Supplementary Material for

# "Weighted approximations of the tail empirical expectile process"

Abdelaati Daouia[a], Stéphane Girard[b] and Gilles Stupfler[c]

[a] Toulouse School of Economics, University of Toulouse Capitole, France

[b] Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK

[c] School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, United Kingdom

Section A illustrates the behavior of the *expectHill* estimator of the tail index with data examples. Section B plots and comments on the asymptotic variance of the *expectHill* estimator. Simulation results are discussed in Section C. Section D applies our expectile-based method to estimate the expected shortfall of three large US investment banks. The proofs of all theoretical results in the main paper and additional technical results are provided in Section E.

## A Examples of tail index estimation

The aim of this section is to illustrate the behavior of the *expectHill* estimator with data examples and to highlight some of the theoretical findings in Section 3 of the main article. First, the purely expectile-based estimator

$$\widetilde{\gamma}_{\tau_n} = \frac{1}{\lfloor n(1-\tau_n)\rfloor} \sum_{i=1}^{\lfloor n(1-\tau_n)\rfloor} \log\left(\frac{\widetilde{\xi}_{1-(i-1)/n}}{\widetilde{\xi}_{1-\lfloor n(1-\tau_n)\rfloor/n}}\right)$$

of the tail index $\gamma$ has exactly the same form as the quantile-based Hill estimator

$$\widehat{\gamma}_{\tau_n} = \frac{1}{\lfloor n(1-\tau_n)\rfloor} \sum_{i=1}^{\lfloor n(1-\tau_n)\rfloor} \log\left(\frac{\widehat{q}_{1-(i-1)/n}}{\widehat{q}_{1-\lfloor n(1-\tau_n)\rfloor/n}}\right)$$

with the tail empirical quantile process $\widehat{q}$ replaced by its least squares analogue $\widetilde{\xi}$. Theorem 4 gives its asymptotic normality. As pointed out in Remark 1, the conditions involving the auxiliary function $A$ in Theorem 4 are also required to derive the asymptotic normality of Hill's estimator $\widehat{\gamma}_{\tau_n}$. These conditions are, however, difficult to check in practice, which makes the choice of the intermediate sequence $\tau_n$ a hard problem. A usual practice for selecting a

reasonable estimate $\widehat{\gamma}_{\tau_n}$ is to set $\tau_n = 1 - k/n$ for a sequence of integers $k$, then to plot the graph of $k \mapsto \widehat{\gamma}_{1-k/n}$, and finally to pick out a value of $k$ corresponding to the first stable part of the plot (see Remark 2). Yet, the Hill plot may be so unstable that reasonable values of $k$ (which would correspond to the true value of $\gamma$) may be hidden in the graph. The least squares analogue $\widetilde{\gamma}_{1-k/n}$ affords a smoother and more stable plot which counteracts the volatility defect of the Hill plot. This is illustrated in the following two examples.

**Example 1.** Figure 1(a)-(c) shows the paths $k \mapsto \widehat{\gamma}_{1-k/n}$ in red and $k \mapsto \widetilde{\gamma}_{1-k/n}$ in blue, for three large US financial institutions. We consider the same investment banks as in the study of Cai *et al.* (2015), namely Goldman Sachs, Morgan Stanley and T. Rowe Price. The dataset consists of the loss returns (*i.e.* minus log-returns) on their equity prices at a daily frequency from July 3rd, 2000, to June 30th, 2010. The chosen stable regime in Cai *et al.* (2015) is $k \in [70, 100]$ for the three Hill plots. To gain stability in the estimates, they took the average of the estimates $\widehat{\gamma}_{1-k/n}$ over this region. The results are reported in the second column of Table 1. As regards the asymmetric least squares estimator $\widetilde{\gamma}_{1-k/n}$, we applied a very simple technique which consists in computing its standard deviations over a moving window of 30 successive values of $k$ [same length as the chosen interval in Cai *et al.* (2015)]; this corresponds to a window large enough to cover around 20% of the possible values of $k$ in the selected range $1 \leqslant k \leqslant 150$. The value of $k$ where the standard deviation (and hence the variation) of the estimates is minimal defines the desired sample fraction $\widetilde{k}$. We found $\widetilde{k} = 72$ in the window $[54, 84]$ for Goldman Sachs, $\widetilde{k} = 80 \in [62, 92]$ for Morgan Stanley, and $\widetilde{k} = 88 \in [68, 98]$ for T. Rowe Price. The final estimates $\widetilde{\gamma}_{1-\widetilde{k}/n}$ are reported in the third column of Table 1. The messages yielded by the two methods are broadly similar, indicating particularly that Morgan Stanley displays a greater variability in loss returns and a much heavier tail of their distribution than do Goldman Sachs and T. Rowe Price.

| *Bank* | $\widehat{\gamma}_{1-k/n}$ | $\widetilde{\gamma}_{1-k/n}$ |
|---|---|---|
| Goldman Sachs | 0.3877 | 0.3720 |
| Morgan Stanley | 0.4645 | 0.4221 |
| T. Rowe Price | 0.3781 | 0.3653 |

Table 1: *Tail index estimates based on daily loss returns* ($n = 2513$).

**Example 2.** Figure 1(d) shows the paths $k \mapsto \widehat{\gamma}_{1-k/n}$ and $k \mapsto \widetilde{\gamma}_{1-k/n}$ for the Society of Actuaries (SOA) group medical insurance large claims. We consider the same database as in Beirlant *et al.* (2004) that contains $75,789$ claim amounts exceeding 25,000 USD, collected

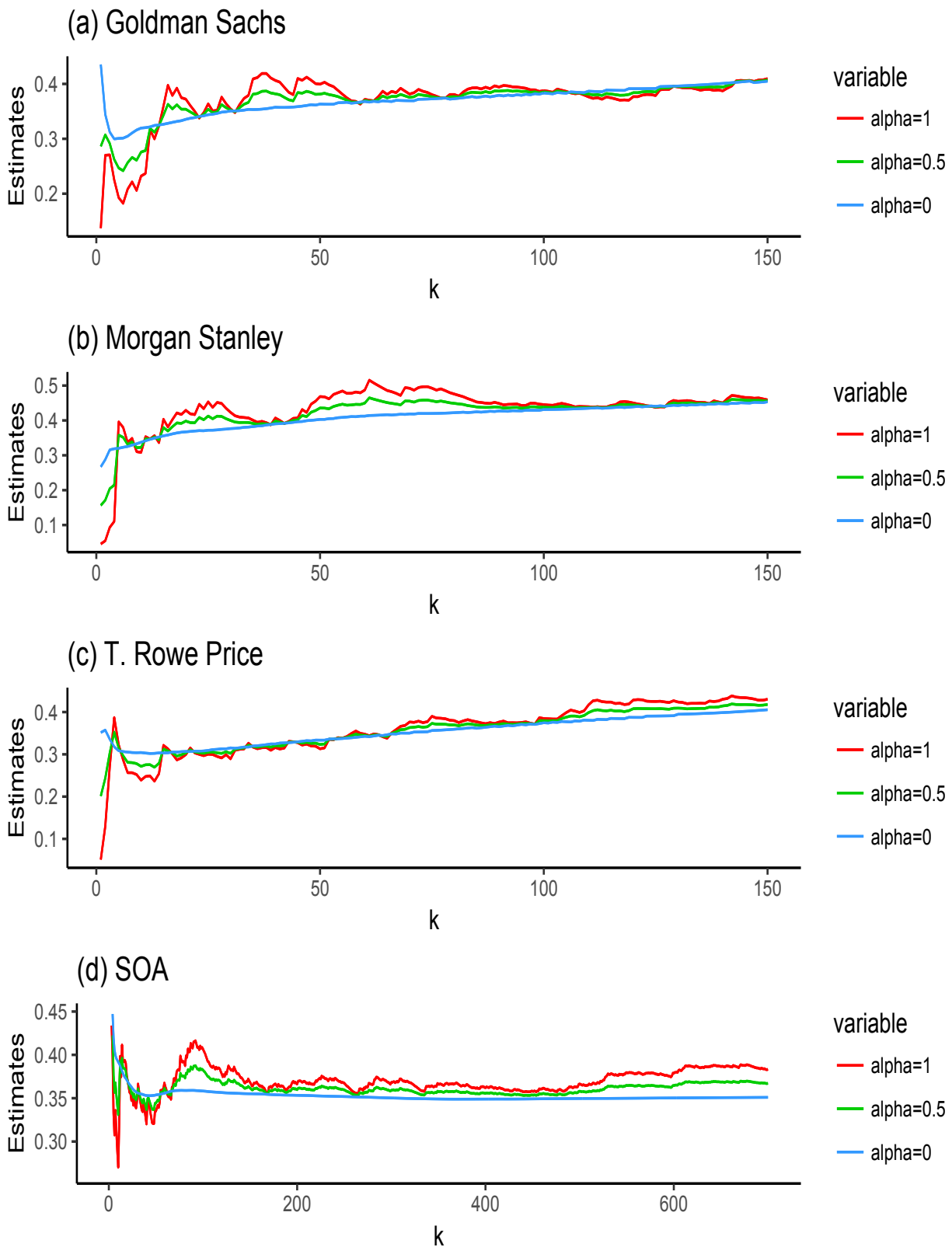



Figure 1: *Plots of* $\widehat{\gamma}_{1-k/n}$ *in red,* $\widetilde{\gamma}_{1-k/n}$ *in blue, and* $\overline{\gamma}_{1-k/n}(1/2)$ *in green, for the three banks in (a)-(c) and for the SOA group medical insurance large claims in (d).*

over the year 1991 from 26 insurers. The minimal standard deviation of the estimates indicates a pointwise estimate $\widehat{\gamma}_{1-k/n}$ around 0.36 achieved over the window $[119, 259]$, and an estimate $\widetilde{\gamma}_{1-k/n}$ around 0.35 attained over $[331, 471]$. Here also the standard deviations were computed over a moving window large enough to cover 20% of the possible values of $k$ in the selected range $1 \leqslant k \leqslant 700$.

In these examples, the purely least asymmetrically weighted squares estimator $\widetilde{\gamma}_{1-k/n}$ seems to be beneficial in producing smoother and more pleasing plots, but these plots may not be more revealing than Hill plots. Already in Figure 1(a)-(c), it may be seen that the smooth paths of $\widetilde{\gamma}_{1-k/n}$ can exhibit a sample-wise monotonic evolution with $k$. This may result in estimates with higher bias than the Hill estimates. One way to reduce this potential defect is by using a linear combination of $\widetilde{\gamma}$ and $\widehat{\gamma}$ for estimating $\gamma$. For $\alpha \in \mathbb{R}$ we have then defined the more general *expectHill* estimator

$$\overline{\gamma}_{\tau_n}(\alpha) = \alpha\widehat{\gamma}_{\tau_n} + (1-\alpha)\widetilde{\gamma}_{\tau_n}. \tag{A.1}$$

For example, as visualized in Figure 1, the simple mean $\overline{\gamma}_{\tau_n}(1/2)$ in green line would represent a reasonable compromise between the use of large asymmetric least squares in $\widetilde{\gamma}_{\tau_n}$ and top order statistics in $\widehat{\gamma}_{\tau_n}$.

# B Asymptotic variance of the *expectHill* estimator

The optimal value of the weighting coefficient $\alpha$ in (A.1), which minimizes the asymptotic variance $v_\alpha$ of $\overline{\gamma}_{\tau_n}(\alpha)$, only depends on the tail index $\gamma$ and has the explicit expression

$$\alpha(\gamma) = \frac{(1-\gamma) - (1-2\gamma)(\gamma^{-1}-1)^{\gamma}}{(1-\gamma)(3-4\gamma) - 2(1-2\gamma)(\gamma^{-1}-1)^{\gamma}}.$$

Its plot against $\gamma \in (0, 1/2)$ is given in Figure 2(a). Interestingly, the optimal $\alpha$ is negative for small values of $\gamma$, say $\gamma \leqslant 0.2$. By contrast, for large values of $\gamma$ (close to 1/2), the optimal $\alpha$ tends to one, favoring thus the robustness of order statistics over the tail sensitivity of asymmetric least squares. It can also be seen that the simple mean $\overline{\gamma}_{\tau_n}(1/2)$ of $\widehat{\gamma}_{\tau_n}$ and $\widetilde{\gamma}_{\tau_n}$, with $\alpha = 1/2$, is optimal for $\gamma = 1/4$. This is unsurprising since both $\widehat{\gamma}_{\tau_n}$ and $\widetilde{\gamma}_{\tau_n}$ have the same asymptotic variance in this case, as illustrated in Figure 2(b). This figure also shows that the mean $\overline{\gamma}_{\tau_n}(1/2)$ of $\widehat{\gamma}_{\tau_n}$ and $\widetilde{\gamma}_{\tau_n}$ affords a middle course between $\widehat{\gamma}_{\tau_n} \equiv \overline{\gamma}_{\tau_n}(1)$ and $\widetilde{\gamma}_{\tau_n} \equiv \overline{\gamma}_{\tau_n}(0)$ in terms of asymptotic variance. In terms of smoothness, $\overline{\gamma}_{\tau_n}(1/2)$ offers a middle course as well, as shown above in Figure 1, where the plot of $\overline{\gamma}_{\tau_n}(1/2)$ is superimposed in green line with the plots of $\widehat{\gamma}_{\tau_n}$ and $\widetilde{\gamma}_{\tau_n}$.

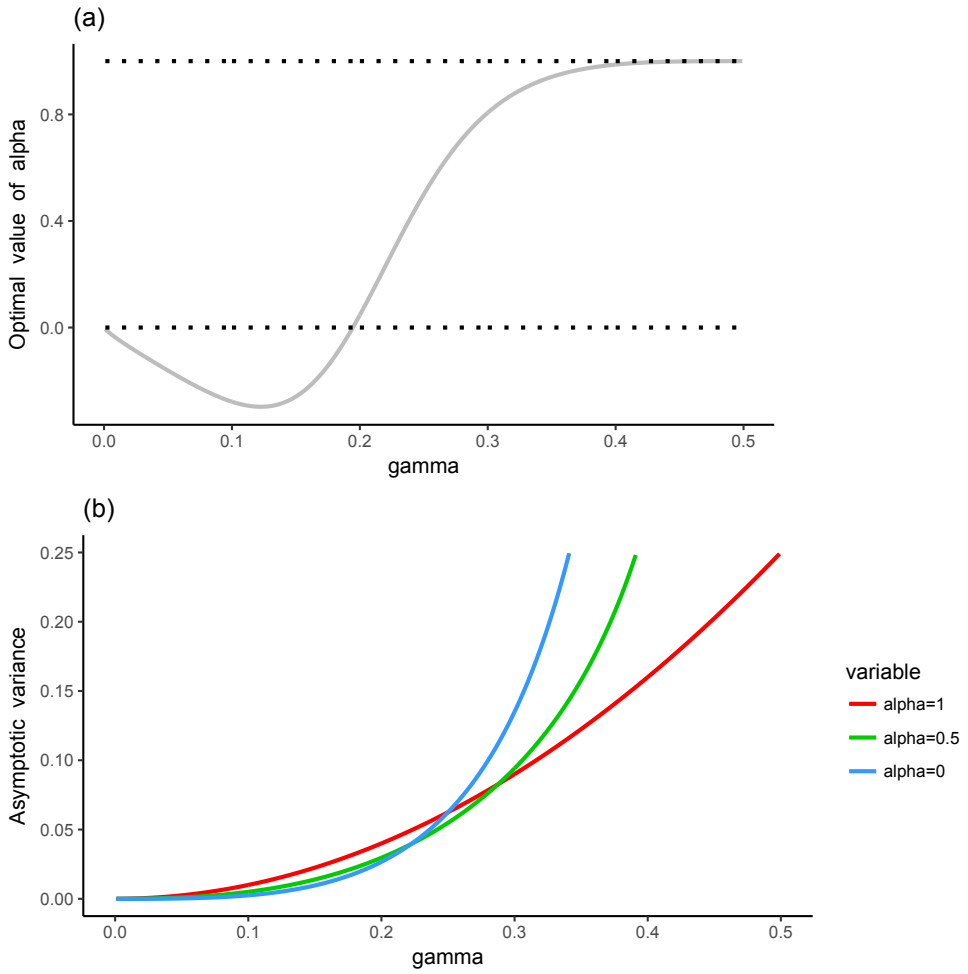



Figure 2: *(a) – Evolution of the optimal value* $\alpha(\gamma)$ *against* $\gamma \in (0, 1/2)$. *The dotted lines represent the values* $\alpha = 0$ *and* $\alpha = 1$. *(b) – Asymptotic variance* $v_\alpha$ *of* $\widehat{\gamma}_{\tau_n}$ *in red* $(\alpha = 1)$, $\widetilde{\gamma}_{\tau_n}$ *in blue* $(\alpha = 0)$, *and* $\overline{\gamma}_{\tau_n}(1/2)$ *in green* $(\alpha = 0.5)$, *as function of* $\gamma \in (0, 1/2)$.

