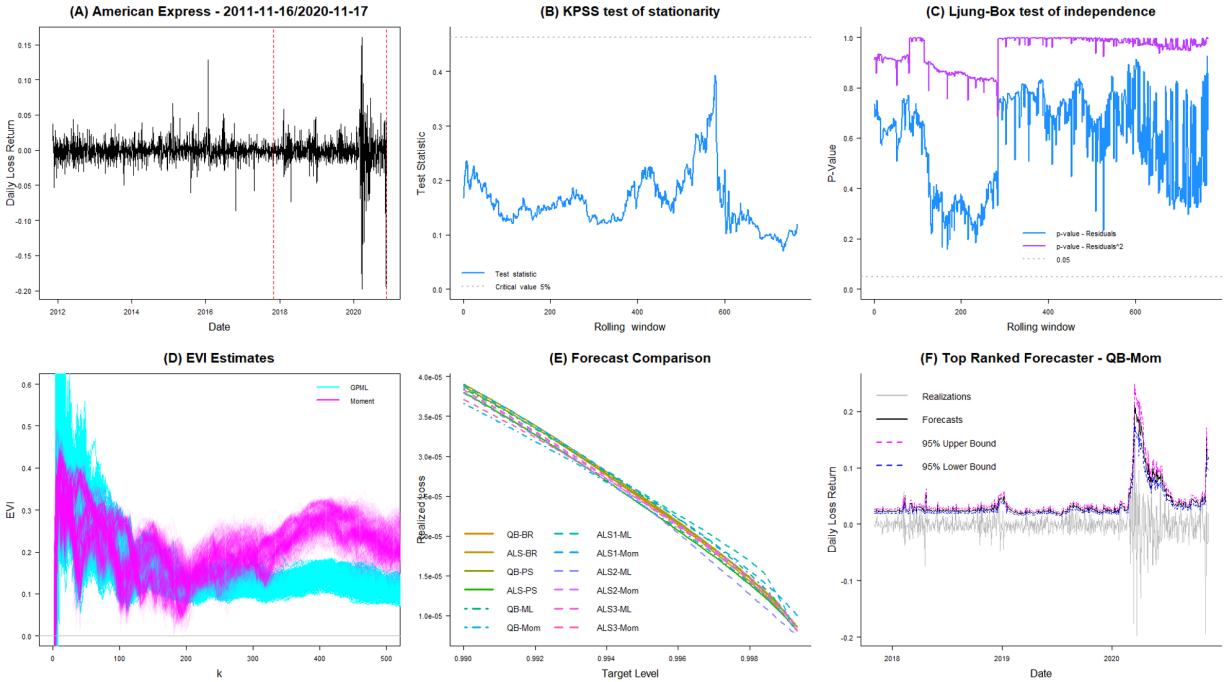


FIGURE 9. American Express data. (A) Observed negative log returns, with realizations wrapped between dashed red lines. (B) KPSS test statistic of stationarity (in blue) with the 5% critical value in dotted gray, both against the rolling window index. (C) p-values of the Ljung–Box test of independence, with 30 lags, for the residuals $\widehat{\varepsilon}_i$ and their squares $\widehat{\varepsilon}_i^2$ against the rolling window index. (D) Collage of $\widehat{\gamma}_n^{\mathrm{ML}}$ and $\widehat{\gamma}_n^{\mathrm{Mom}}$ against k; each curve corresponds to a distinct rolling window. (E) Comparison of the competing estimators on the basis of realized loss $\mathcal{L}_{\tau_n'}^{(m)}$ for $\tau_n' \in \{0.99, \ldots, 1 - 1/n\}$ and $m = 1, \ldots, 12$. (F) Next-day forecasts (black curve) $\check{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \widehat{\sigma}_n^{\mathrm{Mom}})$ at level $\tau_n' = 0.99$, for the daily loss returns computed sequentially over each rolling window, with 95% lower (blue curve) and upper (magenta curve) asymptotic confidence bounds, along with the realization of the future observation (gray curve).

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

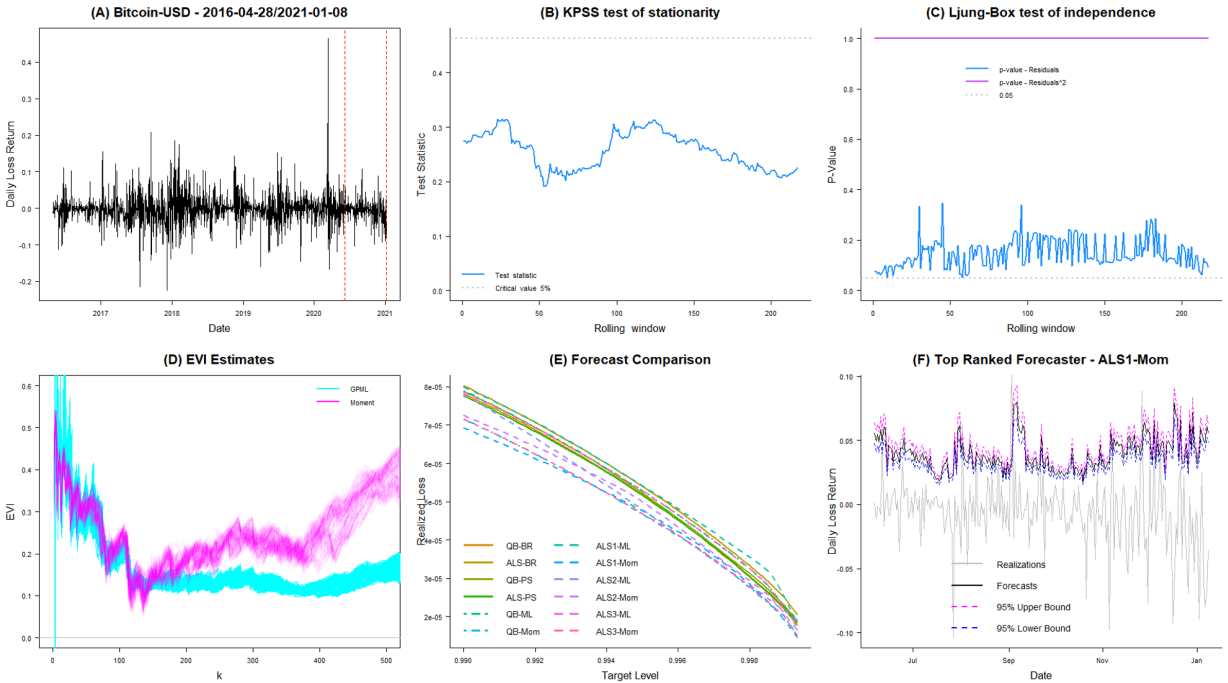


FIGURE 10. Bitcoin-USD data. (A) Observed negative log returns, with realizations wrapped between dashed red lines. (B) KPSS test statistic of stationarity (in blue) with the 5% critical value in dotted gray, both against the rolling window index. (C) p-values of the Ljung–Box test of independence, with 30 lags, for the residuals $\widehat{\varepsilon}_i$ and their squares $\widehat{\varepsilon}_i^2$ against the rolling window index. (D) Collage of $\widehat{\gamma}_n^{\mathrm{ML}}$ and $\widehat{\gamma}_n^{\mathrm{Mom}}$ against k; each curve corresponds to a distinct rolling window. (E) Comparison of the competing estimators on the basis of realized loss $\mathcal{L}_{\tau_n'}^{(m)}$ for $\tau_n' \in \{0.99, \ldots, 1 - 1/n\}$ and $m = 1, \ldots, 12$. (F) Next-day forecasts (black curve) $\breve{\xi}_{\tau_n'}^{\star}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom},1})$ of $\xi_{\tau_n'}$ at level $\tau_n' = 0.99$, for the daily loss returns computed sequentially over each rolling window, with 95% lower (blue curve) and upper (magenta curve) asymptotic confidence bounds, along with the realization of the future observation (gray curve).

asymptotic confidence intervals and the realization of the future observation. The point forecasts and associated confidence bounds seem to follow data trends fairly closely.

5. DISCUSSION

Tail expectiles define a prime candidate for a standard risk measure in finance and insurance thanks to their excellent axiomatic properties and their ability to capture essential information about the severity of extreme observations as well as their probabilities. The problem of correctly estimating and inferring extreme expectile risk is a difficult question due to the least squares nature of expectiles and their sensitivity to tail heaviness, even though a series of recent papers has focused on this problem using Weissman-type extrapolation devices. Our Generalized Pareto approach provides good estimates of tail risk across the board, with reasonably accurate (though elementary) asymptotic Gaussian confidence intervals, in the financial context of heavy-tailed real-valued profit–loss distributions without resorting to any bias correction of our estimators, even though their asymptotic behavior indicates that they are asymptotically biased. The results of the tail risk forecasting exercise we have carried out similarly indicate that, in a dynamic estimation setting, the Generalized Pareto extrapolation method yields encouraging results.

There still remains a lot to be done, especially in the actuarial context of nonnegative heavy-tailed loss variables for which Weissman-type methods seem to outperform our GPD-based approach. As we have highlighted, asymptotic Gaussian inference of extreme expectiles may not perform well in finite samples, because of the difficulty of tracking the statistical uncertainties of the estimated high expectiles and the use of the asymptotic variances arising in the Gaussian limiting distributions. The finite-sample approximations and errors thus made are typically due to (i) the use of the GPD model when no parametric assumption about the exceedance distribution is desired, (ii) the asymptotic approximations motivating the extreme value estimators of the scale and shape parameters, (iii) the use of the asymptotic connection between extreme expectiles and quantiles while ignoring higher-order error terms, (iv) incorrectly neglecting correlations between estimators when the asymptotic behavior of one of them dominates, and/or (v) not accounting for the variance distortions incurred by utilizing the delta method for linearization purposes. Improving our results by providing successive corrections for each of these types of approximation errors is a topic of interest for future research.

APPENDIXES

The Appendix has three parts. Appendix A presents an overview of the current state of research on tail expectile estimation. Appendix B provides further details about bias

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

reduction and variance correction within the Weissman extrapolation framework. Appendix C contains the proofs of all theoretical results in the main article.