# C Some simulation evidence

The aim of this section is to explore some features that were mentioned in Section 6 of the main article. We will illustrate the following points:

(C.1) Estimates of $\gamma$.

(C.2) Estimates of $\mathrm{XES}_{\tau'_n}$.

(C.3) Estimates of $\mathrm{QES}_{p_n}$.

(C.4) Confidence intervals for $\mathrm{QES}_{p_n}$.

In order to illustrate the behavior of the presented estimation procedures, we use the same

considerations as in Section 6 of the main paper. Namely, we consider the Student $t$-distribution with degree of freedom $1/\gamma$, the Fréchet distribution $F(x) = e^{-x^{-1/\gamma}},\ x > 0$, and the Pareto distribution $F(x) = 1 - x^{-1/\gamma},\ x > 1$. The finite-sample performance of the different estimators is evaluated through their relative Mean-Squared Error (MSE) and bias, computed over 200 replications. All the experiments have sample size $n = 500$ and true tail index $\gamma \in \{0.35, 0.45\}$. In our simulations we used the extreme levels $\tau_n' = p_n = 1 - \frac{1}{n}$ and the intermediate level $\tau_n = 1 - \frac{k}{n}$, where the integer $k$ can be viewed as the effective sample size for tail extrapolation.

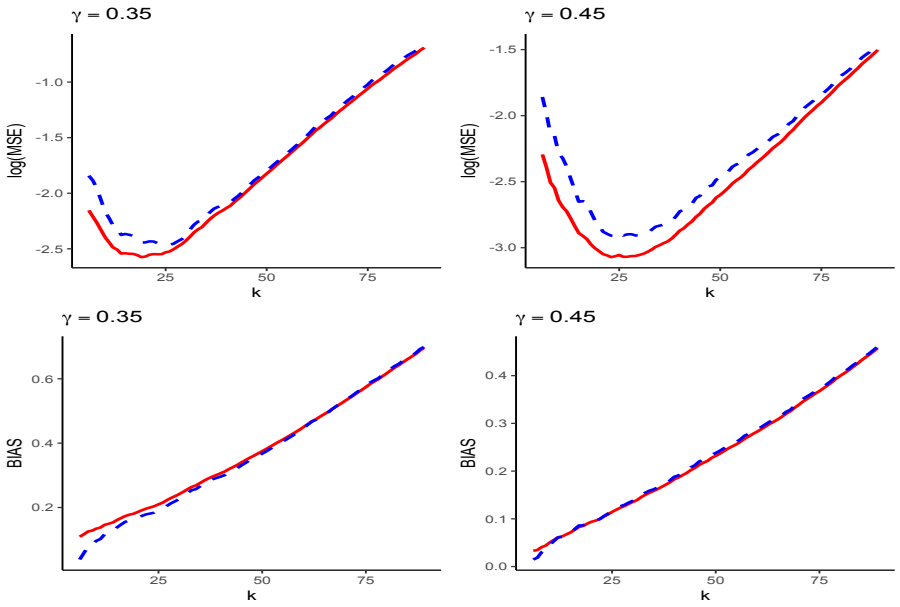
## C.1 Estimation of the tail index

This section provides Monte-Carlo evidence that the *expectHill* estimator $\overline{\gamma}_{1-k/n}(\alpha)$, introduced in (A.1) with the weight $\alpha = 1/2$, is more efficient relative to the standard Hill estimator $\widehat{\gamma}_{1-k/n}$, for both Student and Fréchet distributions. In the case of the real-valued Student distribution, Figure 3(a) gives the evolution of the MSE (in top panels) and the bias (in bottom panels) of $\overline{\gamma}_{1-k/n}(\frac{1}{2})/\gamma$ and $\widehat{\gamma}_{1-k/n}/\gamma$, as functions of the effective sample fraction $k$. It may be seen that $\overline{\gamma}_{1-k/n}(\frac{1}{2})$ performs better than $\widehat{\gamma}_{1-k/n}$ in terms of MSE, for all values of $k$, without sacrificing too much quality in terms of bias, especially for the larger value of $\gamma$. We arrive at the same tentative conclusion in the case of the Fréchet distribution as may be seen from Figure 3(b). By contrast, in the special case of the Pareto distribution, the Hill estimator $\widehat{\gamma}_{1-k/n}$ is exactly the maximum likelihood estimator of $\gamma$ and is unbiased, whereas the *expectHill* estimator $\overline{\gamma}_{1-k/n}(\frac{1}{2}) = \frac{1}{2}(\widehat{\gamma}_{1-k/n} + \widetilde{\gamma}_{1-k/n})$ is biased in this case. Unsurprisingly, the Monte Carlo results obtained in Figure 4 indicate that $\widehat{\gamma}_{1-k/n}$ is as expected the winner in this case.

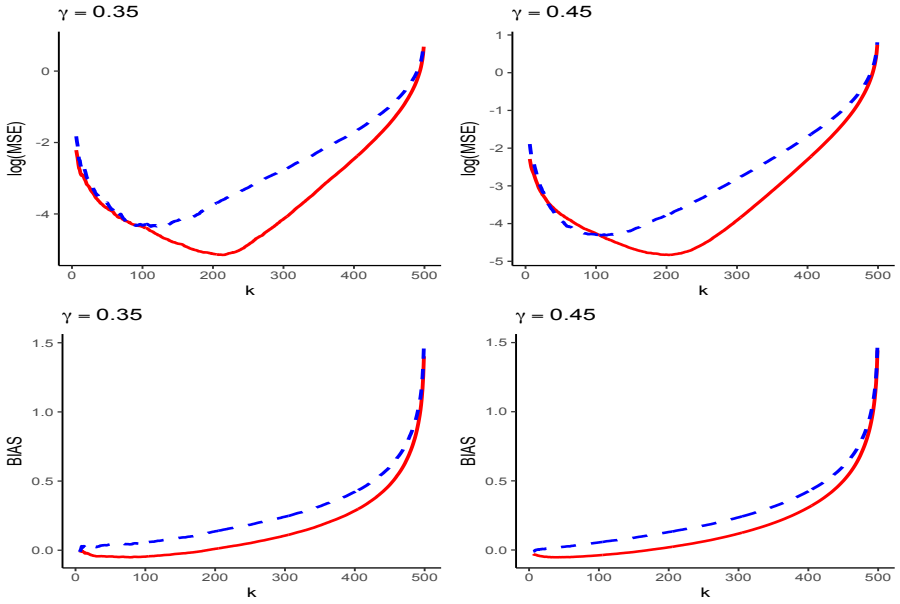
## C.2 Estimates of $\mathrm{XES}_{\tau_n'}$

Before comparing the performance of $\overline{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha, \beta)$, $\widehat{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha, \beta)$ and $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)$ as estimators of $\mathrm{XES}_{\tau_n'}$, we first investigated the accuracy of each estimator in terms of the associated weights $\alpha$ and $\beta$.

Figures 5 and 6 give the evolution of the MSE (in log scale) and bias estimates of $\widetilde{\mathrm{XES}}^{\star}_{\tau_n'}(\alpha)/\mathrm{XES}_{\tau_n'}$, as functions of the sample fraction $k$, for $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$. In the case of Student distribution, it may be seen that the red curve ($\alpha = 1$) gives the best estimates in terms of both MSE and Bias. In the case of Fréchet distribution, it may be seen that the purple curve ($\alpha = 0.5$) performs quite well in terms of MSE for both values of $\gamma$, but the blue curve ($\alpha = 0$) performs clearly better in terms of Bias. In the case of Pareto distribution, it may be seen that the purple ($\alpha = 0.5$) and red ($\alpha = 1$) curves have,




(a) Student distribution




(b) Fréchet distribution

Figure 3: MSE estimates in log scale (top panels) and Bias estimates (bottom panels) of $\overline{\gamma}_{1-k/n}(\alpha)/\gamma$ (solid red line) and $\widehat{\gamma}_{1-k/n}/\gamma$ (dashed blue line), as functions of $k$, for $\gamma = 0.35$ (left) and $\gamma = 0.45$ (right).

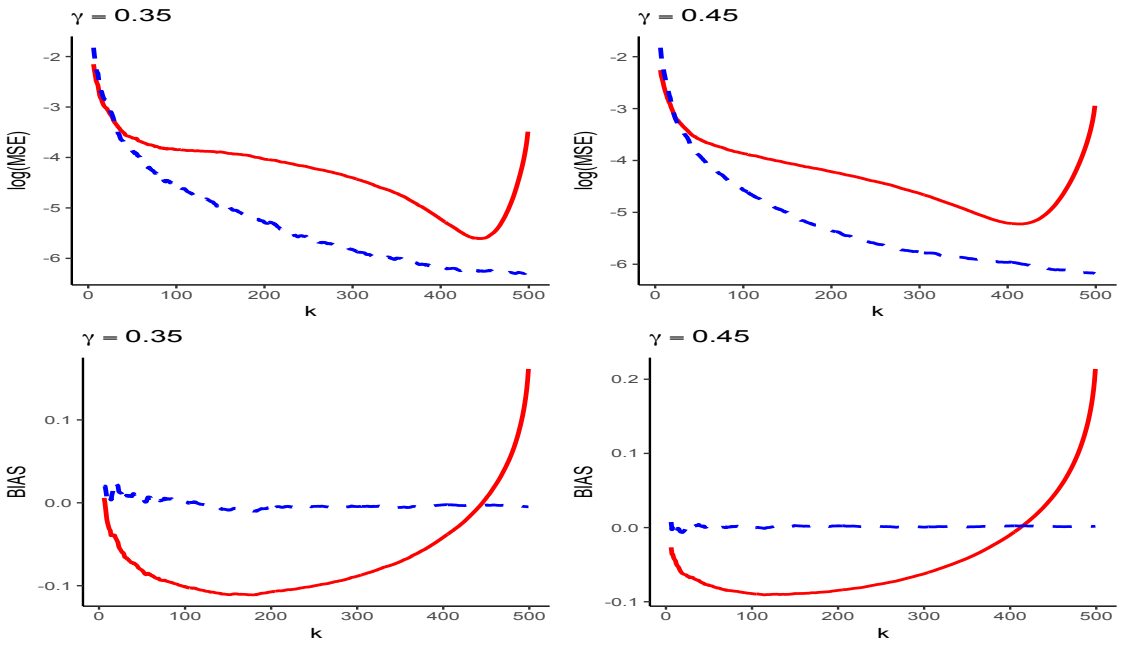



Figure 4: *MSE estimates (top panels) and Bias estimates (bottom panels) of* $\overline{\gamma}_k(\alpha)/\gamma$ *(solid red line) and* $\widehat{\gamma}_k/\gamma$ *(dashed blue line), as functions of* $k$*, for* $\gamma = 0.35$ *(left) and* $\gamma = 0.45$ *(right), in the case of a Pareto distribution.*

respectively, a quite respectable accuracy in terms of MSE for $\gamma = 0.35$ and $\gamma = 0.45$, while the purple curve ($\alpha = 0.5$) behaves clearly better in terms of Bias for both values of $\gamma$.

For $(\alpha, \beta) \in \{(0,0), (0,0.5), (0,1), (0.5,0), (0.5,0.5), (0.5,1), (1,0), (1,0.5), (1,1)\}$, Figures 7 and 8 give, respectively, the MSE (in log scale) and Bias estimates of $\overline{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)/\text{XES}_{\tau'_n}$, against $k$. It may be seen that the winner is the red curve ($\alpha = 0.5, \beta = 0$) in the case of Student distribution, the blue curve ($\alpha = 0.5, \beta = 1$) in the case of Fréchet distribution, and the black curve ($\alpha = \beta = 1$) in the case of Pareto distribution.

The Monte Carlo estimates for $\widehat{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)/\text{XES}_{\tau'_n}$ are displayed in Figures 9 and 10. It may be seen that the winner is the orange curve ($\alpha = 1, \beta = 0$) in the case of Student distribution, the blue curve ($\alpha = 0.5, \beta = 1$) in the case of Fréchet distribution, and the black curve ($\alpha = \beta = 1$) in the case of Pareto distribution.

Finally, when comparing the three estimators $\widetilde{\text{XES}}^{\star}_{\tau'_n}(\alpha)$, $\overline{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)$ and $\widehat{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)$ with each other, we arrive at the following tentative conclusions:

- In the case of Student distribution, the best $\widetilde{\text{XES}}^{\star}_{\tau'_n}(\alpha)$, achieved at $\alpha = 1$, is superior to the best $\widehat{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)$, achieved at $\alpha = 1$ and $\beta = 0$, which in turn is superior to the best $\overline{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)$, achieved at $\alpha = 0.5$ and $\beta = 0$.

- In the case of Fréchet distribution, the best $\widehat{\text{XES}}^{\star}_{\tau'_n}(\alpha,\beta)$, achieved at $\alpha = 0.5$ and

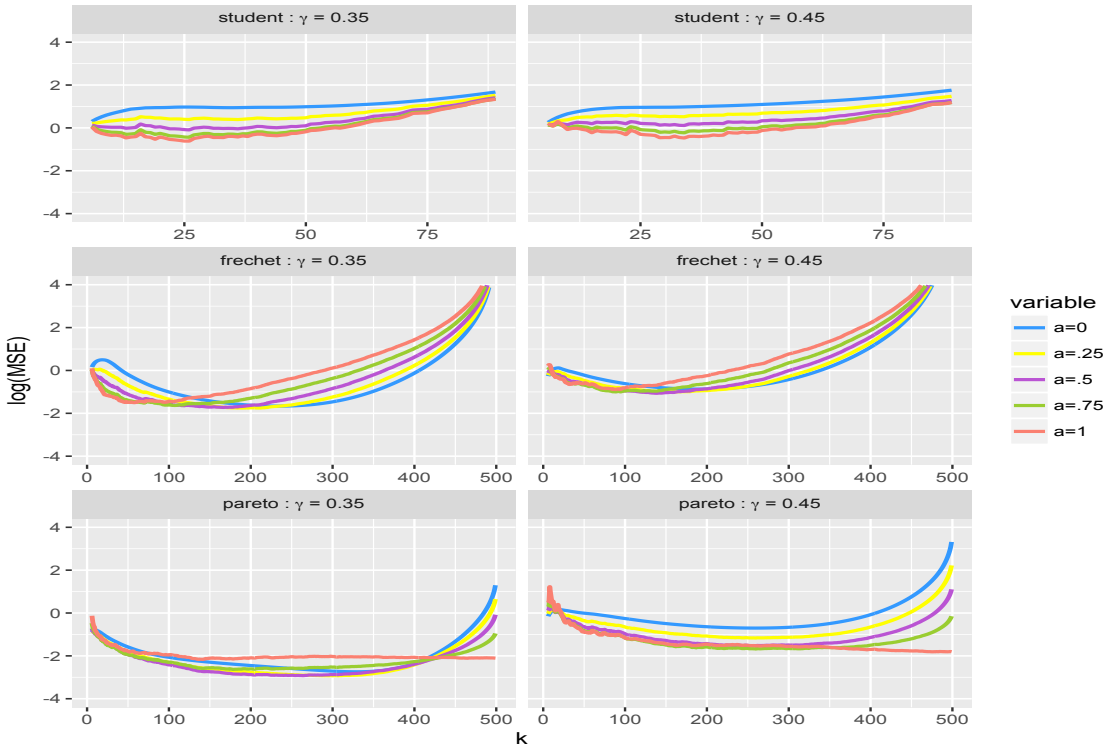



Figure 5: *MSE estimates (in log scale) of* $\widetilde{XES}^{\star}_{\tau'_n}(\alpha)/XES_{\tau'_n}$*, against* $k$*, for Student (top), Fréchet (middle) and Pareto (bottom) distributions, with* $\gamma = 0.35$ *(left) and* $\gamma = 0.45$ *(right).*