A. WEISSMAN-TYPE ESTIMATORS

Let $X_1, \ldots, X_n$ be a sample of observations drawn from F, and for the moment assume that these random copies of X are independent. Let $X_{1,n} \leqslant \cdots \leqslant X_{n,n}$ be their corresponding order statistics. In the intermediate case when $\tau = \tau_n \to 1$ such that $n(1-\tau_n) \to \infty$ as $n \to \infty$, or under the more common discretized setup $\tau_n = 1 - k/n$ for a sequence of integers $k = k_n \to \infty$ with $k_n/n \to 0$ as $n \to \infty$, the population expectile ξ_{τ_n} is extreme but well inside the sample so that it can be consistently estimated by its empirical version

$$\widehat{\xi}_{\tau_n} = \arg\min_{u \in \mathbb{R}} \frac{1}{n} \sum_{i=1}^{n} \eta_{\tau_n}(X_i - u), \tag{A.1}$$

where $\eta_\tau(y) = |\tau - \mathbb{1}\{y \leqslant 0\}|y^2$. The asymptotic normality of $\widehat{\xi}_{\tau_n}$ is derived in Theorem 2 in Daouia et al. (2018) for the class of heavy-tailed distributions (2.2) with $\gamma < 1/2$. In the more general setting where $0 < \gamma < 1$, an indirect asymptotically normal estimator of ξ_{τ_n} can be obtained from the asymptotic connection (2.5), or equivalently, $\xi_{\tau_n} \sim (\gamma^{-1} - 1)^{-\gamma} q_{\tau_n}$ as $n \to \infty$. By substituting in a suitable estimator $\overline{\gamma}$ in place of γ and the empirical quantile $\widehat{q}_{\tau_n} = X_{n-\lfloor n(1-\tau_n) \rfloor, n} = X_{n-k,n}$ in place of q_{τ_n}, the resulting estimator

$$\widetilde{\xi}_{\tau_n} = (\overline{\gamma}^{-1} - 1)^{-\overline{\gamma}} \widehat{q}_{\tau_n} \tag{A.2}$$

is shown to have a $\sqrt{n(1-\tau_n)}$-asymptotic nondegenerate distribution under the conditions formulated in Theorem 1 in Daouia et al. (2018). The EVI $\gamma \in (0,1)$ can be estimated by the celebrated Hill estimator (Hill, 1975) $\widehat{\gamma}_n^{\mathrm{H}} := \frac{1}{\lfloor n(1-\tau_n) \rfloor} \sum_{i=1}^{\lfloor n(1-\tau_n) \rfloor} \log X_{n-i+1,n} - \log X_{n-\lfloor n(1-\tau_n) \rfloor, n}$, or, in the case $\gamma \in (0, 1/2)$, by the following expectile-based estimator introduced and popularized by Girard et al. (2022a, 2022b):

$$\widehat{\gamma}_n^{\mathrm{E}} := \left(1 + \frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n} \right)^{-1} \quad \text{where } \widehat{\overline{F}}_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}\{X_i > x\}. \tag{A.3}$$

In the more challenging extreme case when $\tau = \tau_n' \to 1$ such that $n(1-\tau_n') \to c \in [0, \infty)$ as $n \to \infty$, the population expectile $\xi_{\tau_n'}$ lies at the far tail of the underlying distribution, possibly beyond the sample maximum $X_{n,n}$ in the data. Typical examples include $\tau_n' = 1 - 1/n$, with the associated extreme expectile $\xi_{\tau_n'}$ having the order of magnitude of $X_{n,n}$. In this case, the model assumption of Pareto-type tail (2.2), combined with the asymptotic connection (2.5), yields the Weissman approximation

$$\frac{q_{\tau_n'}}{q_{\tau_n}} = \frac{U((1-\tau_n')^{-1})}{U((1-\tau_n)^{-1})} \approx \left(\frac{1-\tau_n'}{1-\tau_n} \right)^{-\gamma} \quad \text{and hence} \quad \frac{\xi_{\tau_n'}}{\xi_{\tau_n}} \sim \frac{q_{\tau_n'}}{q_{\tau_n}} \approx \left(\frac{1-\tau_n'}{1-\tau_n} \right)^{-\gamma}. \tag{A.4}$$

This approximation motivates the following class of plug-in estimators of $\xi_{\tau_n'}$:

$$\overline{\xi}^{\star}_{\tau_n'} \equiv \overline{\xi}^{\star}_{\tau_n'}(\tau_n) := \left(\frac{1-\tau_n'}{1-\tau_n} \right)^{-\overline{\gamma}} \overline{\xi}_{\tau_n}, \tag{A.5}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

where $\overline{\gamma}$ is a $\sqrt{n(1-\tau_n)}$-consistent estimator of γ, and $\overline{\xi}_{\tau_n}$ is either the empirical asymmetric least squares estimator $\widehat{\xi}_{\tau_n}$ in (A.1) or the quantile-based estimator $\widetilde{\xi}_{\tau_n}$ in (A.2) for the intermediate expectile ξ_{τ_n}. More specifically, the use of the latter indirect intermediate expectile $\overline{\xi}_{\tau_n} \equiv \widetilde{\xi}_{\tau_n}$ in (A.5) yields the extreme expectile estimator $\widetilde{\xi}^{\star}_{\tau'_n} := \overline{\xi}^{\star}_{\tau'_n}$ defined as

$$\widetilde{\xi}^{\star}_{\tau'_n} = \widetilde{\xi}^{\star}_{\tau'_n}(\overline{\gamma}) := \left(\frac{1-\tau'_n}{1-\tau_n}\right)^{-\overline{\gamma}} \widetilde{\xi}_{\tau_n} = \left(\frac{1-\tau_n}{1-\tau'_n}\right)^{\overline{\gamma}} (\overline{\gamma}^{-1}-1)^{-\overline{\gamma}} \widehat{q}_{\tau_n}, \tag{A.6}$$

while the choice of the former direct intermediate expectile $\overline{\xi}_{\tau_n} \equiv \widehat{\xi}_{\tau_n}$ results in

$$\widehat{\xi}^{\star}_{\tau'_n} = \widehat{\xi}^{\star}_{\tau'_n}(\overline{\gamma}) := \left(\frac{1-\tau'_n}{1-\tau_n}\right)^{-\overline{\gamma}} \widehat{\xi}_{\tau_n}. \tag{A.7}$$

Based on an i.i.d. assumption on the sample and an SORV condition $\mathcal{C}_2(\gamma, \tilde{\rho}, \tilde{A})$ with additional mild regularity conditions, Daouia et al. (2018) have shown in their Corollaries 3 and 4 that the estimators (A.6) and (A.7) have the same limit distribution as the chosen tail index estimator $\overline{\gamma}$ with the slightly slower rate of convergence $\log((1-\tau_n)/(1-\tau'_n))/\sqrt{n(1-\tau_n)}$. In the rest of this Appendix, for the sake of notational convenience, we set $\rho = \tilde{\rho}$ and $A = \tilde{A}$, since there is no possible confusion between parameter values in conditions $\mathcal{C}_2$ and $\mathcal{E}_2$.

When using the extrapolated estimator $\widetilde{\xi}^{\star}_{\tau'_n}(\overline{\gamma})$ in (A.6) in conjunction with the Hill estimator $\overline{\gamma} = \widehat{\gamma}^{\mathrm{H}}_n$, and its asymmetric least squares competitor $\widehat{\xi}^{\star}_{\tau'_n}(\overline{\gamma})$ in (A.7) along with the expectile-based estimator $\overline{\gamma} = \widehat{\gamma}^{\mathrm{E}}_n$, it was shown in Daouia et al. (2018) and Girard et al. (2022b) that, if $\rho < 0$ and under the conditions $\lambda_1 = \lim_{n\to\infty} \sqrt{n(1-\tau_n)} A((1-\tau_n)^{-1}) \in \mathbb{R}$, $\lambda_2 = \lim_{n\to\infty} \sqrt{n(1-\tau_n)}/q_{\tau_n} \in \mathbb{R}$ and $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau'_n)) \to \infty$, one has

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau'_n))} \log\left(\frac{\widetilde{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H}}_n)}{\xi_{\tau'_n}}\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(\frac{\lambda_1}{1-\rho}, \gamma^2\right) \tag{A.8}$$

and, when $\mathbb{E}|X_-|^2 < \infty$ and $\gamma < 1/2$,

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau'_n))} \log\left(\frac{\widehat{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{E}}_n)}{\xi_{\tau'_n}}\right)$$
$$\xrightarrow{\mathrm{d}} \mathcal{N}\left(\frac{\gamma(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}\lambda_1 + \gamma^2(\gamma^{-1}-1)^{\gamma+1}\mathbb{E}(X)\lambda_2, \frac{\gamma^3(1-\gamma)}{1-2\gamma}\right). \tag{A.9}$$

Note that when $(X_n, n = 1, 2, \ldots)$ is a strictly stationary, β-mixing time series and with the Hill estimator $\overline{\gamma} = \widehat{\gamma}^{\mathrm{H}}_n$, Davison et al. (2023) have recently proved in their Theorem 3.5 that these estimators, namely, $\widetilde{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H}}_n)$ and $\widehat{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H}}_n)$, are still asymptotically Gaussian with the same rate of convergence, but at the cost of an increased asymptotic variance.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