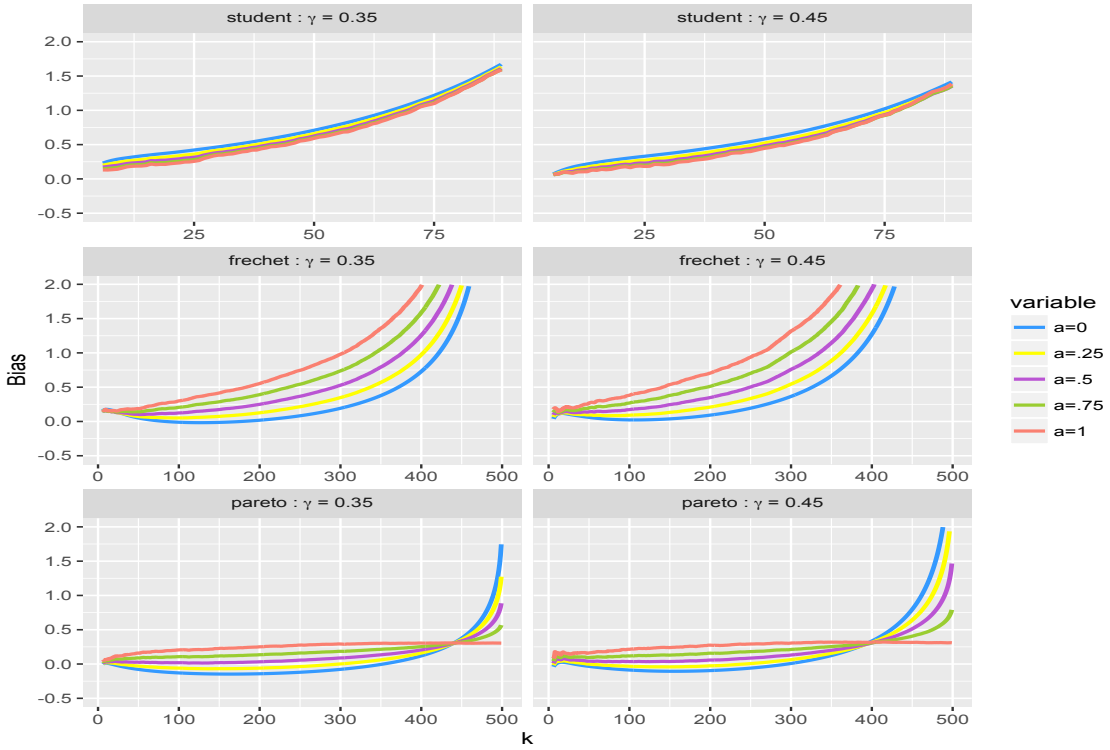



Figure 6: *Bias estimates of* $\widetilde{XES}^{\star}_{\tau'_n}(\alpha)/XES_{\tau'_n}$*.*

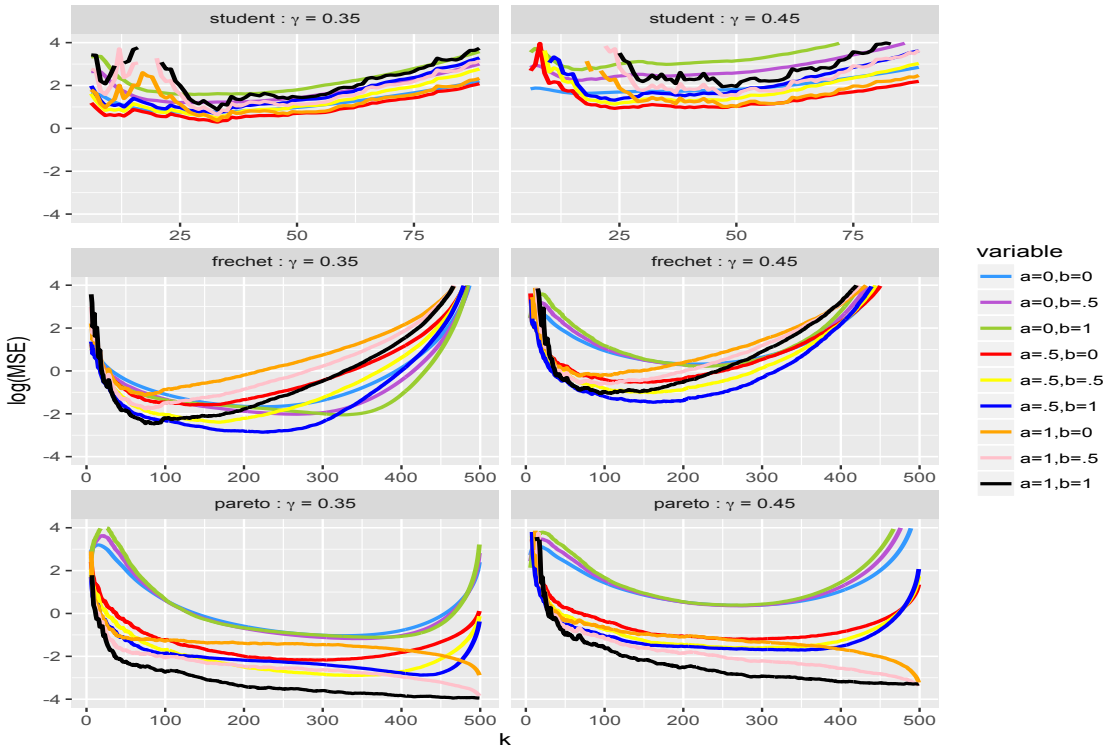



Figure 7: *MSE estimates (in log scale) of* $\overline{XES}^{\star}_{\tau'_n}(\alpha,\beta)/XES_{\tau'_n}$.

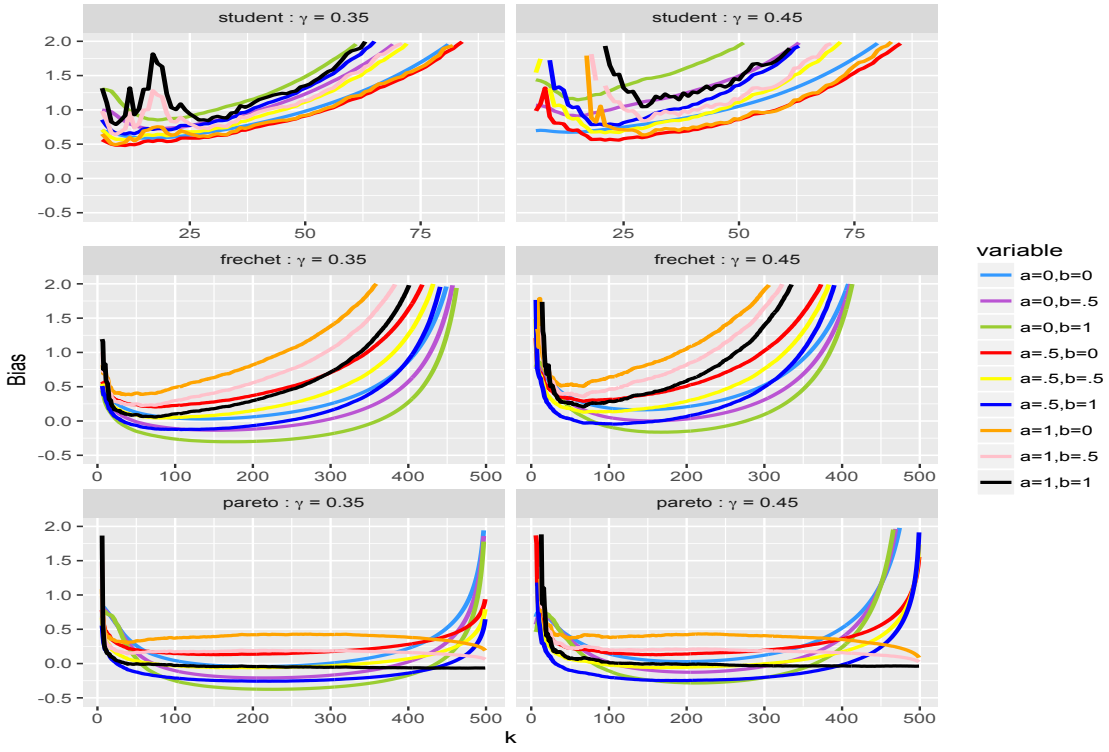



Figure 8: *Bias estimates of* $\overline{XES}^{\star}_{\tau'_n}(\alpha,\beta)/XES_{\tau'_n}$.

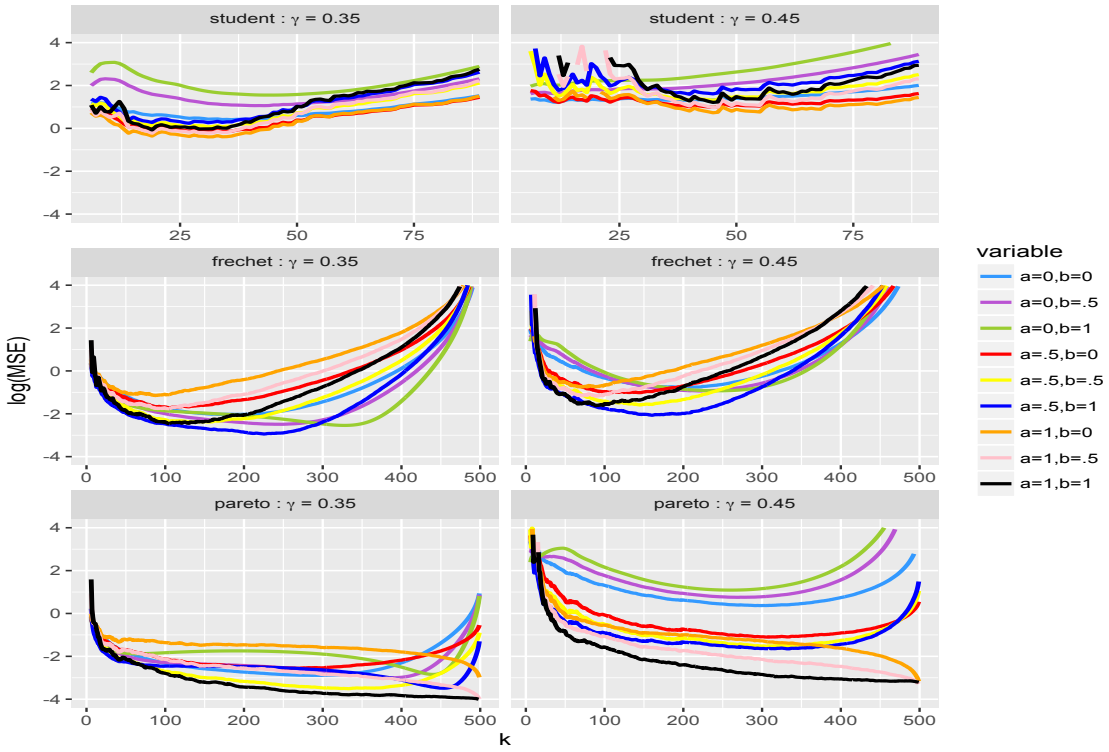



Figure 9: *MSE estimates (in log scale) of* $\widehat{XES}^{\star}_{\tau'_n}(\alpha,\beta)/XES_{\tau'_n}$.

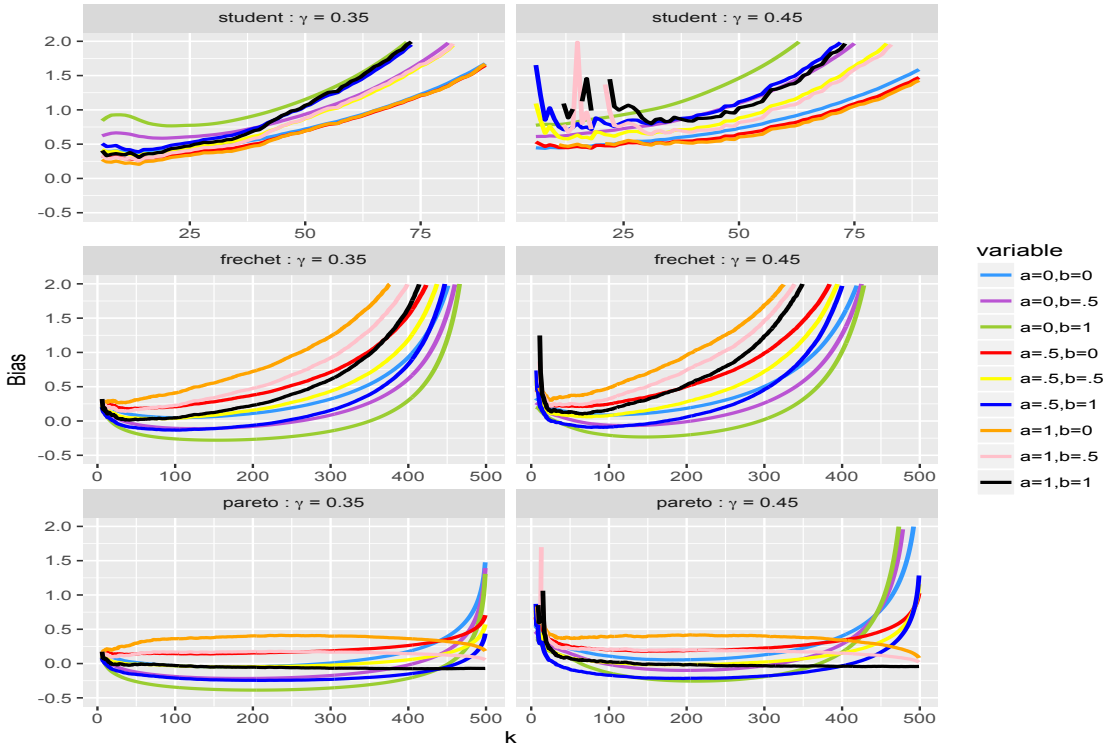



Figure 10: *Bias estimates of* $\widehat{XES}^{\star}_{\tau'_n}(\alpha,\beta)/XES_{\tau'_n}$.

$\beta = 1$, is superior but not by much to the best $\overline{\text{XES}}^{\star}_{\tau'_n}(\alpha, \beta)$, achieved at $\alpha = 0.5$ and $\beta = 1$, which in turn is superior to the best $\widetilde{\text{XES}}^{\star}_{\tau'_n}(\alpha)$ achieved at, say, $\alpha \in \{0, 0.5\}$.

- In the case of Pareto distribution, the best $\widehat{\text{XES}}^{\star}_{\tau'_n}(\alpha, \beta)$, achieved at $\alpha = \beta = 1$, is superior but not by much to the best $\overline{\text{XES}}^{\star}_{\tau'_n}(\alpha, \beta)$, achieved at $\alpha = \beta = 1$, which in turn is superior to the best $\widetilde{\text{XES}}^{\star}_{\tau'_n}(\alpha)$ achieved at, say, $\alpha \in \{0.5, 1\}$.

In particular, it seems that $\widetilde{\text{XES}}^{\star}_{\tau'_n}(\alpha)$ is the winner in the case of the real-valued profit-loss Student distribution for $\alpha = 1$, while $\widehat{\text{XES}}^{\star}_{\tau'_n}(\alpha, \beta)$ is most efficient in the case of the non-negative Fréchet and Pareto loss distributions, for $\alpha \in \{0.5, 1\}$ and $\beta = 1$.

## C.3 Estimates of $\text{QES}_{p_n}$

We have also undertaken simulation experiments to evaluate finite-sample performance of the composite versions $\overline{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, \beta)$, $\widehat{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, \beta)$ and $\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)$ studied in Theorem 11. These composite expectile-based estimators estimate the same conventional expected shortfall $\text{QES}_{p_n}$ as the direct quantile-based estimator $\widehat{\text{QES}}^{\star}_{p_n}(\alpha) \equiv \widehat{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, 1)$. We first examined the accuracy of each estimator for various values of $\alpha$ and $\beta$.

Figures 11 and 12 give the MSE (in log scale) and Bias estimates of $\widehat{\text{QES}}^{\star}_{p_n}(\alpha)/\text{QES}_{p_n}$. The results suggest the choice of $\alpha = 1$ (red curve) for Student and Pareto distributions, and $\alpha = 0.5$ (purple curve) for Fréchet distribution.

Figures 13 and 14 give the Monte Carlo estimates of $\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)/\text{QES}_{p_n}$. It may be seen that the choice of $\alpha = 0$ (blue curve) globally provides quite respectable behavior for the three distributions.

The Monte Carlo estimates of $\overline{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, \beta)/\text{QES}_{p_n}$ are displayed in Figures 15 and 16. It may be seen that the appropriate choice of $(\alpha, \beta)$ is, respectively, $(1, 0)$ for Student distribution (orange curve), $(0.5, 1)$ for Fréchet distribution (blue curve), and $(1, 1)$ for Pareto distribution (black curve).

The results for $\widehat{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, \beta)/\text{QES}_{p_n}$ are graphed in Figures 17 and 18. Here, it may be seen that the appropriate choice of $(\alpha, \beta)$ is, respectively, $(0, 0)$ for Student distribution (light blue), $(0.5, 1)$ for Fréchet distribution (heavy blue), and $(1, 1)$ for Pareto distribution (black curve).