When ignoring the asymptotic bias, under the assumption $\lambda_1 = \lambda_2 = 0$, the resulting Gaussian $100(1-\alpha)\%$ asymptotic confidence intervals for $\xi_{\tau_n'}$ are then

$$\widetilde{I}^{(0)}_{\tau_n'}(\alpha) = \left[\widetilde{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H}}_n)\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}}\sqrt{\widetilde{\sigma}^2_n}\times z_{1-\alpha/2}\right)\right]$$

$$\text{and } \widehat{I}^{(0)}_{\tau_n'}(\alpha) = \left[\widehat{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E}}_n)\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}}\sqrt{\widehat{s}^2_n}\times z_{1-\alpha/2}\right)\right],$$

$$\text{where } \widetilde{\sigma}^2_n = (\widehat{\gamma}^{\mathrm{H}}_n)^2 \text{ and } \widehat{s}^2_n = \frac{(\widehat{\gamma}^{\mathrm{E}}_n)^3(1-\widehat{\gamma}^{\mathrm{E}}_n)}{1-2\widehat{\gamma}^{\mathrm{E}}_n},$$

with $z_{1-\alpha/2}$ being the quantile of level $1-\alpha/2$ for the standard Gaussian distribution. Bias-reduced versions of $\widetilde{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H}}_n)$ in (A.8) and $\widehat{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E}}_n)$ in (A.9) have been recently suggested by Girard et al. (2022b) when the auxiliary function A in their SORV condition $\mathcal{C}_2(\gamma,\rho,A)$ takes the form $A(t) = b\gamma t^{\rho}$ for certain constants $b \neq 0$ and $\rho < 0$. This function A can be estimated by using a consistent tail index estimator $\overline{\gamma}$ and the second-order parameter estimators $\overline{b}$ and $\overline{\rho}$ that were introduced in Gomes and Martins (2002) and Fraga Alves, Gomes, and de Haan (2003). A careful consideration of the different sources of bias in the construction of $\widetilde{\xi}^{\star}_{\tau_n'}$ and $\widehat{\xi}^{\star}_{\tau_n'}$ (see Appendix B) leads to bias-reduced versions $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$ and $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$, whose full expressions are given below in (B.5) and (B.6), and with (see Caeiro, Gomes, and Pestana, 2005)

$$\widehat{\gamma}^{\mathrm{H,BR}}_n = \widehat{\gamma}^{\mathrm{H}}_n\left(1-\frac{\overline{b}}{1-\overline{\rho}}\left(\frac{1}{1-\tau_n}\right)^{\overline{\rho}}\right) \tag{A.10}$$

and (see Girard et al., 2022a)

$$\widehat{\gamma}^{\mathrm{E,BR}}_n = \left(1+\frac{\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n})}{1-\tau_n}\left(1-\frac{\overline{X}_n}{\widehat{\xi}_{\tau_n}}\right)^{-1}(2\tau_n-1)\left(1+\frac{\overline{b}(\widehat{\overline{F}}_n(\widehat{\xi}_{\tau_n}))^{-\overline{\rho}}}{1-\widehat{\gamma}^{\mathrm{E}}_n-\overline{\rho}}\right)\right)^{-1}. \tag{A.11}$$

Gaussian $100(1-\alpha)\%$ asymptotic confidence intervals for $\xi_{\tau_n'}$ based on the purely asymmetric least squares estimator $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ and its quantile-based competitor $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$, also available as part of the R package `Expectrem`, are discussed in Daouia et al. (2024b), with an alternative solution put forward in Padoan and Stupfler (2022) to correct the naive confidence intervals $\widehat{I}^{(0)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(0)}_{\tau_n'}(\alpha)$. We list in Appendix B the intervals constructed from Daouia et al. (2024b) as $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$, $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$, $\widehat{I}^{(2)}_{\tau_n'}(\alpha)$, and $\widetilde{I}^{(2)}_{\tau_n'}(\alpha)$, and from Padoan and Stupfler (2022) as $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$. Appendix B also provides more details about these intervals and their observed finite-sample behavior.

B. BIAS-REDUCED ESTIMATORS AND INFERENCE

The asymptotic distributions of $\widetilde{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{H}}_n)$ in (A.8) and $\widehat{\xi}^{\star}_{\tau_n'}(\widehat{\gamma}^{\mathrm{E}}_n)$ in (A.9) feature bias components due to the semiparametric heavy tail framework. Bias-reduced versions of these estimators have been recently suggested by Girard et al. (2022b) under the i.i.d. assumption in the particular case where the auxiliary function $\tilde{A}$ in their SORV condition $\mathcal{C}_2(\gamma,\tilde{\rho},\tilde{A})$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

takes the form $\tilde{A}(t) = b\gamma t^{\tilde{\rho}}$ for certain constants $b \neq 0$ and $\tilde{\rho} < 0$. Again, in the rest of this Appendix, we set $\rho = \tilde{\rho}$ and $A = \tilde{A}$ for the sake of notational convenience.

This function A can be estimated by using a consistent tail index estimator $\overline{\gamma}$ and the second-order parameter estimators $\overline{b}$ and $\overline{\rho}$ that were introduced in Gomes and Martins (2002) and Fraga Alves et al. (2003) and can directly be calculated from the R package `evt0`. Supplement A of Girard et al. (2022b) provides a nice and comprehensive summary of how these estimators of b and ρ are constructed. As shown by Girard et al. (2022b), the direct and indirect estimators $\widehat{\xi}^{\star}_{\tau'_n}$ and $\widetilde{\xi}^{\star}_{\tau'_n}$ satisfy

$$\log\left(\frac{\widehat{\xi}^{\star}_{\tau'_n}}{\xi_{\tau'_n}}\right) = (\overline{\gamma}-\gamma)\log\left(\frac{1-\tau_n}{1-\tau'_n}\right) + \log\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}\right) - \log\left(\left(\frac{1-\tau'_n}{1-\tau_n}\right)^{\gamma}\frac{\xi_{\tau'_n}}{\xi_{\tau_n}}\right), \tag{B.1}$$

$$\text{and } \log\left(\frac{\widetilde{\xi}^{\star}_{\tau'_n}}{\xi_{\tau'_n}}\right) = (\overline{\gamma}-\gamma)\log\left(\frac{1-\tau_n}{1-\tau'_n}\right) + \log\left(\frac{(\overline{\gamma}^{-1}-1)^{-\overline{\gamma}}}{(\gamma^{-1}-1)^{-\gamma}}\right) + \log\left(\frac{\widehat{q}_{\tau_n}}{q_{\tau_n}}\right) \tag{B.2}$$

$$- \log\left(\left(\frac{1-\tau'_n}{1-\tau_n}\right)^{\gamma}\left(\gamma^{-1}-1\right)^{\gamma}\frac{\xi_{\tau'_n}}{q_{\tau_n}}\right).$$

To deal with the nonrandom bias term in (B.1), they have also shown that

$$\left(\frac{1-\tau'_n}{1-\tau_n}\right)^{\gamma}\frac{\xi_{\tau'_n}}{\xi_{\tau_n}} = \underbrace{\left(\frac{1-\tau'_n}{1-\tau_n}\right)^{\gamma}\frac{q_{\tau'_n}}{q_{\tau_n}}}_{1+B_{1,n}}\underbrace{\left(\gamma^{-1}-1\right)^{-\gamma}\frac{q_{\tau_n}}{\xi_{\tau_n}}}_{1+B_{2,n}}\underbrace{\left(\gamma^{-1}-1\right)^{\gamma}\frac{\xi_{\tau'_n}}{q_{\tau'_n}}}_{1+B_{3,n}} \tag{B.3}$$

for obvious definitions of $B_{1,n}, B_{2,n}$, and $B_{3,n}$. As for the extrapolation bias in (B.2), they have proposed the following correction:

$$\left(\frac{1-\tau'_n}{1-\tau_n}\right)^{\gamma}\left(\gamma^{-1}-1\right)^{\gamma}\frac{\xi_{\tau'_n}}{q_{\tau_n}} = \left(1+B_{1,n}\right)\left(1+B_{3,n}\right). \tag{B.4}$$

The terms $B_{1,n}, B_{2,n}$, and $B_{3,n}$ can be estimated by

$$\overline{B}_{1,n} = \frac{((1-\tau'_n)/(1-\tau_n))^{-\overline{\rho}}-1}{\overline{\rho}}\,\overline{b}\,\overline{\gamma}(1-\tau_n)^{-\overline{\rho}},$$

$$\overline{B}_{2,n} = (1+\overline{r}(\tau_n))^{\overline{\gamma}}\left(1+\frac{\frac{(\overline{\gamma}^{-1}-1)^{-\overline{\rho}}}{(1+\overline{r}(\tau_n))^{\overline{\rho}}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\tau_n)^{-\overline{\rho}}\right)^{-1}-1$$

$$\text{and } \overline{B}_{3,n} = (1+\overline{r}^{\star}(\tau'_n))^{-\overline{\gamma}}\left(1+\frac{\frac{(\overline{\gamma}^{-1}-1)^{-\overline{\rho}}}{(1+\overline{r}^{\star}(\tau'_n))^{\overline{\rho}}}-1}{\overline{\rho}}\overline{b}\overline{\gamma}(1-\tau'_n)^{-\overline{\rho}}\right)-1,$$

$$\text{where } 1+\overline{r}(\tau_n) = \left(1-\frac{\overline{X}_n}{\overline{\xi}_{\tau_n}}\right)\frac{1}{2\tau_n-1}\left(1+\frac{\overline{b}(\widehat{\overline{F}}_n(\overline{\xi}_{\tau_n}))^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}\right)^{-1}$$

$$\text{and } 1+\overline{r}^{\star}(\tau'_n) = \left(1-\frac{\overline{X}_n}{\overline{\xi}^{\star}_{\tau'_n}}\right)\frac{1}{2\tau'_n-1}\left(1+\frac{\overline{b}\left(\overline{\gamma}^{-1}-1\right)^{-\overline{\rho}}}{1-\overline{\gamma}-\overline{\rho}}(1-\tau'_n)^{-\overline{\rho}}\right)^{-1},$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

with $\overline{X}_n$ being the sample mean, $\overline{\xi}_{\tau_n}$ either the direct intermediate estimator $\widehat{\xi}_{\tau_n}$ or indirect intermediate estimator $\widetilde{\xi}_{\tau_n}$, and $\overline{\xi}^{\star}_{\tau'_n}$ the corresponding extrapolated version. Based on (B.1) and (B.3), the direct extrapolated estimator $\widehat{\xi}^{\star}_{\tau'_n}$ can then be corrected for its inherent bias, exclusively due to the heavy-tailed extrapolation, by the bias-reduced version,

$$\begin{aligned}\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n} \equiv \widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\overline{\gamma}) &:= \widehat{\xi}^{\star}_{\tau'_n}(1+\overline{B}_{1,n})(1+\overline{B}_{2,n})(1+\overline{B}_{3,n}) \\ &= \left(\frac{1-\tau'_n}{1-\tau_n}\right)^{-\overline{\gamma}} \widehat{\xi}_{\tau_n}(1+\overline{B}_{1,n})(1+\overline{B}_{2,n})(1+\overline{B}_{3,n}), \qquad \textbf{(B.5)}\end{aligned}$$

where $\overline{\gamma}$ is to be itself a bias-reduced estimator of the EVI γ. Similarly, from (B.2) and (B.4), a bias-corrected version of the indirect extrapolated estimator $\widetilde{\xi}^{\star}_{\tau'_n}$ is obtained as

$$\begin{aligned}\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n} \equiv \widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\overline{\gamma}) &:= \widetilde{\xi}^{\star}_{\tau'_n}(1+\overline{B}_{1,n})(1+\overline{B}_{3,n}) \\ &= \left(\frac{1-\tau'_n}{1-\tau_n}\right)^{-\overline{\gamma}} \left(\overline{\gamma}^{-1}-1\right)^{-\overline{\gamma}} \widehat{q}_{\tau_n}(1+\overline{B}_{1,n})(1+\overline{B}_{3,n}). \qquad \textbf{(B.6)}\end{aligned}$$

Under $\mathcal{C}_2(\gamma,\rho,A)$ and additional mild regularity conditions, if $\sqrt{n(1-\tau_n)}(\overline{\gamma}-\gamma) \to \Gamma$ for a nondegenerate distribution Γ, both bias-corrected estimators (B.5) and (B.6) converge in distribution to the same limit Γ as $\overline{\gamma}$ at the slower rate $\log((1-\tau_n)/(1-\tau'_n))/\sqrt{n(1-\tau_n)}$, as established in Theorem 2 in Girard et al. (2022b). For the choice of the bias-corrected EVI estimator $\overline{\gamma}$, Girard et al. (2022b) have suggested to use either the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,BR}}_n$ in (A.10) from Caeiro et al. (2005), or their bias-reduced version $\widehat{\gamma}^{\mathrm{E,BR}}_n$ in (A.11) of $\widehat{\gamma}^{\mathrm{E}}_n$ from Girard et al. (2022a). As shown in Theorem 3.1 of Caeiro et al. (2005), the bias-reduced Hill estimator $\widehat{\gamma}^{\mathrm{H,BR}}_n$ is $\sqrt{n(1-\tau_n)}$-asymptotically Gaussian with mean zero and the same variance γ^2 as $\widehat{\gamma}^{\mathrm{H}}_n$, while $\widehat{\gamma}^{\mathrm{E,BR}}_n$ is $\sqrt{n(1-\tau_n)}$-asymptotically Gaussian with mean zero and variance $\gamma^3(1-\gamma)\big/(1-2\gamma)$ as established in Theorem 1 of Girard et al. (2022b). The latter estimator is preferable variance-wise only when γ is less than 0.35. The function `tindexp` from the R package `Expectrem` allows one to compute either $\widehat{\gamma}^{E}_n$ (if argument `br=FALSE`) or $\widehat{\gamma}^{\mathrm{E,BR}}_n$ (if `br=TRUE`), while the function `extExpect` computes the estimators $\widehat{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{E}_n)$ (if argument `method= "direct"`) and $\widetilde{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}^{H}_n)$ (if `method= "indirect"`) as well as their bias-reduced versions $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ and $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$ (if `br=TRUE`).

Gaussian $100(1-\alpha)\%$ asymptotic confidence intervals for $\xi_{\tau'_n}$ based on the purely asymmetric least squares estimator $\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{E,BR}}_n)$ and its quantile-based competitor $\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H,BR}}_n)$, also available as part of the R package `Expectrem`, are

$$\widehat{I}^{(1)}_{\tau'_n}(\alpha) = \left[\widehat{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{E,BR}}_n)\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau'_n))}{\sqrt{n(1-\tau_n)}}\sqrt{\widehat{s}^{2,\mathrm{BR}}_n}\times z_{1-\alpha/2}\right)\right]$$

$$\widetilde{I}^{(1)}_{\tau'_n}(\alpha) = \left[\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau'_n}(\widehat{\gamma}^{\mathrm{H,BR}}_n)\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau'_n))}{\sqrt{n(1-\tau_n)}}\sqrt{\widetilde{\sigma}^{2,\mathrm{BR}}_n}\times z_{1-\alpha/2}\right)\right],$$

$$\text{with } \widehat{s}^{2,\mathrm{BR}}_n = \frac{(\widehat{\gamma}^{\mathrm{E,BR}}_n)^3(1-\widehat{\gamma}^{\mathrm{E,BR}}_n)}{1-2\widehat{\gamma}^{\mathrm{E,BR}}_n} \quad \text{and} \quad \widetilde{\sigma}^{2,\mathrm{BR}}_n = (\widehat{\gamma}^{\mathrm{H,BR}}_n)^2.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

These intervals tend, however, to have poor coverage as illustrated very recently by Daouia et al. (2024b): while the bias correction can be reasonably effective, the estimation of the variances of $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n'))\log(\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$ and $(\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau_n')))\log(\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}/\xi_{\tau_n'})$, by $\widehat{s}^{2,\mathrm{BR}}_n$ and $\widetilde{\sigma}^{2,\mathrm{BR}}_n$, respectively, can be a long way off the truth. To improve on this, they have suggested corrected versions of $\widehat{I}^{(1)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(1)}_{\tau_n'}(\alpha)$ defined, respectively, as

$$\widehat{I}^{(2)}_{\tau_n'}(\alpha)=\widehat{I}^{(2)}_{\tau_n'}(\alpha;J)=\left[\widehat{\xi}^{\star,\mathrm{BR}}_{\tau_n'}\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}}\sqrt{\widehat{s}^2_n(J)}\times z_{1-\alpha/2}\right)\right]$$
$$\widetilde{I}^{(2)}_{\tau_n'}(\alpha)=\widetilde{I}^{(2)}_{\tau_n'}(\alpha;J)=\left[\widetilde{\xi}^{\star,\mathrm{BR}}_{\tau_n'}\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}}\sqrt{\widetilde{\sigma}^2_n(J)}\times z_{1-\alpha/2}\right)\right],$$

where the expressions of $\widehat{s}^2_n(J)$ and $\widetilde{\sigma}^2_n(J)$, as well as the rationale behind their formulations, can be found in Daouia et al. (2024b), with $J\geq 1$ being a suitably chosen tuning parameter. The function `CIextExpect` from the R package `Expectrem` computes $\widehat{I}^{(2)}_{\tau_n'}(\alpha)$ with `method= "direct"` and $\widetilde{I}^{(2)}_{\tau_n'}(\alpha)$ with `method= "indirect"`. These confidence intervals have asymptotically the desired correct coverage as established in Theorems 1 and 2 in Daouia et al. (2024b), and they serve as a benchmark for our comparison purposes in Section 3.

An alternative solution put forward earlier in Padoan and Stupfler (2022) to correct the naive confidence intervals $\widehat{I}^{(0)}_{\tau_n'}(\alpha)$ and $\widetilde{I}^{(0)}_{\tau_n'}(\alpha)$ takes the respective forms

$$\widehat{I}^{(3)}_{\tau_n'}(\alpha)=\left[\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}}\sqrt{\overline{s}^2_n}\times z_{1-\alpha/2}\right)\right]$$
$$\text{and }\widetilde{I}^{(3)}_{\tau_n'}(\alpha)=\left[\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}\exp\left(\pm\frac{\log((1-\tau_n)/(1-\tau_n'))}{\sqrt{n(1-\tau_n)}}\sqrt{\overline{\sigma}^2_n}\times z_{1-\alpha/2}\right)\right],$$

where $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ and $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ (resp., $\overline{s}^2_n$ and $\overline{\sigma}^2_n$) are suitably chosen bias-corrected asymmetric least squares and quantile-based estimators of the expectile $\xi_{\tau_n'}$ (resp., the asymptotic variance of $\widehat{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$ and $\widetilde{\xi}^{\star,\mathrm{PS}}_{\tau_n'}$). See Padoan and Stupfler (2022) for full details. The `R` function `CIextExpect` from the `Expectrem` package computes both of these confidence intervals, with `method= "direct_PS"` for $\widehat{I}^{(3)}_{\tau_n'}(\alpha)$ and `method= "indirect_PS"` for $\widetilde{I}^{(3)}_{\tau_n'}(\alpha)$.

C. PROOFS

We will use in the sequel the notation $d_n:=(1-\tau_n)/(1-\tau_n')$ and $\phi_\gamma(t):=\int_1^t s^{\gamma-1}\log(s)\,\mathrm{d}s$. We shall freely use the following fact: under condition $\mathcal{E}_2(\gamma,\rho,a,A)$ with $\gamma>0$ and $\rho<0$, Theorem 3.1 in Fraga Alves et al. (2007) guarantees that there exist $\tilde{A}$ and $\tilde{\rho}<0$ such that the SORV condition $\mathcal{C}_2(\gamma,\tilde{\rho},\tilde{A})$ holds with $|\tilde{A}|$ being regularly varying with index $\tilde{\rho}$, and $\tilde{A}(t)\to 0$ as $t\to\infty$. Moreover, by Theorems 2.1 and 3.1 in Fraga Alves et al. (2007) in the case $\gamma>0$, the auxiliary function $\tilde{A}$ is asymptotically proportional to $\overline{A}(t)=a(t)/U(t)-\gamma$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Proof of Theorem 1. Write

$$\frac{\sqrt{n(1-\tau_n)}}{a(1/\overline{F}(\xi_{\tau_n}))\,\phi_\gamma(d_n)}(\breve{\xi}^{\star}_{\tau'_n}-\xi_{\tau'_n})$$

$$=\frac{\sqrt{n(1-\tau_n)}}{a(1/\overline{F}(\xi_{\tau_n}))\,\phi_\gamma(d_n)}(\widehat{\xi}_{\tau_n}-\xi_{\tau_n}) \tag{C.1}$$

$$+\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{\breve{\sigma}_n}{a(1/\overline{F}(\xi_{\tau_n}))}-1\right)\frac{d_n^{\breve{\gamma}_n}-1}{\breve{\gamma}_n} \tag{C.2}$$

$$+\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{d_n^{\breve{\gamma}_n}-1}{\breve{\gamma}_n}-\frac{d_n^{\gamma}-1}{\gamma}\right) \tag{C.3}$$

$$-\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau'_n}))^{\gamma}-d_n^{\gamma}}{\gamma}\right) \tag{C.4}$$

$$-\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{\xi_{\tau'_n}-\xi_{\tau_n}}{a(1/\overline{F}(\xi_{\tau_n}))}-\frac{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau'_n}))^{\gamma}-1}{\gamma}\right). \tag{C.5}$$

The key argument is that the third term (C.3) converges in distribution to Γ and all the other terms on the right-hand side are either $o(1)$ or $o_{\mathbb{P}}(1)$, as $n\to\infty$.