Finally, when comparing the four estimators $\widehat{\text{QES}}^{\star}_{p_n}(\alpha)$, $\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)$, $\overline{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, \beta)$ and $\widehat{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha, \beta)$ with each other, we arrive at the following tentative conclusions:

- In the case of the real-valued profit-loss Student distribution, the best estimate seems to be $\widetilde{\text{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0)$;

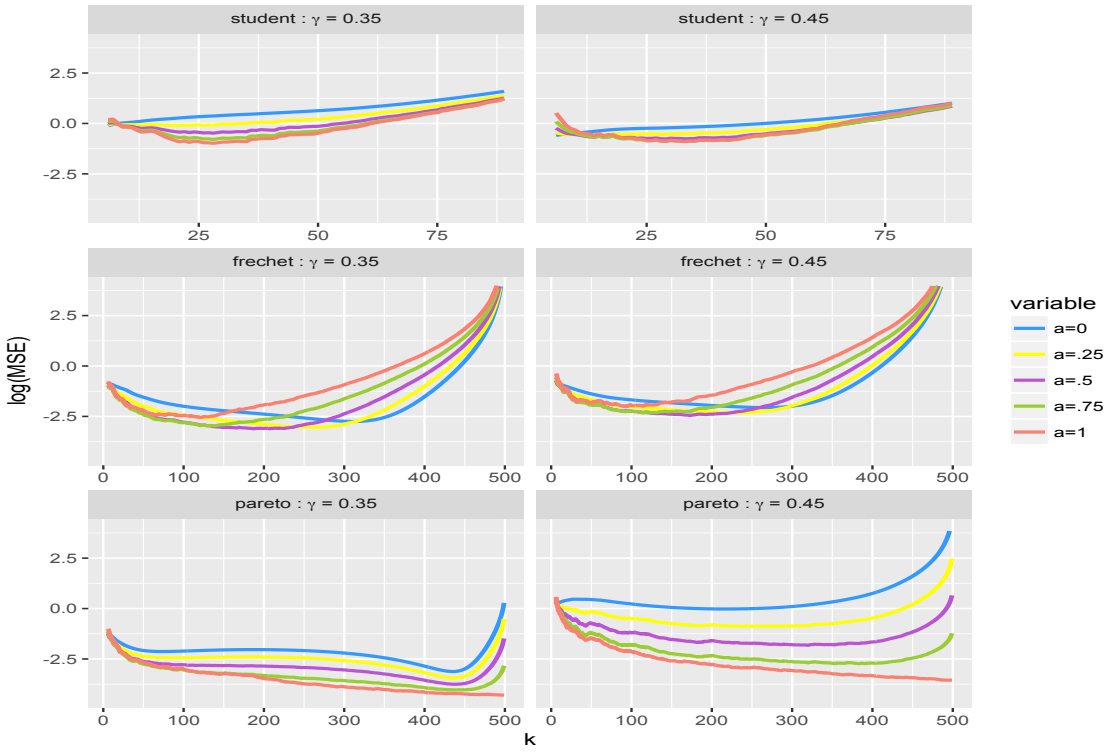



Figure 11: *MSE estimates (in log scale) of* $\widehat{QES}^{\star}_{p_n}(\alpha)/QES_{p_n}$*, against* $k$*, for Student (top), Fréchet (middle) and Pareto (bottom) distributions, with* $\gamma = 0.35$ *(left) and* $\gamma = 0.45$ *(right).*

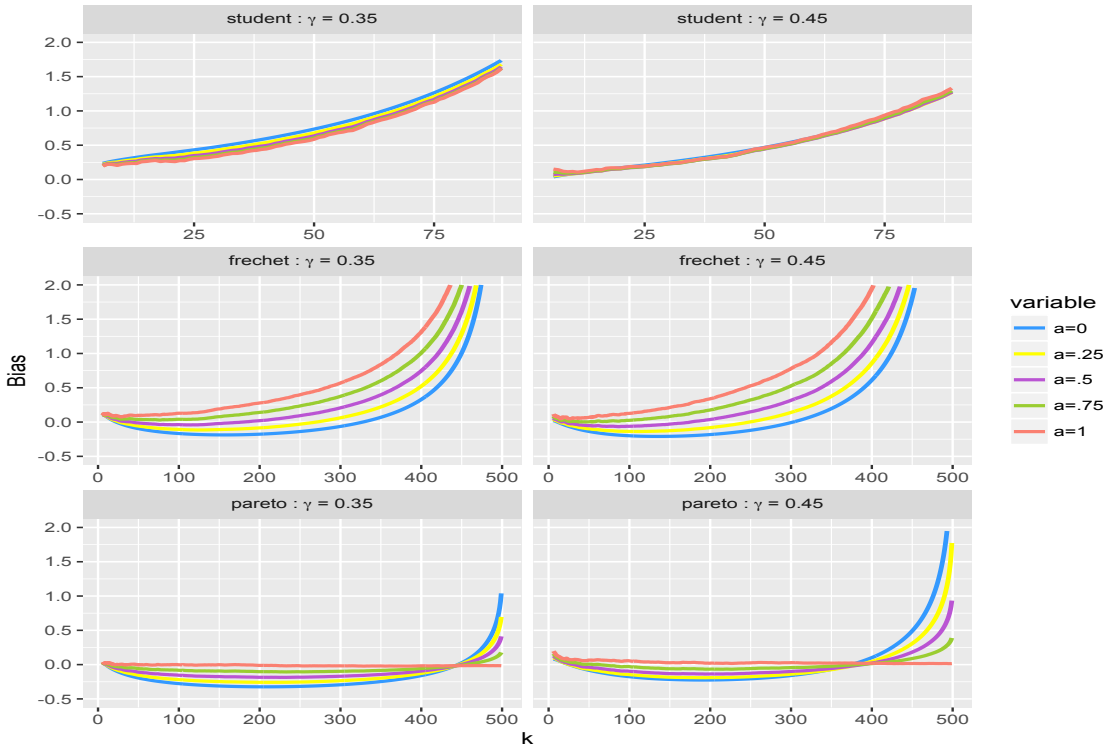



Figure 12: *Bias estimates of* $\widehat{QES}^{\star}_{p_n}(\alpha)/QES_{p_n}$.

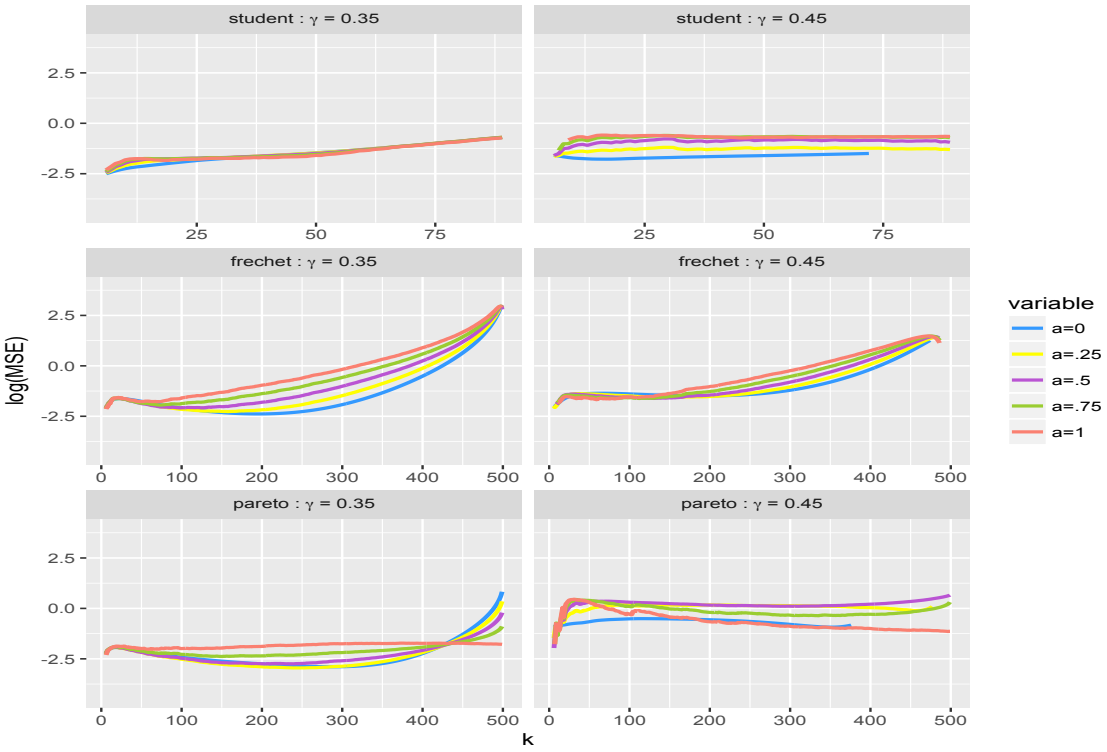



Figure 13: *MSE estimates (in log scale) of* $\widetilde{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)/QES_{p_n}$.

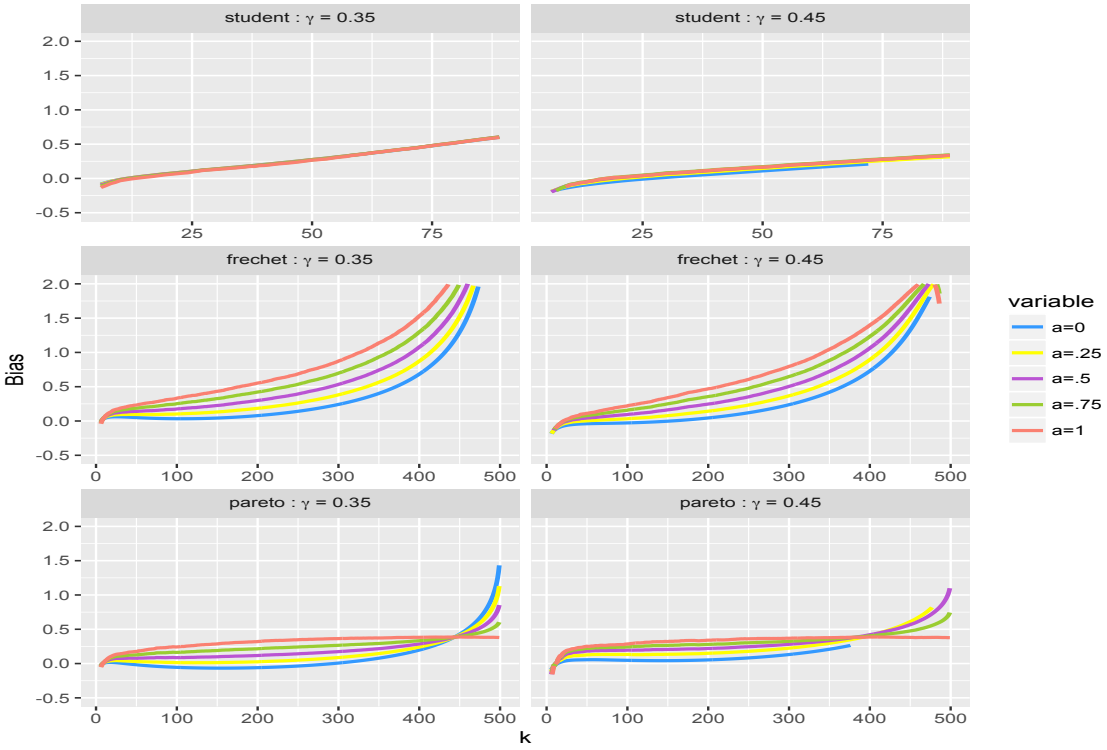



Figure 14: *Bias estimates of* $\widetilde{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)/QES_{p_n}$.

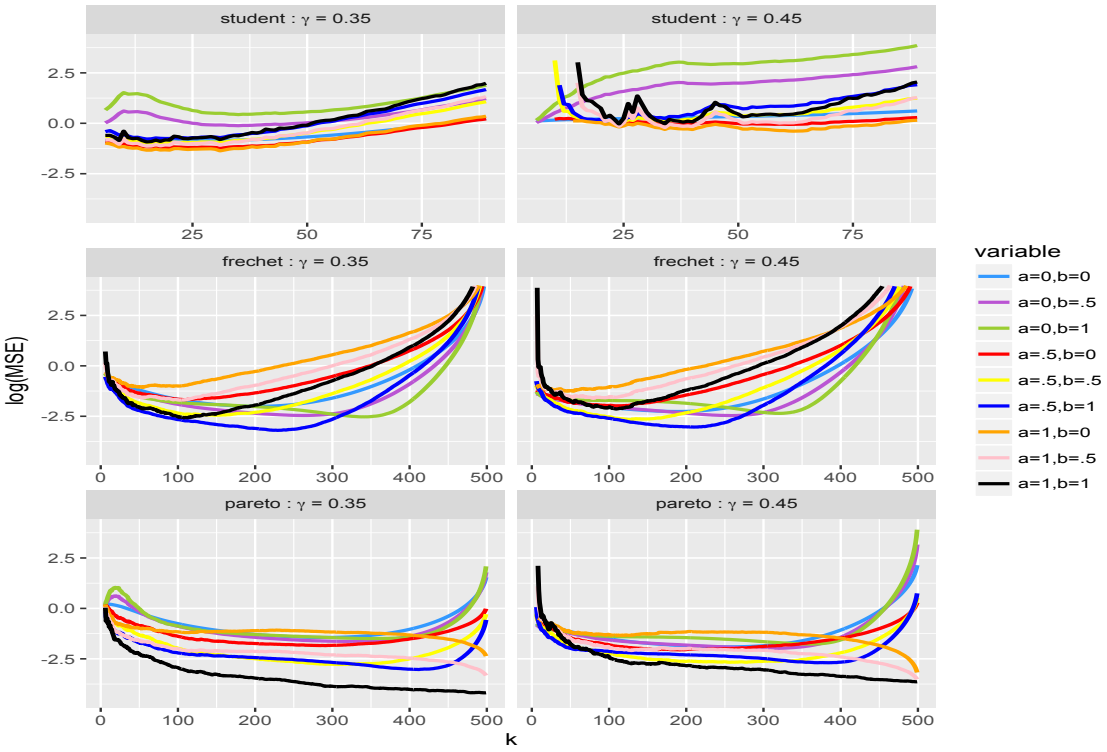



Figure 15: *MSE estimates (in log scale) of* $\overline{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)/QES_{p_n}$.

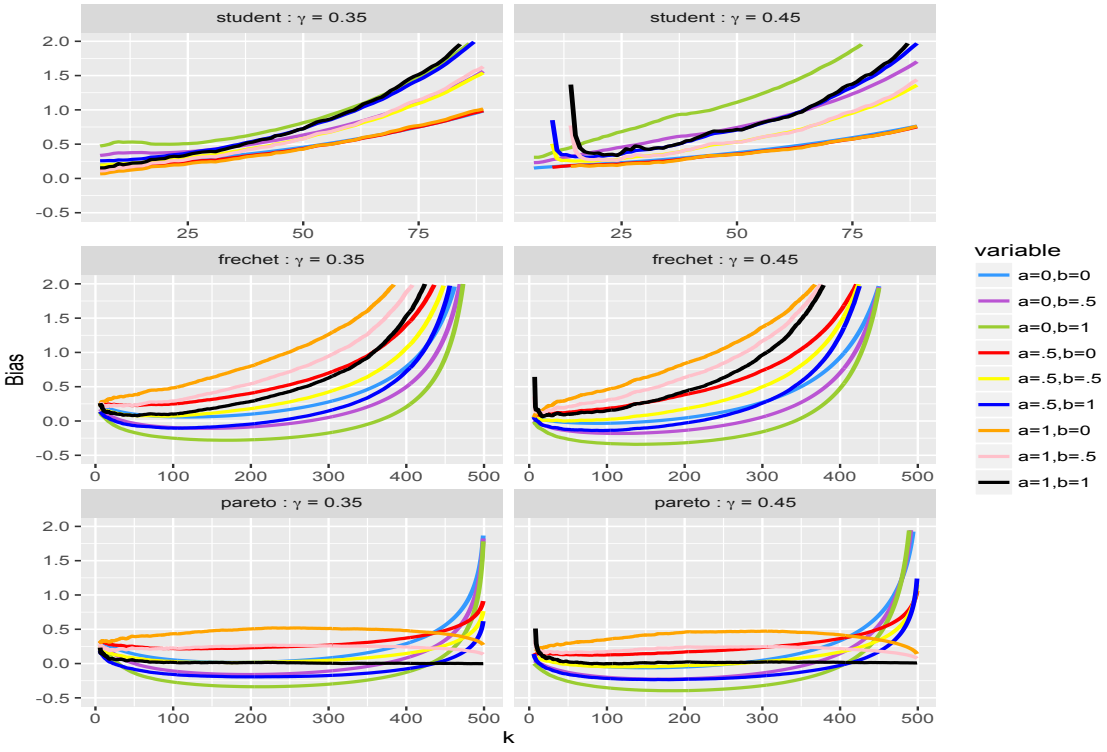



Figure 16: *Bias estimates of* $\overline{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)/QES_{p_n}$.