Indeed, since $d_n\to\infty$, $\log(n(1-\tau'_n))/\sqrt{n(1-\tau_n)}\to 0$ (as a consequence of assumptions $n(1-\tau_n)\to\infty$ and $\sqrt{n(1-\tau_n)}/\log((1-\tau_n)/(1-\tau'_n))\to\infty$) and $\sqrt{n(1-\tau_n)}(\breve{\gamma}_n-\gamma)\xrightarrow{\mathrm{d}}\Gamma$, we know by the proof of Theorem 4.3.1 on pp. 136–137 in de Haan and Ferreira (2006) that the term (C.3) has the same limit distribution Γ as $\sqrt{n(1-\tau_n)}(\breve{\gamma}_n-\gamma)$ (see the analysis of term II therein). In particular, since $n(1-\tau_n)\to\infty$, this implies that

$$\frac{1}{\phi_\gamma(d_n)}\left(\frac{d_n^{\breve{\gamma}_n}-1}{\breve{\gamma}_n}-\frac{d_n^{\gamma}-1}{\gamma}\right)=o_{\mathbb{P}}(1)\ \text{ as } n\to\infty. \tag{C.6}$$

On the other hand, by Remark 4.3.3 on p. 135 in de Haan and Ferreira (2006), we have

$$\phi_\gamma(d_n)\sim\frac{d_n^{\gamma}}{\gamma}\log(d_n)\ \text{ as } n\to\infty. \tag{C.7}$$

In addition, under the condition $\mathcal{E}_2(\gamma,\rho,a,A)$ with $\gamma>0$, we have according to the proof of Theorem 1.1.6 on pp. 10–11 in de Haan and Ferreira (2006) that $(t-U(1/\overline{F}(t)))/a(1/\overline{F}(t))\to 0$ as $t\to\infty$. Since $a(t)/U(t)\to\gamma$ as $t\to\infty$ in view of Lemma 1.2.9 on p. 22 in de Haan and Ferreira (2006), we conclude that $U(1/\overline{F}(\xi_\tau))\sim\xi_\tau$ as $\tau\to 1$ and then that

$$a(1/\overline{F}(\xi_{\tau_n}))\sim\gamma\,U(1/\overline{F}(\xi_{\tau_n}))\sim\gamma\,\xi_{\tau_n}\ \text{ as } n\to\infty. \tag{C.8}$$

Thus, using (C.7), the first term (C.1) can be expressed as

$$\frac{\sqrt{n(1-\tau_n)}}{a(1/\overline{F}(\xi_{\tau_n}))\,\phi_\gamma(d_n)}(\widehat{\xi}_{\tau_n}-\xi_{\tau_n})=O_{\mathbb{P}}\left(\frac{\sqrt{n(1-\tau_n)}}{d_n^{\gamma}\log(d_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)\right)\ \text{ as } n\to\infty.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Since $\sqrt{n(1-\tau_n)}\left(\frac{\widehat{\xi}_{\tau_n}}{\xi_{\tau_n}}-1\right)=\mathrm{O}_{\mathbb{P}}(1)$ and $d_n^{\gamma}\log(d_n)\to\infty$, the first term (C.1) is then $\mathrm{o}_{\mathbb{P}}(1)$.

The second term (C.2) is handled as follows: we have by (C.6) and (C.7) that

$$\frac{d_n^{\check{\gamma}_n}-1}{\check{\gamma}_n\,\phi_\gamma(d_n)}=\frac{d_n^{\gamma}-1}{\gamma\,\phi_\gamma(d_n)}+\mathrm{o}_{\mathbb{P}}(1)=\frac{1}{\log(d_n)}+\mathrm{o}_{\mathbb{P}}(1)=\mathrm{o}_{\mathbb{P}}(1)\ \text{ as } n\to\infty.$$

Thus, the second term (C.2) is $\mathrm{o}_{\mathbb{P}}(1)$ because $\sqrt{n(1-\tau_n)}\left(\frac{\check{\sigma}_n}{a(1/\overline{F}(\xi_{\tau_n}))}-1\right)=\mathrm{O}_{\mathbb{P}}(1)$.

Let us now turn to the term (C.4) which, due to (C.7), satisfies, as $n\to\infty$,

$$\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{d_n^{\gamma}-(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}}{\gamma}\right)\sim\frac{\sqrt{n(1-\tau_n)}}{\log(d_n)}\left(1-\left(\frac{\overline{F}(\xi_{\tau_n})}{d_n\overline{F}(\xi_{\tau_n'})}\right)^{\gamma}\right).\tag{C.9}$$

Since condition $\mathcal{C}_2(\gamma,\tilde{\rho},\tilde{A})$ holds, Proposition 1 in Daouia et al. (2018) implies that

$$\begin{aligned}&\frac{\overline{F}(\xi_\tau)}{1-\tau}=(\gamma^{-1}-1)(1+\varepsilon(\tau))\\ &\text{with }\ \varepsilon(\tau)=-\frac{(\gamma^{-1}-1)^{\gamma}}{q_\tau}(\mathbb{E}(X)+\mathrm{o}(1))-\frac{(\gamma^{-1}-1)^{-\tilde{\rho}}}{\gamma(1-\gamma-\tilde{\rho})}\tilde{A}((1-\tau)^{-1})(1+\mathrm{o}(1))\end{aligned}\tag{C.10}$$

as $\tau\to1$. It follows that $\frac{\overline{F}(\xi_{\tau_n})}{\overline{F}(\xi_{\tau_n'})}=d_n\frac{1+\varepsilon(\tau_n)}{1+\varepsilon(\tau_n')}$ as $n\to\infty$. Since $\varepsilon(\tau)\to0$ as $\tau\to1$, and $\varepsilon(\tau_n')=\mathrm{O}(\varepsilon(\tau_n))$ in view of the regular variation properties of U and of $|\tilde{A}|$, a Taylor expansion yields $\left(\frac{\overline{F}(\xi_{\tau_n})}{d_n\overline{F}(\xi_{\tau_n'})}\right)^{\gamma}-1=\mathrm{O}(\varepsilon(\tau_n))$ as $n\to\infty$. Thus, (C.9) becomes

$$\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{d_n^{\gamma}-(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}}{\gamma}\right)=\mathrm{O}\left(\frac{\sqrt{n(1-\tau_n)}}{\log(d_n)}\varepsilon(\tau_n)\right).\tag{C.11}$$

It remains to show that $\sqrt{n(1-\tau_n)}\varepsilon(\tau_n)=\mathrm{O}(1)$. For this, recall that $\tilde{A}$ is asymptotically proportional to $t\mapsto\overline{A}(t)=a(t)/U(t)-\gamma$; since $\sqrt{n(1-\tau_n)}\overline{A}((1-\tau_n)^{-1})\to\mu$, we have $\sqrt{n(1-\tau_n)}\tilde{A}((1-\tau_n)^{-1})=\mathrm{O}(1)$. It follows from the condition $\sqrt{n(1-\tau_n)}/q_{\tau_n}\to\lambda_2$ that $\sqrt{n(1-\tau_n)}\varepsilon(\tau_n)=\mathrm{O}(1)$, and hence the term (C.4) is o(1) by (C.11).

The last term (C.5) can be rewritten as

$$\begin{aligned}&-\frac{\sqrt{n(1-\tau_n)}}{\phi_\gamma(d_n)}\left(\frac{\xi_{\tau_n'}-\xi_{\tau_n}}{a(1/\overline{F}(\xi_{\tau_n}))}-\frac{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}{\gamma}\right)\\ &=-\sqrt{n(1-\tau_n)}A(1/\overline{F}(\xi_{\tau_n}))\frac{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}{\gamma\,\phi_\gamma(d_n)}\\ &\times\frac{1}{A(1/\overline{F}(\xi_{\tau_n}))}\left(\frac{\xi_{\tau_n'}-\xi_{\tau_n}}{a(1/\overline{F}(\xi_{\tau_n}))}\,\frac{\gamma}{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}-1\right).\end{aligned}\tag{C.12}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Under $\mathcal{E}_2(\gamma,\rho,a,A)$ with $\rho<0$, Lemma 4.3.5 on p. 135 in de Haan and Ferreira (2006) gives

$$\lim_{n\to\infty}\frac{1}{A(1/\overline{F}(\xi_{\tau_n}))}\left(\frac{U(1/\overline{F}(\xi_{\tau_n'}))-U(1/\overline{F}(\xi_{\tau_n}))}{a(1/\overline{F}(\xi_{\tau_n}))}\,\frac{\gamma}{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}-1\right)=-1/\rho.$$

In particular, combined with (C.8), and recalling that $\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'})\sim d_n\to\infty$, this yields

$$\left(\frac{\overline{F}(\xi_{\tau_n'})}{\overline{F}(\xi_{\tau_n})}\right)^{\gamma}\frac{U(1/\overline{F}(\xi_{\tau_n'}))}{U(1/\overline{F}(\xi_{\tau_n}))}\to 1\ \text{ as } n\to\infty. \tag{C.13}$$

Moreover, the proof of Theorem B.3.19 on p. 401 in de Haan and Ferreira (2006) yields $\lim_{t\to\infty}\frac{1}{A(1/\overline{F}(t))}\times\frac{U(1/\overline{F}(t))-t}{a(1/\overline{F}(t))}=0$. Then, on the one hand,

$$\lim_{n\to\infty}\frac{1}{A(1/\overline{F}(\xi_{\tau_n}))}\times\frac{U(1/\overline{F}(\xi_{\tau_n}))-\xi_{\tau_n}}{a(1/\overline{F}(\xi_{\tau_n}))}\,\frac{\gamma}{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}=0.$$

On the other hand, using again (C.8) together with (C.13) and the regular variation property of $|A|$ with index $\rho<0$ (see Theorem 2.3.3 on page 44 in de Haan and Ferreira, 2006), we find

$$\begin{aligned}&\frac{1}{A(1/\overline{F}(\xi_{\tau_n}))}\times\frac{U(1/\overline{F}(\xi_{\tau_n'}))-\xi_{\tau_n'}}{a(1/\overline{F}(\xi_{\tau_n}))}\,\frac{\gamma}{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}\\&=\mathrm{o}\left(\frac{1}{A(1/\overline{F}(\xi_{\tau_n'}))}\times\frac{U(1/\overline{F}(\xi_{\tau_n'}))-\xi_{\tau_n'}}{a(1/\overline{F}(\xi_{\tau_n'}))}\right)=\mathrm{o}(1).\end{aligned}$$

As a consequence,

$$\lim_{n\to\infty}\frac{1}{A(1/\overline{F}(\xi_{\tau_n}))}\left(\frac{\xi_{\tau_n'}-\xi_{\tau_n}}{a(1/\overline{F}(\xi_{\tau_n}))}\,\frac{\gamma}{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}-1\right)=-1/\rho. \tag{C.14}$$

Then, applying again (C.7), we arrive at

$$\frac{(\overline{F}(\xi_{\tau_n})/\overline{F}(\xi_{\tau_n'}))^{\gamma}-1}{\gamma\,\phi_{\gamma}(d_n)}\sim\frac{d_n^{\gamma}}{\gamma\,\phi_{\gamma}(d_n)}\sim\frac{1}{\log(d_n)}\to 0\ \text{ as } n\to\infty. \tag{C.15}$$

Now, write $\sqrt{n(1-\tau_n)}A(1/\overline{F}(\xi_{\tau_n}))=\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1})\frac{A(1/\overline{F}(\xi_{\tau_n}))}{A((1-\tau_n)^{-1})}$. Since $|A|$ is regularly varying with index ρ and $\overline{F}(\xi_{\tau_n})/(1-\tau_n)\to\gamma^{-1}-1$, one has, by local uniformity of the regular variation property, that $\left|A(1/\overline{F}(\xi_{\tau_n}))\right|\sim(\gamma^{-1}-1)^{-\rho}\left|A((1-\tau_n)^{-1})\right|$ as $n\to\infty$. Therefore, as $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1})=\mathrm{O}(1)$ by assumption, we get

$$\sqrt{n(1-\tau_n)}A(1/\overline{F}(\xi_{\tau_n}))=\mathrm{O}(1)\ \text{ as } n\to\infty. \tag{C.16}$$

Thus, combining (C.12) and (C.14)–(C.16), we find that the last term (C.5) is indeed $\mathrm{o}(1)$.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Let us now show that an equivalent statement of the convergence in Theorem 1 is

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\check{\xi}^{\star}_{\tau_n'}}{\xi_{\tau_n'}}-1\right)\xrightarrow{\mathrm{d}}\Gamma.$$

Since $\mathcal{C}_2(\gamma,\tilde{\rho},\tilde{A})$ holds with $\tilde{\rho}<0$, we know from p. 49 in de Haan and Ferreira (2006) that, for some constant $\mathfrak{c}>0$, we have $U(t)\sim \mathfrak{c}t^{\gamma}$ as $t\to\infty$. Then

$$\frac{q_{\tau_n'}}{q_{\tau_n}}=\frac{U((1-\tau_n')^{-1})}{U((1-\tau_n)^{-1})}\sim d_n^{\gamma}\quad\text{as } n\to\infty. \tag{C.17}$$

Since $\xi_\tau\sim(\gamma^{-1}-1)^{-\gamma}q_\tau$ as $\tau\to1$, we obtain the stronger version of (C.13): $\frac{\xi_{\tau_n'}}{\xi_{\tau_n}}\sim\frac{q_{\tau_n'}}{q_{\tau_n}}\sim d_n^{\gamma}$ as $n\to\infty$. The desired convergence result follows by using (C.7) and (C.8). □

Proof of Theorem 2. Write

$$\sqrt{n(1-\tau_n)}\left(\frac{\check{\sigma}_n}{a(1/\overline{F}(\xi_{\tau_n}))}-1\right)$$

$$=\frac{a((1-\tau_n)^{-1})}{a(1/\overline{F}(\xi_{\tau_n}))}\left(\check{\gamma}_n^{-1}-1\right)^{-\check{\gamma}_n}\times\sqrt{n(1-\tau_n)}\left(\frac{\widehat{a}((1-\tau_n)^{-1})}{a((1-\tau_n)^{-1})}-1\right)$$

$$+\frac{a((1-\tau_n)^{-1})}{a(1/\overline{F}(\xi_{\tau_n}))}\times\sqrt{n(1-\tau_n)}\left(\left(\check{\gamma}_n^{-1}-1\right)^{-\check{\gamma}_n}-\left(\gamma^{-1}-1\right)^{-\gamma}\right)$$

$$-\frac{a((1-\tau_n)^{-1})}{a(1/\overline{F}(\xi_{\tau_n}))}\times\sqrt{n(1-\tau_n)}\left(\frac{a(1/\overline{F}(\xi_{\tau_n}))}{a((1-\tau_n)^{-1})}-\left(\gamma^{-1}-1\right)^{-\gamma}\right).$$

By Theorem 2.3.3 on p. 44 in de Haan and Ferreira (2006), the function $a(\cdot)$ is (locally uniformly) regularly varying with index γ, so that $(1-\tau_n)/\overline{F}(\xi_{\tau_n})=(\gamma^{-1}-1)^{-1}+\mathrm{o}(1)$ implies that

$$\frac{a(1/\overline{F}(\xi_{\tau_n}))}{a((1-\tau_n)^{-1})}\sim\left(\gamma^{-1}-1\right)^{-\gamma}\quad\text{as } n\to\infty.$$

Hence, it suffices to show that

$$\left(\check{\gamma}_n^{-1}-1\right)^{-\check{\gamma}_n}\times\sqrt{n(1-\tau_n)}\left(\frac{\widehat{a}((1-\tau_n)^{-1})}{a((1-\tau_n)^{-1})}-1\right) \tag{C.18}$$

$$+\sqrt{n(1-\tau_n)}\left(\left(\check{\gamma}_n^{-1}-1\right)^{-\check{\gamma}_n}-\left(\gamma^{-1}-1\right)^{-\gamma}\right) \tag{C.19}$$

$$-\sqrt{n(1-\tau_n)}\left(\frac{a(1/\overline{F}(\xi_{\tau_n}))}{a((1-\tau_n)^{-1})}-\left(\gamma^{-1}-1\right)^{-\gamma}\right) \tag{C.20}$$

$$=\mathrm{O}_{\mathbb{P}}(1)\quad\text{as } n\to\infty.$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

The first term (C.18) is $\mathrm{O}_{\mathbb{P}}(1)$, since $\sqrt{n(1-\tau_n)}\left(\frac{\widehat{a}((1-\tau_n)^{-1})}{a((1-\tau_n)^{-1})}-1\right)=\mathrm{O}_{\mathbb{P}}(1)$, and, by the consistency of $\check{\gamma}_n$ as an estimator of $\gamma\in(0,1)$, we have $\left(\check{\gamma}_n^{-1}-1\right)^{-\check{\gamma}_n}=\mathrm{O}_{\mathbb{P}}(1)$. Likewise, the second term (C.19) is $\mathrm{O}_{\mathbb{P}}(1)$ by applying the delta method in conjunction with $\sqrt{n(1-\tau_n)}(\check{\gamma}_n-\gamma)\xrightarrow{\mathrm{d}}\Gamma$. As for the last term (C.20), we have again by Theorem 2.3.3 on p. 44 in de Haan and Ferreira (2006) (and Theorem B.2.18 on page 383 in de Haan and Ferreira, 2006, which guarantees local uniformity of this SORV property) that

$$\lim_{n\to\infty}\frac{1}{A((1-\tau_n)^{-1})}\left[\frac{a(1/\overline{F}(\xi_{\tau_n}))}{a((1-\tau_n)^{-1})}-\left(\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)^{-\gamma}\right]=\left(\gamma^{-1}-1\right)^{-\gamma}\frac{\left(\gamma^{-1}-1\right)^{-\rho}-1}{\rho}.$$

Moreover, from (C.10) and a Taylor expansion,

$$\left(\frac{\overline{F}(\xi_{\tau_n})}{1-\tau_n}\right)^{-\gamma}-\left(\gamma^{-1}-1\right)^{-\gamma}=\mathrm{O}(1/q_{\tau_n})+\mathrm{O}(\tilde{A}((1-\tau_n)^{-1})).$$

Recalling that $\sqrt{n(1-\tau_n)}A((1-\tau_n)^{-1})=\mathrm{O}(1)$, $\sqrt{n(1-\tau_n)}/q_{\tau_n}=\mathrm{O}(1)$ and $\sqrt{n(1-\tau_n)}$ $(a((1-\tau_n)^{-1})/q_{\tau_n}-\gamma)=\mathrm{O}(1)$, it follows that (C.20) is $\mathrm{O}(1)$ as $n\to\infty$, which ends the proof. □

Proof of Theorem 3. Recall from (2.9) that $\check{\xi}^{\star}_{\tau'_n}=(\widehat{\gamma}_n^{-1}-1)^{-\widehat{\gamma}_n}\check{q}^{\star}_{\tau'_n}$ and write

$$\begin{aligned}
&\frac{\sqrt{n(1-\tau_n)}}{a((1-\tau_n)^{-1})\,\phi_\gamma(d_n)}(\check{\xi}^{\star}_{\tau'_n}-\xi_{\tau'_n})\\
&=\frac{\sqrt{n(1-\tau_n)}}{a((1-\tau_n)^{-1})\,\phi_\gamma(d_n)}\left(\check{q}^{\star}_{\tau'_n}-q_{\tau'_n}\right)\times(\widehat{\gamma}_n^{-1}-1)^{-\widehat{\gamma}_n} &\textbf{(C.21)}\\
&+\frac{\sqrt{n(1-\tau_n)}}{a((1-\tau_n)^{-1})\,\phi_\gamma(d_n)}\left((\widehat{\gamma}_n^{-1}-1)^{-\widehat{\gamma}_n}-(\gamma^{-1}-1)^{-\gamma}\right)q_{\tau'_n} &\textbf{(C.22)}\\
&+\frac{\sqrt{n(1-\tau_n)}}{a((1-\tau_n)^{-1})\,\phi_\gamma(d_n)}\left((\gamma^{-1}-1)^{-\gamma}q_{\tau'_n}-\xi_{\tau'_n}\right). &\textbf{(C.23)}
\end{aligned}$$