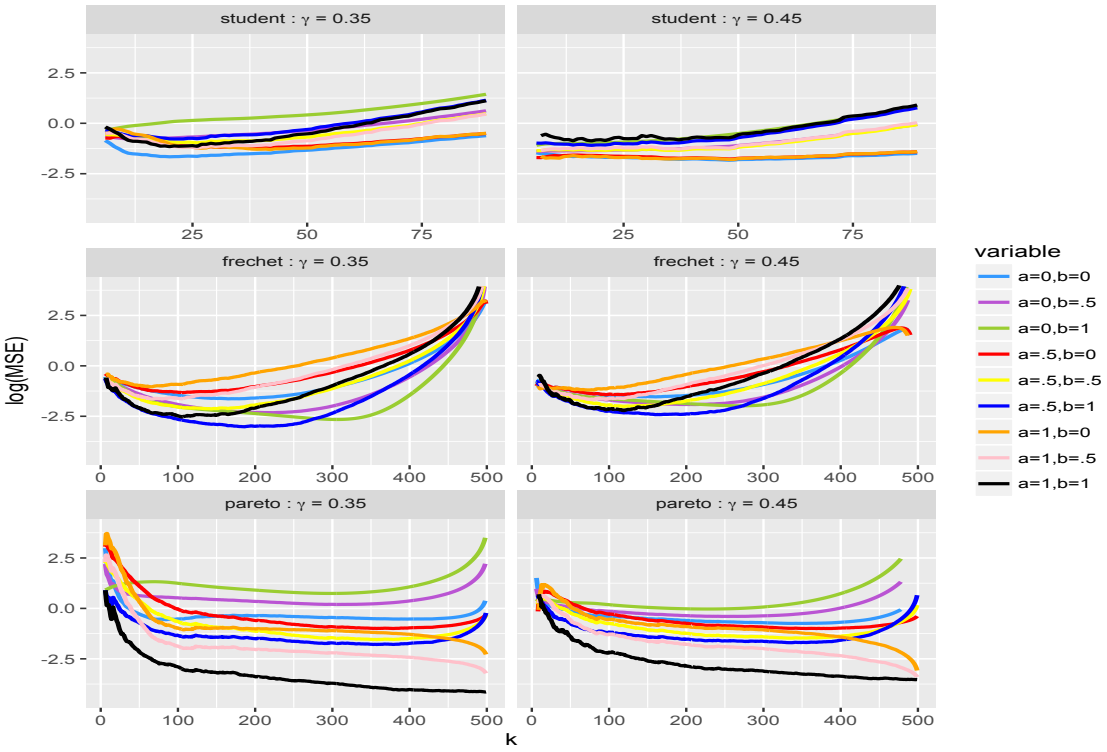



Figure 17: *MSE estimates (in log scale) of* $\widehat{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)/QES_{p_n}$.

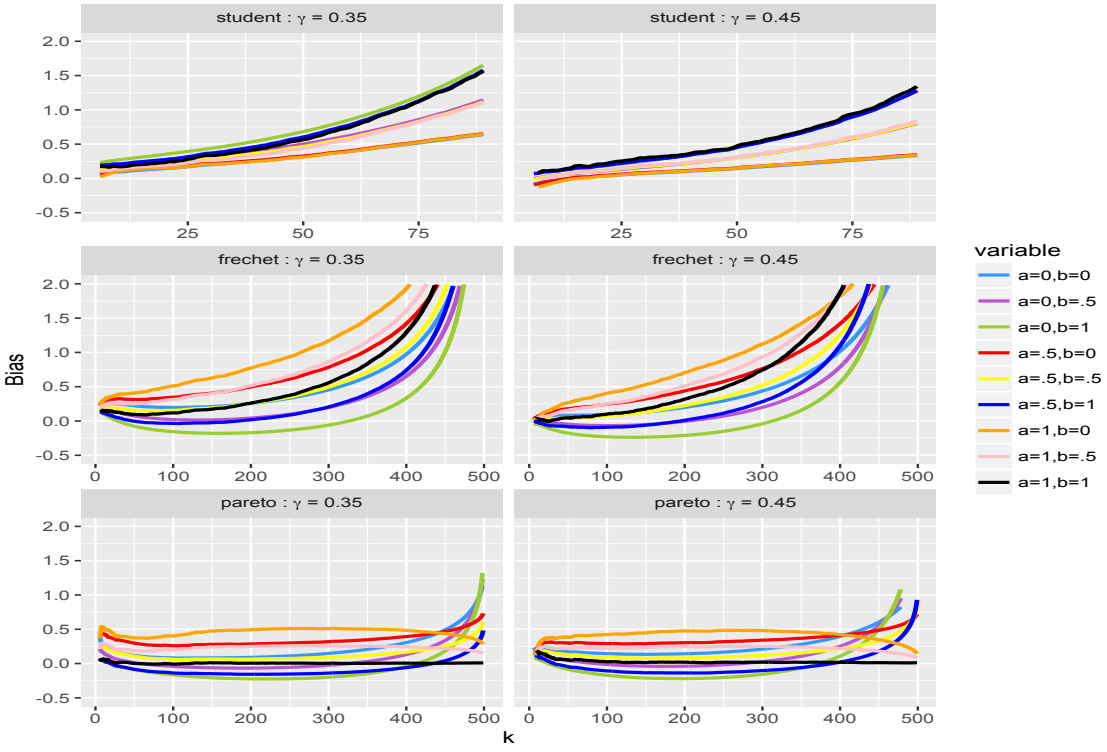



Figure 18: *Bias estimates of* $\widehat{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha,\beta)/QES_{p_n}$.

- In the case of the non-negative Fréchet and Pareto loss distributions, the best estimates seem to be $\overline{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0.5, \beta = 1)$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha = 1) \equiv \widehat{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 1, \beta = 1)$, respectively.

## C.4 Confidence intervals for $\mathrm{QES}_{p_n}$

We also investigated the performance of the three asymptotic confidence intervals described in Section 6.2.3, namely $\overline{\mathrm{CI}}_{0.95}(k)$, $\widehat{\mathrm{CI}}_{0.95}(k)$ and $\widetilde{\mathrm{CI}}_{0.95}(k)$. For the classical 95% confidence level, we take the value $z_{0.975} \approx 1.960$.

Figures 19 and 20 provide the average lengths and the achieved coverages of $\overline{\mathrm{CI}}_{0.95}(k)$ against $k$. In the case of Student distribution, the orange plots ($\alpha = 1, \beta = 0$) seem to provide better confidence intervals in terms of both average lengths and achieved coverages. In the case of Fréchet distribution, it may be seen that the black plots ($\alpha = 1, \beta = 1$) give a quite good global impression, though they are not clearly the winners in each scenario. In the case of Pareto distribution, it may be seen that the pink plots ($\alpha = 1, \beta = 0.5$) give a quite good global impression.

Figures 21 and 22 display the results for the asymptotic confidence intervals $\widehat{\mathrm{CI}}_{0.95}(k)$ against $k$. Our tentative conclusions are very similar to those obtained above for $\overline{\mathrm{CI}}_{0.95}(k)$.

The results for the asymptotic confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ are graphed in Figures 23 and 24 against $k$. Here, the red plots ($\alpha = 1$) seem to provide a reasonable global compromise between average lengths and achieved coverages.

Finally, when comparing the three 95% asymptotic confidence intervals $\overline{\mathrm{CI}}_{0.95}(k)$, $\widehat{\mathrm{CI}}_{0.95}(k)$ and $\widetilde{\mathrm{CI}}_{0.95}(k)$ with each other, it seems that

- $\widetilde{\mathrm{CI}}_{0.95}(k)$ performs better in the case of Student distribution, for the selected weight $\alpha = 1$;
- $\widehat{\mathrm{CI}}_{0.95}(k)$ performs quite well in the case of Fréchet distribution, for the selected weights $\alpha = 1$ and $\beta = 1$;
- $\overline{\mathrm{CI}}_{0.95}(k)$ performs quite well in the case of Pareto distribution, for the selected weights $\alpha = 1$ and $\beta = 0.5$.

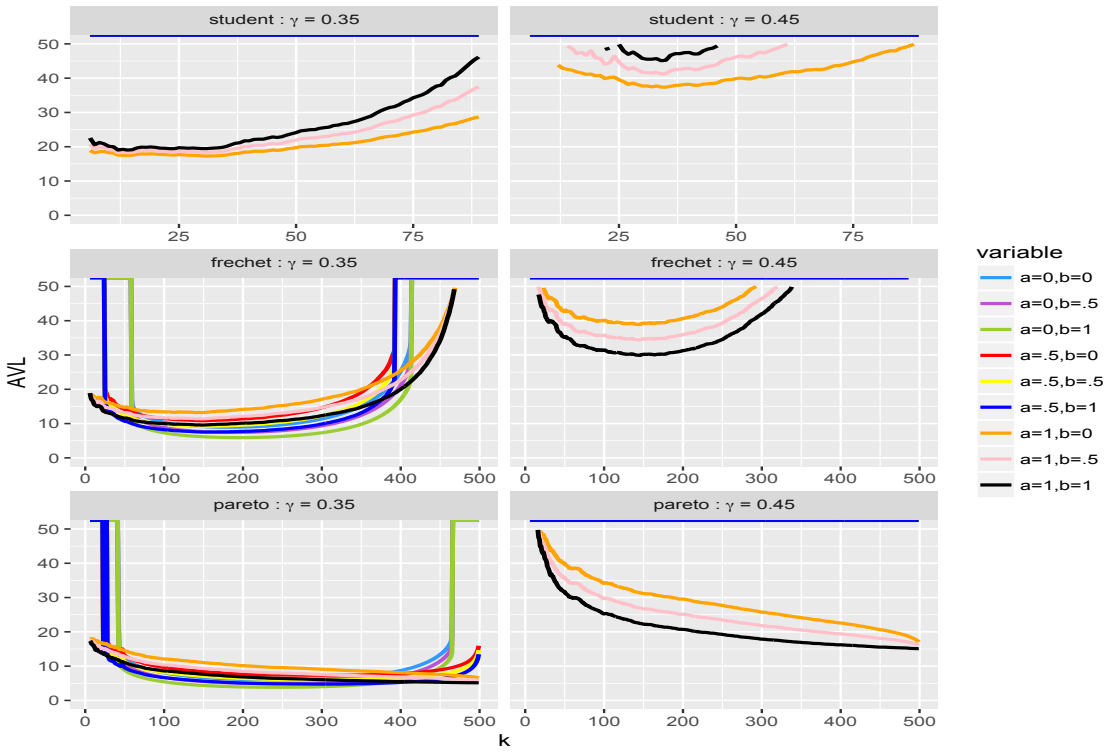



Figure 19: *Average lengths of* $\overline{CI}_{0.95}(k)$ *against* $k$.

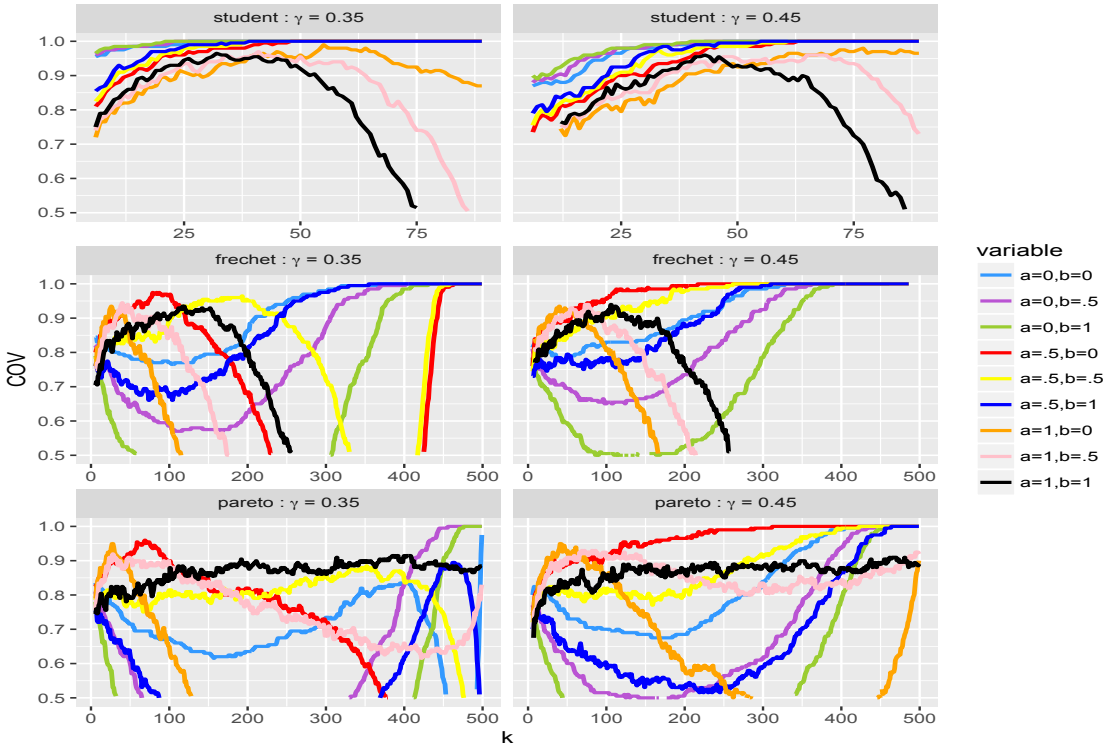



Figure 20: *Achieved coverages of* $\overline{CI}_{0.95}(k)$ *against* $k$.

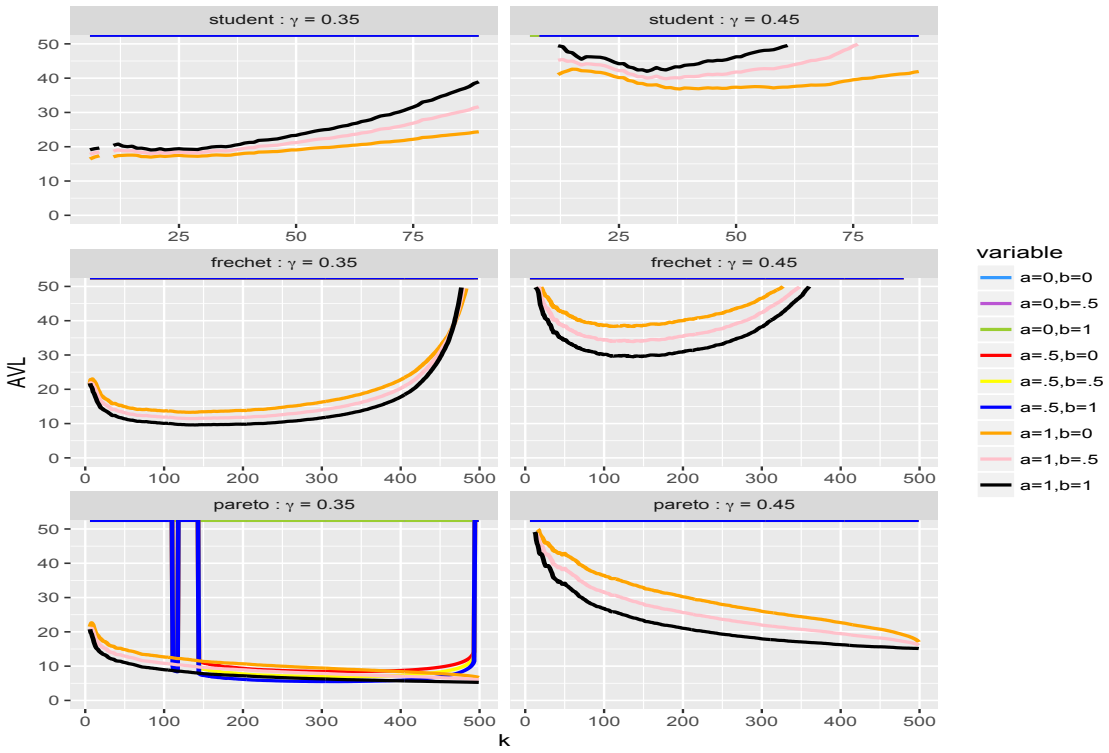



Figure 21: *Average lengths of* $\widehat{CI}_{0.95}(k)$ *against* $k$.

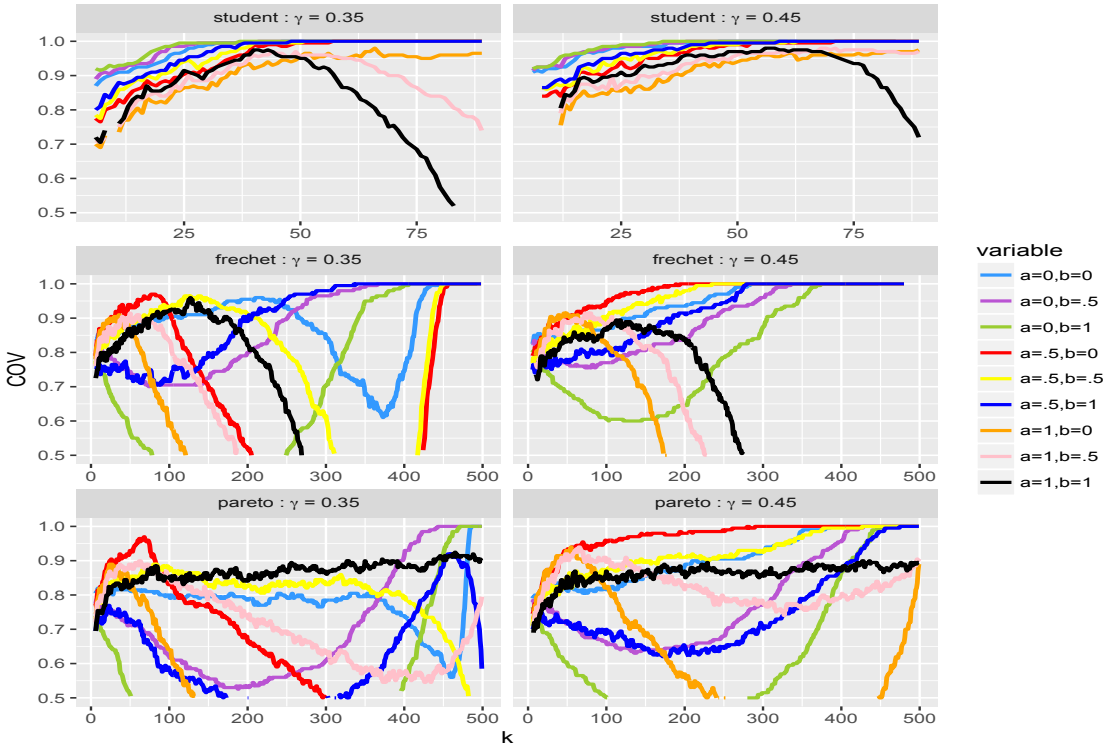



Figure 22: *Achieved coverages of* $\widehat{CI}_{0.95}(k)$ *against* $k$.

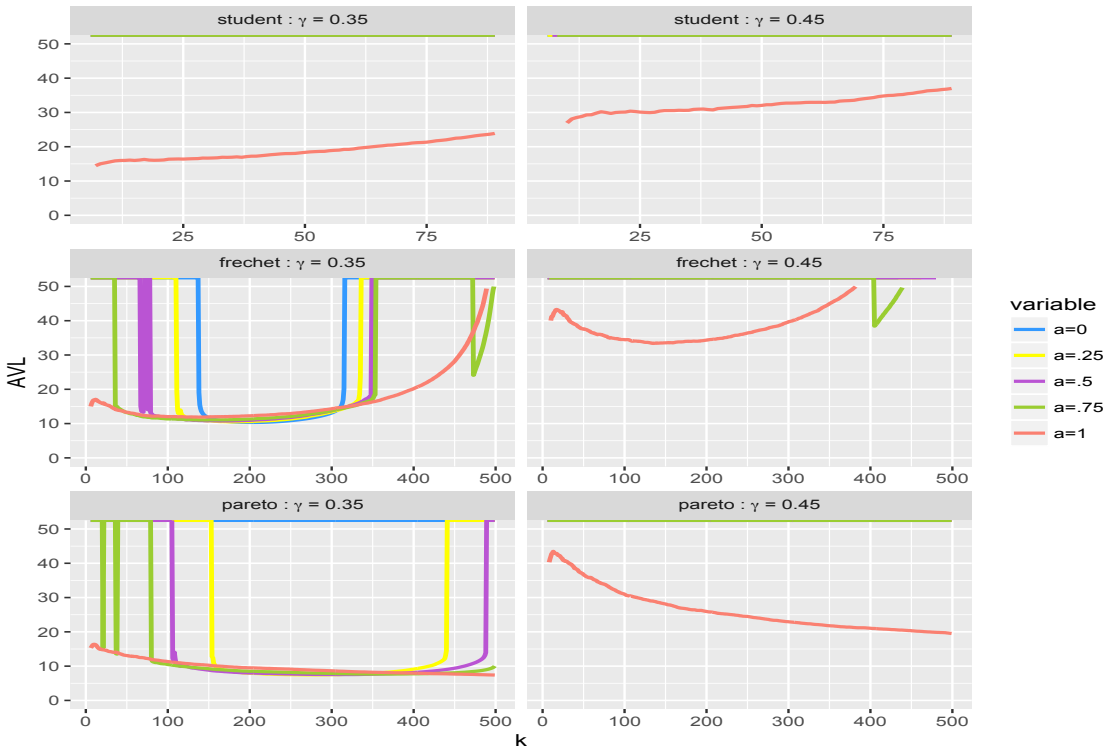



Figure 23: *Average lengths of* $\widetilde{CI}_{0.95}(k)$ *against* $k$.

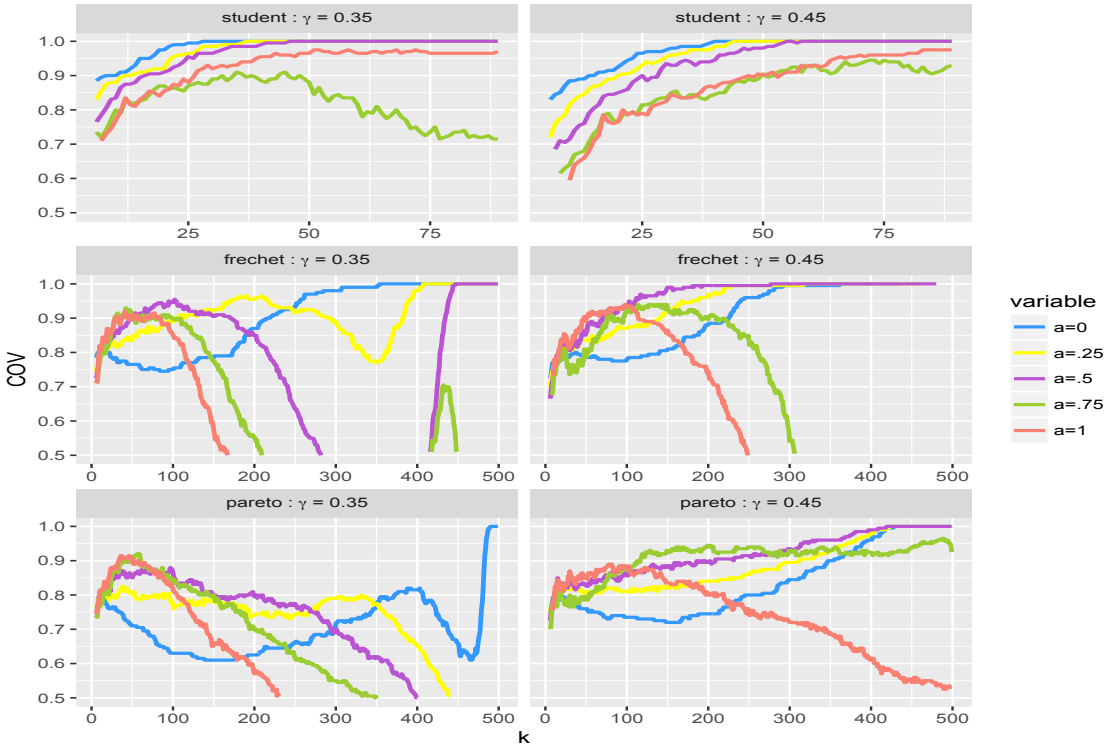



Figure 24: *Achieved coverages of* $\widetilde{CI}_{0.95}(k)$ *against* $k$.

# D ES for financial institutions

In this section, we apply our method to estimate the ES for three large US financial institutions. We consider the same investment banks as in Example 1, namely Goldman Sachs, Morgan Stanley and T. Rowe Price. All of these banks had a market capitalization greater than US $5 billion at the end of June 2007. The dataset consists of the negative log-returns $(Y_i)$ on their equity prices at a daily frequency during 10 years from July 3rd, 2000, to June 30th, 2010. The choice of the frequency of data and time horizon follows the same set-up as in Cai *et al.* (2015) and Daouia *et al.* (2018). This results in the sample size $n = 2513$. We use our composite expectile-based method to estimate the standard quantile-based expected shortfall $\mathrm{QES}_{p_n}$, or equivalently the expectile-based expected shortfall $\mathrm{XES}_{\tau'_n(p_n)}$, with an extreme relative frequency $p_n = 1 - \frac{1}{n}$ that corresponds to a once-per-decade rare event.

In this situation of real-valued profit-loss distributions, our experience with simulated data indicates that the composite estimator $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)$ provides the best $\mathrm{QES}_{p_n}$ estimates in terms of MSE and bias for the special weight $\alpha = 0$, while it provides reasonably good asymptotic 95% confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ for the different weight $\alpha = 1$. In the estimation, we employ the intermediate sequence $\tau_n = 1 - k/n$ as before, for the selected range of values $k = 1, \ldots, 150$. For our comparison purposes, we use as a benchmark the direct quantile-based estimator $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha = 1) \equiv \widehat{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 1, \beta = 1)$ of El Methni *et al.* (2014), as well as the corresponding asymptotic 95% confidence interval $\widehat{\mathrm{CI}}_{0.95}(k)$. We will denote in the sequel the rival estimates $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0)$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}(\alpha = 1)$ simply as $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}$.

For each bank, we superimpose in Figure 25 the plots of the two estimates $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and $\widehat{\mathrm{QES}}^{\star}_{p_n}$ against $k$, as rainbow and dashed black curves respectively, along with the competing confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ in dotted blue lines and $\widehat{\mathrm{CI}}_{0.95}(k)$ in solid grey lines. The effect of the *expectHill* estimate $\overline{\gamma}_{1-k/n}(\alpha = 0) \equiv \widetilde{\gamma}_{1-k/n}$ on the efficient estimate $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ is highlighted by a colour-scheme, ranging from dark red (low $\widetilde{\gamma}_{1-k/n}$) to dark violet (high $\widetilde{\gamma}_{1-k/n}$).

We have already provided some Monte Carlo evidence that the composite expectile-based estimates $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and confidence intervals $\widetilde{\mathrm{CI}}_{0.95}(k)$ are efficient and accurate relative to the pure quantile-based estimates $\widehat{\mathrm{QES}}^{\star}_{p_n}$ and confidence intervals $\widehat{\mathrm{CI}}_{0.95}(k)$, respectively. Their superiority in terms of plots' stability and confidence intervals' length can clearly be visualized in Figure 25 for the three banks. The final ES levels based on minimizing the standard deviations of the estimates, computed over a moving window covering 20% of the possible values of $k$, are reported in Table 2, along with the asymptotic 95% confidence intervals of the ES. Based on the reliable $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ estimates (in the second column), the ES

levels for Goldman Sachs and T. Rowe Price seem to be very close (around $-30\%$ to $-34\%$), whereas the ES level for Morgan Stanley is almost twice higher (around $-60\%$). The $\widehat{\mathrm{QES}}^{\star}_{p_n}$ estimates (in the fourth column) point also towards similar pessimistic results. The lower confidence bands (in third and fifth columns) are themselves quite conservative since they are almost equal to the maximum losses (in the last column) for the three banks.

The theory for our ES estimator $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ and for the estimator $\widehat{\mathrm{QES}}^{\star}_{p_n}$ of El Methni *et al.* (2014) is derived for independent and identically distributed random variables $Y_1, \ldots, Y_n$. For this application to financial returns, the potential serial dependence may then affect the estimation results. Similarly to our extreme value analysis under mixing conditions in Daouia *et al.* (2017), our convergence results may work under serial dependence with enlarged asymptotic variances. A practical solution already employed by Cai *et al.* (2015) to reduce substantially the potential serial dependence in this particular dataset is by using weekly loss returns in the same sample period (*i.e.* sums of the daily loss returns during each week). This results in a sample of size $n = 522$. The plots of the two estimates and the asymptotic 95% confidence intervals, against $k \in [1, 80]$, are superimposed in Figure 26 for the three banks, along with the new sample maxima. The final pointwise results are reported in Table 3. By comparing the obtained estimates for the daily and weekly losses, it may be seen that the results are qualitatively robust to the change from daily to weekly data. In particular, the $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ levels for Goldman Sachs and T. Rowe Price are still almost equal, while the estimated level for Morgan Stanley remains almost twice higher. Quantitatively, these ES estimates are much more conservative: around $-40\%$ to $-43\%$ for Goldman Sachs and T. Rowe Price, and around $-87\%$ for Morgan Stanley.

| *Bank* | $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ | $\widetilde{\mathrm{CI}}_{0.95}$ | $\widehat{\mathrm{QES}}^{\star}_{p_n}$ | $\widehat{\mathrm{CI}}_{0.95}$ | $Y_{n,n}$ |
|---|---|---|---|---|---|
| Goldman Sachs | 0.345 | (0.210, 0.506) | 0.393 | (0.235, 0.544) | 0.210 |
| Morgan Stanley | 0.598 | (0.376, 0.785) | 0.601 | (0.316, 0.984) | 0.299 |
| T. Rowe Price | 0.308 | (0.171, 0.411) | 0.301 | (0.177, 0.437) | 0.197 |

Table 2: *ES levels of the three investment banks, with the* 95% *confidence intervals and the sample maxima. Results based on daily loss returns, with* $n = 2513$ *and* $p_n = 1 - \frac{1}{n}$.

| *Bank* | $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}$ | $\widetilde{\mathrm{CI}}_{0.95}$ | $\widehat{\mathrm{QES}}^{\star}_{p_n}$ | $\widehat{\mathrm{CI}}_{0.95}$ | $Y_{n,n}$ |
|---|---|---|---|---|---|
| Goldman Sachs | 0.436 | (0.194, 0.620) | 0.495 | (0.226, 0.680) | 0.365 |
| Morgan Stanley | 0.874 | (0.384, 1.305) | 0.883 | (0.366, 1.478) | 0.904 |
| T. Rowe Price | 0.401 | (0.213, 0.511) | 0.407 | (0.216, 0.548) | 0.305 |

Table 3: *Results based on weekly loss returns, with* $n = 522$ *and* $p_n = 1 - \frac{1}{n}$.

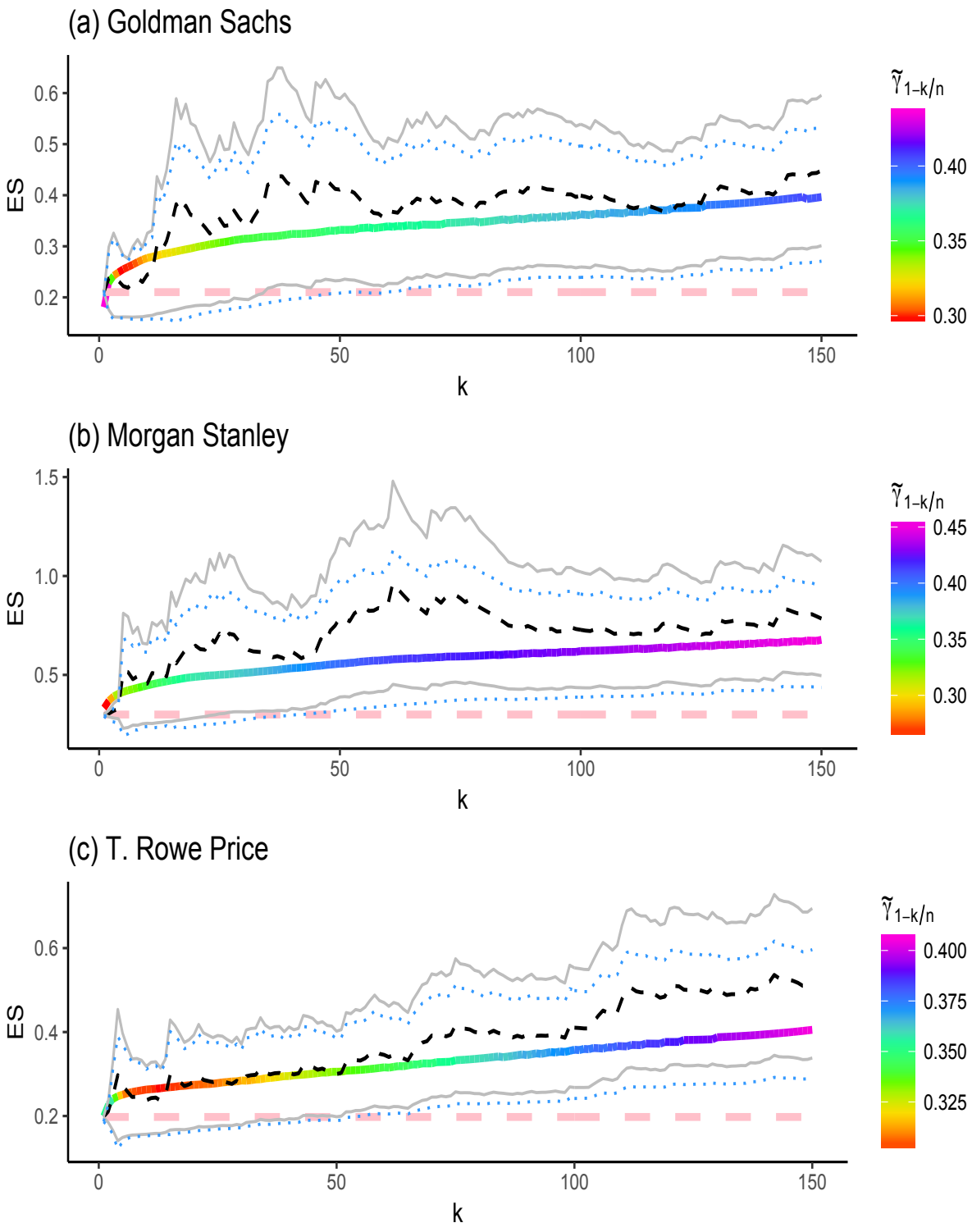



Figure 25: *Results based on daily loss returns of the three investment banks: (a) Goldman Sachs, (b) Morgan Stanley, and (c) T. Rowe Price, with* $n = 2513$ *and* $p_n = 1 - 1/n$. *The estimates* $\widetilde{XES}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha = 0)$ *as rainbow curve and* $\widehat{QES}^{\star}_{p_n}(\alpha = 1)$ *as dashed black curve, along with the asymptotic* 95% *confidence intervals* $\widetilde{CI}_{0.95}(k)$ *in dotted blue lines and* $\widehat{CI}_{0.95}(k)$ *in solid grey lines. The sample maximum* $Y_{n,n}$ *indicated in horizontal dashed pink line.*