By Theorem 4.3.1 on p. 134 in de Haan and Ferreira (2006) and the continuous mapping theorem, the first term (C.21) converges weakly to $\Gamma/(\gamma^{-1}-1)^{\gamma}$. In order to control the second term (C.22), note that since the SORV condition $\mathcal{C}_2(\gamma,\tilde{\rho},\tilde{A})$ holds, one has $a(t)/U(t)\to\gamma$ as $t\to\infty$ by Lemma 1.2.9 on p. 22 in de Haan and Ferreira (2006). Using the delta method and recalling (C.7) and then (C.17), one gets

$$\begin{aligned}
\frac{\sqrt{n(1-\tau_n)}}{a((1-\tau_n)^{-1})\,\phi_\gamma(d_n)}\left((\widehat{\gamma}_n^{-1}-1)^{-\widehat{\gamma}_n}-(\gamma^{-1}-1)^{-\gamma}\right)q_{\tau'_n}&=\mathrm{O}_{\mathbb{P}}\left(\frac{1}{\log(d_n)}\times d_n^{-\gamma}\frac{q_{\tau'_n}}{q_{\tau_n}}\right)\\
&=\mathrm{O}_{\mathbb{P}}\left(\frac{1}{\log(d_n)}\right).
\end{aligned}$$

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

It follows that the term (C.22) is $o_{\mathbb{P}}(1)$. We conclude with the control of the third term: by Proposition 1 in Daouia et al. (2020), $\xi_{\tau_n'} = (\gamma^{-1}-1)^{-\gamma} q_{\tau_n'}(1+r(\tau_n'))$, where, as $\tau \to 1$,

$$r(\tau) = \frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_\tau}\,(\mathbb{E}(X)+o(1)) + \left(\frac{(\gamma^{-1}-1)^{-\tilde\rho}}{1-\gamma-\tilde\rho} + \frac{(\gamma^{-1}-1)^{-\tilde\rho}-1}{\tilde\rho} + o(1)\right)\tilde A((1-\tau)^{-1}).$$

Recall that $\tilde A$ is asymptotically proportional to $t \mapsto \overline{A}(t) = a(t)/U(t) - \gamma$; since $\sqrt{n(1-\tau_n)}\overline{A}((1-\tau_n)^{-1}) \to \mu$ by assumption, we have $\sqrt{n(1-\tau_n)}\tilde A((1-\tau_n)^{-1}) = O(1)$, and then, from the condition $\sqrt{n(1-\tau_n)}/q_{\tau_n} \to \lambda_2$ and the regular variation properties of U and $|\overline{A}|$, that $\sqrt{n(1-\tau_n)}\, r(\tau_n') = o(1)$. As a consequence,

$$\frac{\sqrt{n(1-\tau_n)}}{a((1-\tau_n)^{-1})\,\phi_\gamma(d_n)}\left((\gamma^{-1}-1)^{-\gamma} q_{\tau_n'} - \xi_{\tau_n'}\right)$$
$$= o\left(\frac{q_{\tau_n'}}{a((1-\tau_n)^{-1})\phi_\gamma(d_n)}\right) = O\left(\frac{1}{\log(d_n)} \times d_n^{-\gamma}\,\frac{q_{\tau_n'}}{q_{\tau_n}}\right) = O\left(\frac{1}{\log(d_n)}\right)$$

and hence the term (C.23) is o(1). The equivalent statement

$$\frac{\sqrt{n(1-\tau_n)}}{\log((1-\tau_n)/(1-\tau_n'))}\left(\frac{\check\xi^{\star}_{\tau_n'}}{\xi_{\tau_n'}} - 1\right) \xrightarrow{\mathrm{d}} \Gamma$$

is obtained by using once again the asymptotic proportionality relationship $\xi_{\tau_n'} \sim (\gamma^{-1}-1)^{-\gamma} q_{\tau_n'}$ and arguing along the final lines of the proof of Theorem 1. □

Proof of Corollary 1. We obtain the two convergence results by applying Theorem 1. To do so, we first need to show that the last condition of this theorem is satisfied. Recall that under assumption $\mathcal{E}_2$ and the tail-heaviness condition (2.2) with $0 < \gamma < 1/2$, condition $\mathcal{C}_2$ is satisfied. Then, under the assumptions that $\mathbb{E}|X_-|^2 < \infty$, $\tau_n = 1 - k_n/n \to 1$ with $n(1-\tau_n) = k_n \to \infty$ and $\sqrt{n(1-\tau_n)}\tilde A((1-\tau_n)^{-1}) = O(1)$, one has, by Theorem 1 in Daouia et al. (2020), that

$$\sqrt{k_n}\left(\frac{\widehat\xi_{\tau_n}}{\xi_{\tau_n}} - 1\right) \xrightarrow{\mathrm{d}} \mathcal{N}\left(0, \frac{2\gamma^3}{1-2\gamma}\right), \text{ as } n \to \infty. \tag{C.24}$$

Note that $\sqrt{n(1-\tau_n)}\tilde A((1-\tau_n)^{-1}) = O(1)$ is guaranteed by the condition $\sqrt{n(1-\tau_n)}(a((1-\tau_n)^{-1})/q_{\tau_n} - \gamma) \to \mu \in \mathbb{R}$, see the proof of Theorem 1. Then, according to Theorem 2, it suffices to show that the two conditions of this theorem hold for both the GPML and Moment estimators of $a(n/k_n)$ and γ. Following Theorem 3.4.2 on p. 92 in de Haan and Ferreira (2006), the GPML estimators satisfy the joint convergence (2.12) and hence we obtain

$$\sqrt{k_n}\left(\frac{\widehat\sigma_n^{\mathrm{ML}}}{a(n/k_n)} - 1\right) = O_{\mathbb{P}}(1) \text{ and } \sqrt{k_n}(\widehat\gamma_n^{\mathrm{ML}} - \gamma) \xrightarrow{\mathrm{d}} \Gamma^{\mathrm{ML}}, \tag{C.25}$$

where Γ^{ML} is the normal distribution $\mathcal{N}\left(\frac{\lambda_1(1+\gamma)}{(1-\rho)(1+\gamma-\rho)}, (1+\gamma)^2\right)$. On the other hand, under the first condition of Theorem 1, we know from Lemma B.3.16 on p. 397 in de Haan

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

and Ferreira (2006) that a second-order condition $\mathcal{E}_2(\gamma_- = 0, \rho, a/U, Q)$ holds for $\log U(t)$ with an auxiliary function Q defined by

$$Q(t) = \begin{cases} \gamma - \dfrac{a(t)}{U(t)} & \text{if } \gamma = -\rho \text{ or } (0 < \gamma < -\rho \text{ and } l \neq 0), \\ \dfrac{\rho}{\gamma+\rho} A(t) & \text{if } 0 < -\rho < \gamma \text{ or } (0 < \gamma < -\rho \text{ and } l = 0), \end{cases}$$

where the quantity l is defined as $l = \lim_{t\to\infty} U(t) - a(t)/\gamma$, see Theorem 2.1 in Fraga Alves et al. (2007) for the existence of this limit. Then, as $n \to \infty$,

$$\sqrt{k_n}\, Q\left(\frac{n}{k_n}\right) \to \lambda^{\mathrm{Mom}} := \begin{cases} -\mu & \text{if } \gamma = -\rho \text{ or } (0 < \gamma < -\rho \text{ and } l \neq 0), \\ \dfrac{\lambda_1 \rho}{\gamma+\rho} & \text{if } 0 < -\rho < \gamma \text{ or } (0 < \gamma < -\rho \text{ and } l = 0). \end{cases}$$

Therefore, all conditions of Corollary 4.2.2 on p. 133 in de Haan and Ferreira (2006) are satisfied, which yields the joint convergence (2.13), and so we obtain

$$\sqrt{k_n}\left(\frac{\widehat{\sigma}_n^{\mathrm{Mom}}}{a(n/k_n)} - 1\right) = \mathrm{O}_{\mathbb{P}}(1) \text{ and } \sqrt{k_n}(\widehat{\gamma}_n^{\mathrm{Mom}} - \gamma) \xrightarrow{\mathrm{d}} \Gamma^{\mathrm{Mom}}, \tag{C.26}$$

where, by Theorem 3.5.4 on p. 104 in de Haan and Ferreira (2006), Γ^{Mom} stands for the normal distribution $\mathcal{N}\left(\lambda^{\mathrm{Mom}} b_{\gamma,\rho}, \gamma^2 + 1\right)$, with

$$b_{\gamma,\rho} = \begin{cases} -\dfrac{\gamma}{(1+\gamma)^2} & \text{if } 0 < \gamma < -\rho \text{ and } l \neq 0, \\ \dfrac{\gamma - \gamma\rho + \rho}{\rho(1-\rho)^2} & \text{if } 0 < -\rho \leq \gamma \text{ or } (0 < \gamma < -\rho \text{ and } l = 0). \end{cases}$$

Thus, by applying Theorem 2 in conjunction with (C.25) and (C.26), and then Theorem 1 in conjunction with (C.24), we get

$$\sqrt{k_n}\frac{\breve{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}_n^{\mathrm{ML}}, \breve{\sigma}_n^{\mathrm{ML}}) - \xi_{\tau'_n}}{a(1/\overline{F}(\xi_{\tau_n}))\,\phi_\gamma(d_n)} \xrightarrow{\mathrm{d}} \Gamma^{\mathrm{ML}} \text{ and } \sqrt{k_n}\frac{\breve{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}_n^{\mathrm{Mom}}, \breve{\sigma}_n^{\mathrm{Mom}}) - \xi_{\tau'_n}}{a(1/\overline{F}(\xi_{\tau_n}))\,\phi_\gamma(d_n)} \xrightarrow{\mathrm{d}} \Gamma^{\mathrm{Mom}}.$$

Since $\breve{\sigma}_n^{\mathrm{ML}}/a(1/\overline{F}(\xi_{\tau_n})) \xrightarrow{\mathbb{P}} 1$ and $\breve{\sigma}_n^{\mathrm{Mom}}/a(1/\overline{F}(\xi_{\tau_n})) \xrightarrow{\mathbb{P}} 1$, and since $\phi_{\widehat{\gamma}_n^{\mathrm{ML}}}(d_n)/\phi_\gamma(d_n) \xrightarrow{\mathbb{P}} 1$ and $\phi_{\widehat{\gamma}_n^{\mathrm{Mom}}}(d_n)/\phi_\gamma(d_n) \xrightarrow{\mathbb{P}} 1$ in view of Corollary 4.3.2 on p. 135 in de Haan and Ferreira (2006), the conclusion follows by using Slutsky's lemma. □