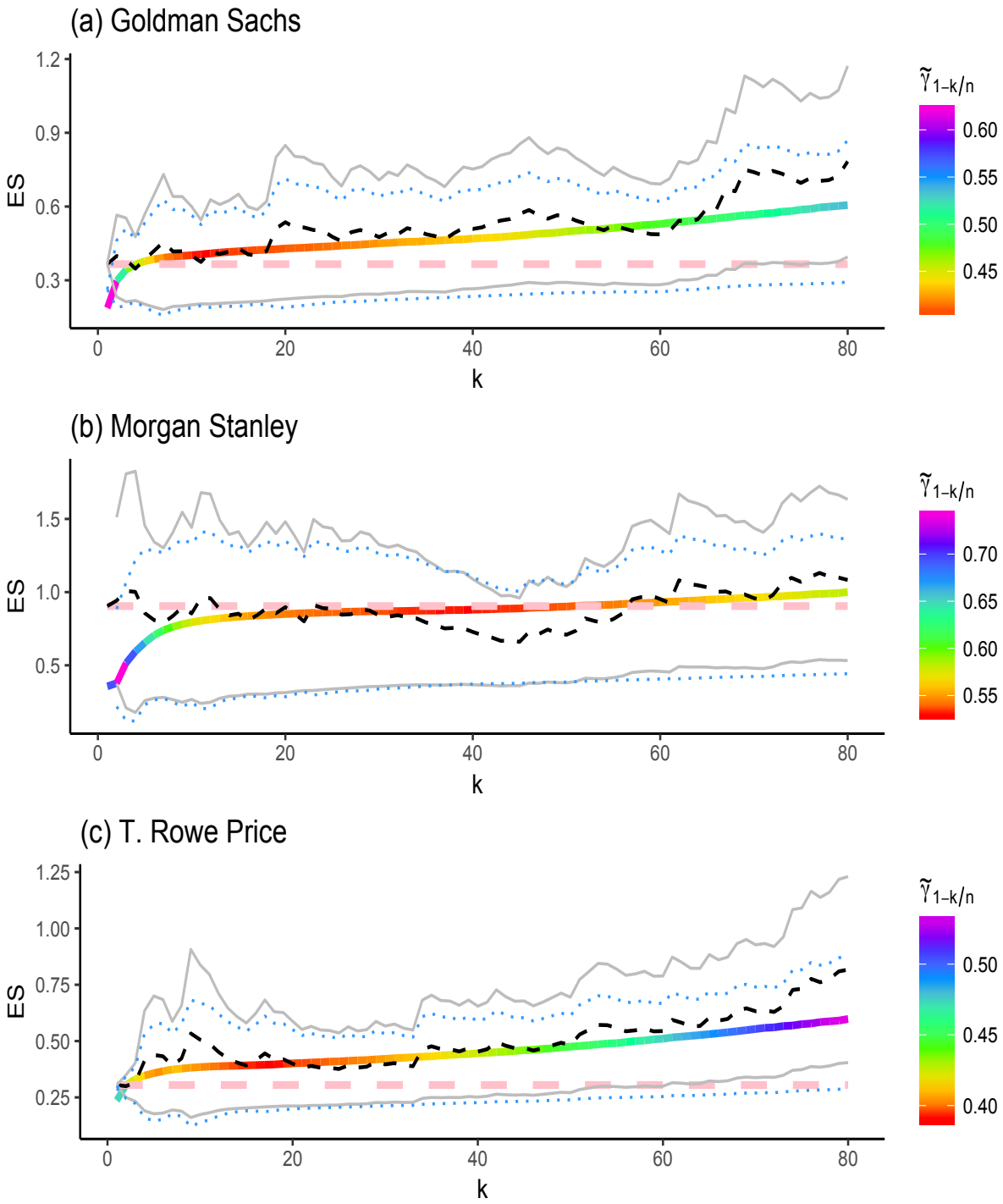



Figure 26: *Results based on weekly loss returns of the three investment banks, with* $n = 522$ *and* $p_n = 1 - 1/n$.

# E Proofs

In all proofs, the sequence $\tau_n$ is replaced by the sequence $k = n(1-\tau_n)$.

**Proof of Proposition 1.** We start by showing (i). By Proposition 1 in Daouia *et al.* (2018):

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

Using this convergence together with local uniformity of condition $\mathcal{C}_2(\gamma,\rho,A)$, we find that

$$\frac{1}{A((1-\tau)^{-1})}\left[\frac{U(1/\overline{F}(\xi_\tau))}{U((1-\tau)^{-1})} - (\gamma^{-1}-1)^{-\gamma}(1+\varepsilon(\tau))^{-\gamma}\right] \to (\gamma^{-1}-1)^{-\gamma}\frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho}$$

as $\tau \to 1$, or equivalently

$$\begin{aligned}\frac{U(1/\overline{F}(\xi_\tau))}{q_\tau} &= (\gamma^{-1}-1)^{-\gamma}\left(1+\frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_\tau}(\mathbb{E}(Y)+\mathrm{o}(1))\right. \\ &\quad \left. + \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho} + \mathrm{o}(1)\right)A((1-\tau)^{-1})\right) \text{ as } \tau \to 1.\end{aligned}$$

A use of Lemma 1 at $t = \xi_\tau$ makes it possible to replace $U(1/\overline{F}(\xi_\tau))$ by $\xi_\tau$ asymptotically, thus completing the proof of (i).

To show (ii), first note that if $s = 1$, there is nothing to prove. Otherwise, write

$$\frac{\xi_{1-ks/n}}{\xi_{1-k/n}} = \frac{\xi_{1-ks/n}}{q_{1-ks/n}} \times \frac{q_{1-k/n}}{\xi_{1-k/n}} \times \frac{q_{1-ks/n}}{q_{1-k/n}}. \tag{E.2}$$

With alternatively $\tau = 1-k/n$ and $\tau = 1-ks/n$ in (i), we obtain

$$\begin{aligned}\frac{\xi_{1-k/n}}{q_{1-k/n}} &= (\gamma^{-1}-1)^{-\gamma}\left(1+\frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_{1-k/n}}(\mathbb{E}(Y)+\mathrm{o}(1))\right. \\ &\quad \left. + \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho} + \mathrm{o}(1)\right)A(n/k)\right)\end{aligned}$$

and

$$\begin{aligned}\frac{\xi_{1-ks/n}}{q_{1-ks/n}} &= (\gamma^{-1}-1)^{-\gamma}\left(1+s^{\gamma}\frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_{1-k/n}}(\mathbb{E}(Y)+\mathrm{o}(1))\right. \\ &\quad \left. + \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho} + \mathrm{o}(1)\right)s^{-\rho}A(n/k)\right)\end{aligned}$$

because of the regular variation property of $t \mapsto q_{1-t^{-1}}$ and $|A|$. Besides, it is a consequence of condition $\mathcal{C}_2(\gamma,\rho,A)$ that

$$\frac{q_{1-ks/n}}{q_{1-k/n}} = \frac{U(n/ks)}{U(n/k)} = s^{-\gamma}\left(1+A(n/k)\frac{s^{-\rho}-1}{\rho} + \mathrm{o}(A(n/k))\right).$$

Combining these three expansions with (E.2) yields the desired result. ∎

**Proof of Theorem 2.** We use Lemma 2 with $\tau = 1 - ks/n$, $s \in (0,1]$, in order to write

$$\frac{ks/n}{1-2ks/n}(\widetilde{\xi}_{1-ks/n} - \overline{Y}_n) = \int_{\widetilde{\xi}_{1-ks/n}}^{\infty} \widehat{\overline{F}}_n(u)du \tag{E.3}$$

where $\overline{Y}_n$ denotes the empirical mean and $\widehat{\overline{F}}_n(u) = n^{-1}\sum_{i=1}^n \mathbb{1}_{\{Y_i>u\}}$ is the empirical survival function of the sample. The idea is now to obtain a uniform (in $s$) "asymptotic expansion" of the integral on the right-hand side.

Our main tool will be Lemma 3(ii): we may enlarge the underlying sample space and choose a suitable version of the empirical process $\widehat{F}_n$ so that there is a sequence of standard Brownian motions $\widetilde{W}_n$ such that for any $\varepsilon > 0$ small enough (which we shall fix later):

$$\frac{n}{k}\widehat{\overline{F}}_n\left(xq_{1-k/n}\right) - x^{-1/\gamma} = \frac{1}{\sqrt{k}}\left(\widetilde{W}_n(x^{-1/\gamma}) + \sqrt{k}A(n/k)x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho} + x^{(\varepsilon-1/2)/\gamma}\,\mathrm{o}_{\mathbb{P}}(1)\right)$$

uniformly in half-lines of the form $x \in [x_0, \infty)$, for $x_0 > 0$. Note then that, as a consequence of the monotonicity of expectiles together with convergence

$$\frac{\xi_\tau}{q_\tau} \to \left(\gamma^{-1}-1\right)^{-\gamma} \quad \text{as} \quad \tau \to 1 \tag{E.4}$$

(see Bellini and Di Bernardino, 2017) and Lemma 4, we have

$$\forall s \in (0,1],\ \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} \geqslant \frac{\widetilde{\xi}_{1-k/n}}{q_{1-k/n}} \xrightarrow{\mathbb{P}} (\gamma^{-1}-1)^{-\gamma} \quad \text{as} \quad n \to \infty.$$

Consequently

$$\mathbb{P}\left(\forall s \in (0,1],\ \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} > \frac{1}{2}(\gamma^{-1}-1)^{-\gamma}\right) \to 1 \quad \text{as} \quad n \to \infty. \tag{E.5}$$

It then follows from the above approximation by a sequence of Brownian motions that, with arbitrarily large probability:

$$\begin{aligned}
&\int_{\widetilde{\xi}_{1-ks/n}}^{\infty} \widehat{\overline{F}}_n(u)du \\
&= q_{1-k/n}\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} \widehat{\overline{F}}_n(xq_{1-k/n})dx \\
&= \frac{k}{n}q_{1-k/n}\left(\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{-1/\gamma}\,dx + \frac{1}{\sqrt{k}}\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\,dx\right. \\
&+ \left. A(n/k)\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho}\,dx + \mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{(\varepsilon-1/2)/\gamma}\,dx\right)\right)
\end{aligned} \tag{E.6}$$

uniformly in $s \in (0,1]$. Note that the last term is indeed well-defined, if $\varepsilon$ is taken close enough to 0, because $\gamma \in (0,1/2)$. We choose such an $\varepsilon$ here and in the sequel.

The next step is to use Lemma 5, primarily to remove the randomness in the lower bound of the integral of the Brownian motion $\widetilde{W}_n$ in (E.6). Lemma 5 only allows us to do so on the restricted range $[k^{-1+\delta},1]$, and we therefore focus on this case for now; we will take care of the case $s \in (0,k^{-1+\delta})$ separately afterwards. Use first (E.5) to get, for any sufficiently small $\delta > 0$ and with arbitrarily large probability irrespective of $s \in [k^{-1+\delta},1]$:

$$\begin{aligned} & s^{\gamma} \left| \int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\,dx - \int_{(\gamma^{-1}-1)^{-\gamma}s^{-\gamma}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\,dx \right| \\ = & \left| \int_{s^{\gamma}\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} \widetilde{W}_n(s\,u^{-1/\gamma})\,du - \int_{(\gamma^{-1}-1)^{-\gamma}}^{\infty} \widetilde{W}_n(s\,u^{-1/\gamma})\,du \right| \\ \leqslant & s^{\gamma} \left| \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} - (\gamma^{-1}-1)^{-\gamma}s^{-\gamma} \right| \times \sup_{0 \leqslant t \leqslant (\gamma^{-1}-1)/2^{-1/\gamma}} |\widetilde{W}_n(st)|. \end{aligned}$$

Self-similarity of the Brownian motion $\widetilde{W}_n$ w.r.t. scaling gives

$$\sup_{0 \leqslant t \leqslant (\gamma^{-1}-1)/2^{-1/\gamma}} |\widetilde{W}_n(st)| \stackrel{d}{=} \sqrt{s} \sup_{0 \leqslant t \leqslant (\gamma^{-1}-1)/2^{-1/\gamma}} |\widetilde{W}_n(t)| = \mathrm{O}_{\mathbb{P}}(\sqrt{s})$$

uniformly in $s$, because a standard Brownian motion is almost surely bounded on any compact interval by almost sure continuity of its sample paths. A use of Lemma 5 then entails

$$\sup_{k^{-1+\delta} \leqslant s \leqslant 1} s^{\gamma-1/2} \left| \int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\,dx - \int_{(\gamma^{-1}-1)^{-\gamma}s^{-\gamma}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\,dx \right| = \mathrm{o}_{\mathbb{P}}(1).$$

Similarly,

$$\begin{aligned} \sup_{k^{-1+\delta} \leqslant s \leqslant 1} s^{\gamma-1} \left| \int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{-1/\gamma} \frac{x^{\rho/\gamma}-1}{\gamma\rho}\,dx - \int_{(\gamma^{-1}-1)^{-\gamma}s^{-\gamma}}^{\infty} x^{-1/\gamma} \frac{x^{\rho/\gamma}-1}{\gamma\rho}\,dx \right| &= \mathrm{o}_{\mathbb{P}}(1) \\ \text{and} \quad \sup_{k^{-1+\delta} \leqslant s \leqslant 1} s^{\gamma-1/2+\varepsilon} \left| \int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{(\varepsilon-1/2)/\gamma}\,dx - \int_{(\gamma^{-1}-1)^{-\gamma}s^{-\gamma}}^{\infty} x^{(\varepsilon-1/2)/\gamma}\,dx \right| &= \mathrm{o}_{\mathbb{P}}(1). \end{aligned}$$

Therefore, we have, uniformly in $s \in [k^{-1+\delta}, 1]$ and with arbitrarily large probability, that

$$\begin{aligned}
&\int_{\widetilde{\xi}_{1-ks/n}}^{\infty} \widehat{\overline{F}}_n(u)du \\
&= \frac{k}{n} q_{1-k/n} \left( \frac{\gamma}{1-\gamma} \left[ \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} \right]^{1-1/\gamma} + \frac{1}{\sqrt{k}} \int_{(\gamma^{-1}-1)^{-\gamma} s^{-\gamma}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\, dx \right. \\
&+ A(n/k) \int_{(\gamma^{-1}-1)^{-\gamma} s^{-\gamma}}^{\infty} x^{-1/\gamma} \frac{x^{\rho/\gamma}-1}{\gamma\rho}\, dx + \mathrm{o}_{\mathbb{P}}\left( \frac{1}{\sqrt{k}} \int_{(\gamma^{-1}-1)^{-\gamma} s^{-\gamma}}^{\infty} x^{(\varepsilon-1/2)/\gamma}\, dx \right) \\
&\left. + \mathrm{o}_{\mathbb{P}}\left( \frac{s^{-\gamma+1/2-\varepsilon}}{\sqrt{k}} \right) \right).
\end{aligned}$$

We now rewrite each integral as follows: firstly, a change of variables and self-similarity of the Brownian motion w.r.t. scaling yield

$$\begin{aligned}
\int_{(\gamma^{-1}-1)^{-\gamma} s^{-\gamma}}^{\infty} \widetilde{W}_n(x^{-1/\gamma})\, dx &= (\gamma^{-1}-1)^{-\gamma}\gamma \int_0^s \widetilde{W}_n((\gamma^{-1}-1)t)\, t^{-\gamma-1}\, dt \\
&= (\gamma^{-1}-1)^{1/2-\gamma}\gamma \int_0^s W_n(t)\, t^{-\gamma-1}\, dt \qquad \text{(E.7)}
\end{aligned}$$

where $W_n(t) := (\gamma^{-1}-1)^{-1/2}\widetilde{W}_n((\gamma^{-1}-1)t)$ defines another sequence of standard Brownian motions. Secondly, a straightforward integration gives

$$\int_{(\gamma^{-1}-1)^{-\gamma} s^{-\gamma}}^{\infty} x^{-1/\gamma} \frac{x^{\rho/\gamma}-1}{\gamma\rho}\, dx = \frac{(\gamma^{-1}-1)^{1-\gamma} s^{1-\gamma}}{\rho} \left[ \frac{(\gamma^{-1}-1)^{-\rho} s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma} \right].$$

Thirdly and finally, another direct integration entails

$$\int_{(\gamma^{-1}-1)^{-\gamma} s^{-\gamma}}^{\infty} x^{(\varepsilon-1/2)/\gamma}\, dx = \mathrm{O}\left( s^{-\gamma+1/2-\varepsilon} \right).$$