Proof of Corollary 2. According to Section 4.3.1 in de Haan and Ferreira (2006), see pp. 139–140, we have, under the first three conditions of Theorem 3, when the scale and shape GPML estimators are used, that $\frac{\sqrt{k_n}}{a(n/k_n)\phi_\gamma(d_n)}(\breve{q}^{\star}_{\tau'_n} - q_{\tau'_n}) \xrightarrow{\mathrm{d}} \Gamma^{\mathrm{ML}}$. Then, following the proof of Theorem 3, we obtain $\sqrt{k_n}\frac{\breve{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}_n^{\mathrm{ML}}, \widehat{\sigma}_n^{\mathrm{ML}}) - \xi_{\tau'_n}}{a(n/k_n)\phi_\gamma(d_n)} \xrightarrow{\mathrm{d}} \frac{\Gamma^{\mathrm{ML}}}{(\gamma^{-1}-1)^\gamma}$. Besides, as seen in the proof of Corollary 1, under the first and second conditions of Theorem 3, the Moment estimators satisfy the joint convergence (2.13) that corresponds to our last required condition

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

of Theorem 3. Therefore,

$$\sqrt{k_n}\frac{\check{\xi}^{\star}_{\tau'_n}(\widehat{\gamma}_n^{\mathrm{Mom}},\widehat{\sigma}_n^{\mathrm{Mom}})-\xi_{\tau'_n}}{a(n/k_n)\phi_\gamma(d_n)} \xrightarrow{\mathrm{d}} \frac{\Gamma^{\mathrm{Mom}}}{(\gamma^{-1}-1)^{\gamma}}.$$

Since $\widehat{\sigma}_n^{\mathrm{ML}}/a(n/k_n) \xrightarrow{\mathbb{P}} 1$, $\phi_{\widehat{\gamma}_n^{\mathrm{ML}}}(d_n)/\phi_\gamma(d_n) \xrightarrow{\mathbb{P}} 1$, $\widehat{\sigma}_n^{\mathrm{Mom}}/a(n/k_n) \xrightarrow{\mathbb{P}} 1$, and $\phi_{\widehat{\gamma}_n^{\mathrm{Mom}}}(d_n)/\phi_\gamma(d_n) \xrightarrow{\mathbb{P}} 1$, the proof is complete following a use of Slutsky's lemma. □

FUNDING STATEMENT

This research was supported by the French National Research Agency under the grants ANR-18-EURE-0023 (EUR MINT), ANR-17-EURE-0010 (EUR CHESS, Investissements d'Avenir Program), ANR-23-CE40-0009 (EXSTA Project), and ANR-11-LABX-0020-01 (Centre Henri Lebesgue). G.S. acknowledges further support from the Chair Stress Test, RISK Management and Financial Steering of the Foundation Ecole Polytechnique.

COMPETING INTERESTS

The authors declare that no competing interests exist.

REFERENCES

Acerbi, C. (2002). Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking and Finance*, 26(7), 1505–1518.

Acerbi, C., & Székely, B. (2014). Backtesting expected shortfall. MSCI, 1–34.

Acerbi, C., & Tasche, D. (2002). On the coherence of expected shortfall. *Journal of Banking and Finance*, 26(7), 1487–1503.

Aigner, D. J., Amemiya, T., & Poirier, D. J. (1976). On the estimation of production frontiers: Maximum likelihood estimation of the parameters of a discontinuous density function. *International Economic Review*, 17(2), 377–396.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203–228.

Beirlant, J., Goegebeur, Y., Segers, J., & Teugels, J. L. (2004). *Statistics of extremes: Theory and applications*. Wiley.

Bellini, F., & Di Bernardino, E. (2017). Risk management with expectiles. *The European Journal of Finance*, 23(6), 487–506.

Bellini, F., Klar, B., Müller, A., & Gianin, E. R. (2014). Generalized quantiles as risk measures. *Insurance: Mathematics and Economics*, 54, 41–48.

Belzile, L., Wadsworth, J. L., Northrop, P. J., Huser, R., Grimshaw, S. D., Zhang, J., Stephens, M. A., & Owen, A. B. (2025). mev: Modelling of extreme values. R package version 2.1.

Breckling, J., & Chambers, R. (1988). M-quantiles. *Biometrika*, 75(4), 761–772.

Caeiro, F., Gomes, M. I., & Pestana, D. (2005). Direct reduction of bias of the classical Hill estimator. *REVSTAT-Statistical Journal*, 3(2), 113–136.

Coles, S. (2001). *An introduction to statistical Modeling of extreme values*. Springer-Verlag.

Daníelsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O., & Shin, H. S. (2001). An academic response to Basel II. Special paper no. 130, Financial Markets Group, London School of Economics.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Daouia, A., Girard, S., & Stupfler, G. (2018). Estimation of tail risk based on extreme expectiles. *Journal of the Royal Statistical Society: Series B*, 80(2), 263–292.

Daouia, A., Girard, S., & Stupfler, G. (2019). Extreme M-quantiles as risk measures: From L^1 to L^p optimization. *Bernoulli*, 25(1), 264–309.

Daouia, A., Girard, S., & Stupfler, G. (2020). Tail expectile process and risk assessment. *Bernoulli*, 26(1), 531–556.

Daouia, A., Girard, S., & Stupfler, G. (2021). ExpectHill estimation, extreme risk and heavy tails. *Journal of Econometrics*, 221(1), 97–117.

Daouia, A., Padoan, S. A., & Stupfler, G. (2024a). Extreme expectile estimation for short-tailed data. *Journal of Econometrics*, 241(2), 105770.

Daouia, A., Stupfler, G., & Usseglio-Carleve, A. (2024b). Bias-reduced and variance-corrected asymptotic Gaussian inference about extreme expectiles. *Statistics and Computing*, 34(4), 130.

Daouia, A., Stupfler, G., & Usseglio-Carleve, A. (2025). Corrected inference about the extreme expected shortfall in the general max-domain of attraction. *Information and Inference*, 14(3), iaaf023.

Davison, A. C., Padoan, S. A., & Stupfler, G. (2023). Tail risk inference via expectiles in heavy-tailed time series. *Journal of Business & Economic Statistics*, 41(3), 876–889.

de Haan, L., & Ferreira, A. (2006). *Extreme value theory: An introduction*. Springer-Verlag.

Dekkers, A. L. M., Einmahl, J. H. J., & de Haan, L. (1989). A moment estimator for the index of an extreme-value distribution. *Annals of Statistics*, 17(4), 1833–1855.

Drees, H. (2003). Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, 9(4), 617–657.

Drees, H., Ferreira, A., & de Haan, L. (2004). On maximum likelihood estimation of the extreme value index. *Annals of Applied Probability*, 14(3), 1179–1201.

Fissler, T., & Ziegel, J. F. (2016). Higher order elicitability and Osband's principle. *Annals of Statistics*, 44(4), 1680–1707.

Fraga Alves, M. I., Gomes, M. I., & de Haan, L. (2003). A new class of semi-parametric estimators of the second order parameter. *Portugaliæ Mathematica*, 60(2), 193–213.

Fraga Alves, M. I., Gomes, M. I., de Haan, L., & Neves, C. (2007). A note on second order conditions in extreme value theory: Linking general and heavy tail conditions. *REVSTAT-Statistical Journal*, 5(3), 285–304.

Girard, S., Stupfler, G., & Usseglio-Carleve, A. (2022a). Nonparametric extreme conditional expectile estimation. *Scandinavian Journal of Statistics*, 49(1), 78–115.

Girard, S., Stupfler, G., & Usseglio-Carleve, A. (2022b). On automatic bias reduction for extreme expectile estimation. *Statistics & Computing*, 32(4), 64.

Gneiting, T. (2011). Making and evaluating point forecasts. *Journal of the American Statistical Association*, 106(494), 746–762.

Gomes, M. I., & Martins, M. J. (2002). "Asymptotically unbiased" estimators of the tail index based on external estimation of the second order parameter. *Extremes*, 5(1), 5–31.

He, Y., Peng, L., Zhang, D., & Zhao, Z. (2022). Risk analysis via generalized Pareto distributions. *Journal of Business & Economic Statistics*, 40(2), 852–867.

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

Kuan, C.-M., Yeh, J.-H., & Hsu, Y.-C. (2009). Assessing value at risk with CARE, the conditional autoregressive expectile models. *Journal of Econometrics*, 150(2), 261–270.

Liu, X., Tan, Z., Wu, Y., & Zhou, Y. (2023). The financial risk measurement EVaR based on DTARCH models. *Entropy*, 25(8), 1204.

McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *Journal of Empirical Finance*, 7(3–4), 271–300.

Newey, W. K., & Powell, J. L. (1987). Asymmetric least squares estimation and testing. *Econometrica*, 55(4), 819–847.

Northrop, P. J., & Coleman, C. L. (2014). Improved threshold diagnostic plots for extreme value analyses. *Extremes*, 17(2), 289–303.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668

Padoan, S. A., & Stupfler, G. (2022). Joint inference on extreme expectiles for multivariate heavy-tailed distributions. *Bernoulli*, 28(2), 1021–1048.

Philipps, C. (2022). Interpreting expectiles. Working Papers 2022-01, Department of Economics and Geosciences, US Air Force Academy.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distributions. *Journal of Banking and Finance*, 26(7), 1443–1471.

Schneider, L. F., Krajina, A., & Krivobokova, T. (2021). Threshold selection in univariate extreme value analysis. *Extremes*, 24(4), 881–913.

Smith, R. L. (1987). Estimating tails of probability distributions. *Annals of Statistics*, 15(3), 1174–1207.

Wadsworth, J. L. (2016). Exploiting structure of maximum likelihood estimators for extreme value threshold selection. *Technometrics*, 58(1), 116–126.

Wang, R., & Zitikis, R. (2021). An axiomatic foundation for the expected shortfall. *Management Science*, 67(3), 1413–1429.

Yılmaz, A., Kara, M., & Özdemir, O. (2021). Comparison of different estimation methods for extreme value distribution. *Journal of Applied Statistics*, 48(13–15), 2259–2284.

Zaevski, T. S., & Nedeltchev, D. C. (2023). From Basel III to Basel IV and beyond: Expected shortfall and expectile risk measures. *International Review of Financial Analysis*, 87, 102645.

Ziegel, J. F. (2016). Coherence and elicitability. *Mathematical Finance*, 26(4), 901–918.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 15 Nov 2026 at 15:09:38, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S0266466626100668