All in all, and combining these calculations with (E.3), we obtain, uniformly in $s \in [k^{-1+\delta}, 1]$:

$$\begin{aligned}
&\frac{s}{1-2ks/n} \left( \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} - \frac{\overline{Y}_n}{q_{1-k/n}} \right) \\
&= \frac{\gamma}{1-\gamma} \left[ \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} \right]^{1-1/\gamma} + \frac{1}{\sqrt{k}} (\gamma^{-1}-1)^{1/2-\gamma}\gamma \int_0^s W_n(t)\, t^{-\gamma-1}\, dt \\
&+ A(n/k) \frac{(\gamma^{-1}-1)^{1-\gamma} s^{1-\gamma}}{\rho} \left[ \frac{(\gamma^{-1}-1)^{-\rho} s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma} \right] + \mathrm{o}_{\mathbb{P}}\left( \frac{s^{-\gamma+1/2-\varepsilon}}{\sqrt{k}} \right). \qquad \text{(E.8)}
\end{aligned}$$

Recall now the following equivalent characterisation of population expectiles:

$$\xi_\tau - \mathbb{E}(Y) = \frac{2\tau-1}{1-\tau} \mathbb{E}((Y-\xi_\tau)\mathbb{1}_{\{Y>\xi_\tau\}}). \qquad \text{(E.9)}$$

We use this identity with $\tau = 1 - k/n$ to get:

$$\begin{aligned}
& \frac{1}{1-2k/n}\left(\frac{\xi_{1-k/n}}{q_{1-k/n}} - \frac{\mathbb{E}(Y)}{q_{1-k/n}}\right) \\
= \; & \frac{n}{k} \times \frac{1}{q_{1-k/n}} \int_{\xi_{1-k/n}}^{\infty} \overline{F}(u)du \\
= \; & \frac{n}{k}\overline{F}(q_{1-k/n}) \left( \frac{\gamma}{1-\gamma}\left[\frac{\xi_{1-k/n}}{q_{1-k/n}}\right]^{1-1/\gamma} + A(n/k)\frac{(\gamma^{-1}-1)^{1-\gamma}}{\rho}\left[\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma}\right]\right) \\
+ \; & \mathrm{o}(A(n/k))
\end{aligned}$$

thanks to convergence (E.4), the asymptotic equivalence $\overline{F}(q_{1-k/n}) \sim k/n$ following from Lemma 1(ii) and used inside the regularly varying function $A$, Lemma 6 and calculations identical to those we have carried out so far. Using the condition $\sqrt{k}A(n/k) = \mathrm{O}(1)$ and the convergence

$$\lim_{n\to\infty} \frac{1}{A(n/k)}\left(\frac{n}{k}\overline{F}(q_{1-k/n}) - 1\right) = 0$$

which follows from Lemma 1(ii), we obtain

$$\begin{aligned}
& \frac{1}{1-2k/n}\left(\frac{\xi_{1-k/n}}{q_{1-k/n}} - \frac{\mathbb{E}(Y)}{q_{1-k/n}}\right) \\
= \; & \frac{\gamma}{1-\gamma}\left[\frac{\xi_{1-k/n}}{q_{1-k/n}}\right]^{1-1/\gamma} + A(n/k)\frac{(\gamma^{-1}-1)^{1-\gamma}}{\rho}\left[\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma}\right] + \mathrm{o}\left(\frac{1}{\sqrt{k}}\right). \qquad \text{(E.10)}
\end{aligned}$$

Dividing (E.8) by (E.10) and using convergence (E.4) together with a Taylor expansion, we

get

$$
\begin{aligned}
& s\frac{1-2k/n}{1-2ks/n} \times \frac{\widetilde{\xi}_{1-ks/n} - \overline{Y}_n}{\xi_{1-k/n} - \mathbb{E}(Y)} \\
= & \left[ \frac{\gamma}{1-\gamma} \left[ \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} \right]^{1-1/\gamma} + \frac{1}{\sqrt{k}} (\gamma^{-1}-1)^{1/2-\gamma} \gamma \int_0^s W_n(t)\, t^{-\gamma-1}\, dt \right. \\
+ & \left. A(n/k) \frac{(\gamma^{-1}-1)^{1-\gamma} s^{1-\gamma}}{\rho} \left[ \frac{(\gamma^{-1}-1)^{-\rho} s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma} \right] + \mathrm{o}_{\mathbb{P}}\left( \frac{s^{-\gamma+1/2-\varepsilon}}{\sqrt{k}} \right) \right] \\
\times & \frac{1-\gamma}{\gamma} \left[ \frac{q_{1-k/n}}{\xi_{1-k/n}} \right]^{1-1/\gamma} \left( 1 - A(n/k) \frac{\gamma^{-1}-1}{\rho} \left[ \frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma} \right] + \mathrm{o}\left( \frac{1}{\sqrt{k}} \right) \right) \\
= & \left[ \frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} \right]^{1-1/\gamma} + \frac{1}{\sqrt{k}} \gamma \sqrt{\gamma^{-1}-1} \int_0^s W_n(t)\, t^{-\gamma-1}\, dt \\
+ & A(n/k) \frac{(\gamma^{-1}-1) s^{1-\gamma}}{\rho} \left[ \frac{(\gamma^{-1}-1)^{-\rho} s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma} \right] \\
- & A(n/k) \left[ \frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} \right]^{1-1/\gamma} \frac{\gamma^{-1}-1}{\rho} \left[ \frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma} \right] + \mathrm{o}_{\mathbb{P}}\left( \frac{s^{-\gamma+1/2-\varepsilon}}{\sqrt{k}} \right).
\end{aligned}
$$

Define now a random process $s \mapsto r_n(s)$ by the equality

$$
s^{\gamma} \frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} = 1 + r_n(s).
$$

We know, by a combination of convergence (E.4) and Lemma 5, that $r_n(s) \xrightarrow{\mathbb{P}} 0$ uniformly in $s \in [k^{-1+\delta}, 1]$. The above expansion then simplifies as

$$
\begin{aligned}
& s\frac{1-2k/n}{1-2ks/n} \times \frac{\widetilde{\xi}_{1-ks/n} - \overline{Y}_n}{\xi_{1-k/n} - \mathbb{E}(Y)} \\
= & \left[ \frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} \right]^{1-1/\gamma} + \frac{1}{\sqrt{k}} \gamma \sqrt{\gamma^{-1}-1} \int_0^s W_n(t)\, t^{-\gamma-1}\, dt \\
+ & A(n/k) \times \frac{(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho} \times s^{1-\gamma} \frac{s^{-\rho}-1}{\rho} + \mathrm{o}_{\mathbb{P}}\left( \frac{s^{-\gamma+1/2-\varepsilon}}{\sqrt{k}} \right). \qquad \text{(E.11)}
\end{aligned}
$$

We now work on the left-hand side of the above identity. Note that we can write, uniformly in $s \in (0, 1]$:

$$
\frac{1-2k/n}{1-2ks/n} = 1 - \frac{2k}{n} \times \frac{1-s}{1-2ks/n} = 1 - \frac{2k}{n}(1-s) \left[ 1 + \mathrm{O}\left( \frac{k}{n} \right) \right].
$$

Moreover,

$$\begin{aligned}\frac{\widetilde{\xi}_{1-ks/n}-\overline{Y}_n}{\xi_{1-k/n}-\mathbb{E}(Y)} &= \left(\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}}-1\right)\left(1+\frac{\mathbb{E}(Y)}{\xi_{1-k/n}-\mathbb{E}(Y)}\right)+\frac{\xi_{1-k/n}-\overline{Y}_n}{\xi_{1-k/n}-\mathbb{E}(Y)}\\ &= 1+\left(\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}}-1\right)\left(1+\frac{(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y)}{q_{1-k/n}}(1+\mathrm{o}(1))\right)+\mathrm{O}_{\mathbb{P}}\left(\frac{1}{q_{1-k/n}\sqrt{n}}\right)\end{aligned}$$

by asymptotic proportionality of $q_{1-k/n}$ and $\xi_{1-k/n}$, and the central limit theorem. Since $\gamma<1/2$, we have by regular variation of $t\mapsto q_{1-t^{-1}}$ that

$$\frac{1}{q_{1-k/n}\sqrt{n}}\Big/\frac{1}{\sqrt{k}}=\sqrt{\frac{k}{n}}\,q_{1-k/n}=\mathrm{o}(1).$$

Consequently

$$s^{\gamma}\frac{\widetilde{\xi}_{1-ks/n}-\overline{Y}_n}{\xi_{1-k/n}-\mathbb{E}(Y)}=1+r_n(s)\,(1+\mathrm{o}_{\mathbb{P}}(1))+(1-s^{\gamma})\frac{(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y)}{q_{1-k/n}}(1+\mathrm{o}_{\mathbb{P}}(1))+\mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).$$

Notice finally that, by the mean value theorem:

$$1\leqslant\sup_{0\leqslant s<1}\left\{\frac{1-s}{1-s^{\gamma}}\right\}<\infty$$

so that, using the relationship $q_{1-k/n}=\mathrm{o}(n/k)$, we get again that

$$\begin{aligned}&\frac{1-2k/n}{1-2ks/n}\times s^{\gamma}\frac{\widetilde{\xi}_{1-ks/n}-\overline{Y}_n}{\xi_{1-k/n}-\mathbb{E}(Y)}\\ =\;& 1+r_n(s)\,(1+\mathrm{o}_{\mathbb{P}}(1))+(1-s^{\gamma})\frac{(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y)}{q_{1-k/n}}(1+\mathrm{o}_{\mathbb{P}}(1))+\mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).\end{aligned}$$

Because, uniformly in $s\in[k^{-1+\delta},1]$,

$$\left[\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}}\right]^{1-1/\gamma}=s^{1-\gamma}(1+r_n(s))^{1-1/\gamma}=s^{1-\gamma}\left(1+\left[1-\frac{1}{\gamma}\right]r_n(s)(1+\mathrm{o}_{\mathbb{P}}(1))\right)$$

we obtain using (E.11) that:

$$\begin{aligned}&1+r_n(s)\,(1+\mathrm{o}_{\mathbb{P}}(1))+(1-s^{\gamma})\frac{(\gamma^{-1}-1)^{\gamma}}{q_{1-k/n}}(\mathbb{E}(Y)+\mathrm{o}_{\mathbb{P}}(1))+\mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right)\\ =\;& 1+\left[1-\frac{1}{\gamma}\right]r_n(s)(1+\mathrm{o}_{\mathbb{P}}(1))+\frac{1}{\sqrt{k}}\gamma\sqrt{\gamma^{-1}-1}\,s^{\gamma-1}\int_0^s W_n(t)\,t^{-\gamma-1}\,dt\\ &+\;A(n/k)\times\frac{(\gamma^{-1}-1)^{1-\rho}}{1-\gamma-\rho}\times\frac{s^{-\rho}-1}{\rho}+\mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k}}\right).\end{aligned}$$

Rearrange and solve for $r_n(s)$ to get, uniformly in $s \in [k^{-1+\delta}, 1]$:

$$\begin{aligned} r_n(s) &= (s^\gamma - 1)\frac{\gamma(\gamma^{-1}-1)^\gamma}{q_{1-k/n}}(\mathbb{E}(Y) + \mathrm{o}_{\mathbb{P}}(1)) \\ &+ \frac{1}{\sqrt{k}}\gamma^2\sqrt{\gamma^{-1}-1}\, s^{\gamma-1}\int_0^s W_n(t)\, t^{-\gamma-1}\, dt \\ &+ A(n/k) \times \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} \times \frac{s^{-\rho}-1}{\rho} + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k}}\right). \end{aligned}$$

This is precisely what we wanted to show, but in the restricted case $s \in [k^{-1+\delta}, 1]$.

We conclude the proof by focusing on the case $s \in (0, k^{-1+\delta})$. To this end, we choose $\delta \in (0, \varepsilon/(2\varepsilon + 1 + 2\gamma))$ and we note that $\sqrt{k}\, s^{1/2+\varepsilon} \to 0$ uniformly in $s \in (0, k^{-1+\delta})$. It then follows that, by a direct calculation:

$$\sqrt{k} \sup_{0<s<k^{-1+\delta}} s^{1/2+\varepsilon}\left(\frac{1}{q_{1-k/n}} + \frac{1}{\sqrt{k}}s^{\gamma-1}\left|\int_0^s W_n(t)\, t^{-\gamma-1}\, dt\right| + |A(n/k)|\right) \xrightarrow{\mathbb{P}} 0.$$

It is then enough to show that

$$\sqrt{k} \sup_{0<s<k^{-1+\delta}} s^{1/2+\varepsilon}\left|s^\gamma \frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} - 1\right| = \sqrt{k} \sup_{0<s<k^{-1+\delta}} s^{\gamma+1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} - s^{-\gamma}\right| \xrightarrow{\mathbb{P}} 0.$$

Recall that expectiles of an arbitrary distribution are monotonically increasing and exactly cover its support, and apply this to the empirical distribution to get $\widetilde{\xi}_{1-ks/n} \leqslant \widetilde{\xi}_1 = Y_{n,n}$ for any $s \in (0,1)$. Write then

$$\sqrt{k} \sup_{0<s<k^{-1+\delta}} s^{\gamma+1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} - s^{-\gamma}\right| \leqslant k^{1/2+(-1+\delta)(\gamma+1/2+\varepsilon)}\frac{Y_{n,n}}{\xi_{1-k/n}} + \mathrm{o}(1).$$

Using Lemma 3(i) with $s = 1/(2k)$ and $\varepsilon/2$ in place of what was an arbitrary $\eta$ there, gives:

$$\frac{Y_{n,n}}{q_{1-k/n}} = \mathrm{o}_{\mathbb{P}}(k^{\gamma+\varepsilon/2})$$

and therefore, by a use of (E.4) again, we get

$$\sqrt{k} \sup_{0<s<k^{-1+\delta}} s^{\gamma+1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} - s^{-\gamma}\right| = \mathrm{o}_{\mathbb{P}}\left(k^{\delta(\gamma+1/2+\varepsilon)-\varepsilon/2}\right) + \mathrm{o}(1).$$

Recalling that $\delta < \varepsilon/(2\varepsilon + 1 + 2\gamma)$, we obtain

$$\sqrt{k} \sup_{0<s<k^{-1+\delta}} s^{\gamma+1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}} - s^{-\gamma}\right| = \mathrm{o}_{\mathbb{P}}(1).$$

This concludes the proof of the approximation result for the tail expectile process.

To complete the proof, just note that the sequence $W_n$ has the closed form expression

$$W_n(t) = \frac{1}{\sqrt{\gamma^{-1}-1}}\,\widetilde{W}_n\left((\gamma^{-1}-1)t\right),$$

where $\widetilde{W}_n$ denotes the sequence of standard Brownian motions appearing in Lemma 3(ii), see (E.7). This sequence of Brownian motions is also the one appearing in Lemma 3(i), which is nothing but the Gaussian approximation of the tail quantile process. We omit the remaining straightforward technical details. ∎

**Proof of Theorem 3.** The idea is to use (E.8) in the proof of Theorem 2 together with an analogue of (E.10), with $\xi_{1-k/n}$ replaced by $\xi_{1-ks/n}$ and valid uniformly in $s \in (0,1]$. To prove such an analogue relationship, note first that

$$\frac{\xi_{1-ks/n}}{q_{1-k/n}} = \frac{\xi_{1-ks/n}}{q_{1-ks/n}} \times \frac{U(n/ks)}{U(n/k)}.$$

Recall that since $\rho < 0$, the function $t \mapsto U(t)$ is equivalent to a constant multiple of $t \mapsto t^\gamma$ in a neighbourhood of infinity, see p.49 of de Haan and Ferreira (2006). Using (E.4), we obtain

$$\frac{\xi_{1-ks/n}}{q_{1-k/n}} = (\gamma^{-1}-1)^{-\gamma} s^{-\gamma}(1+\mathrm{o}(1)) \tag{E.12}$$

uniformly in $s \in (0,1]$. Use then (E.9) with $\tau = 1-ks/n$ to get

$$\frac{s}{1-2ks/n}\left(\frac{\xi_{1-ks/n}}{q_{1-k/n}} - \frac{\mathbb{E}(Y)}{q_{1-k/n}}\right) = \frac{n}{k} \times \frac{1}{q_{1-k/n}} \int_{\xi_{1-ks/n}}^{\infty} \overline{F}(u)du.$$

Use now the asymptotic equivalence $\overline{F}(q_{1-k/n}) \sim k/n$ following from Lemma 1(ii) and used inside the regularly varying function $A$ together with Lemma 6 to obtain, for any small $\kappa > 0$,

$$\begin{aligned}
&\frac{s}{1-2ks/n}\left(\frac{\xi_{1-ks/n}}{q_{1-k/n}} - \frac{\mathbb{E}(Y)}{q_{1-k/n}}\right) \\
= \;& \frac{n}{k}\overline{F}(q_{1-k/n})\left[\frac{\xi_{1-ks/n}}{q_{1-k/n}}\right]^{1-1/\gamma}\left(\frac{\gamma}{1-\gamma} + A(n/k)\frac{1}{\rho}\left[\frac{1}{1-\gamma-\rho}\left[\frac{\xi_{1-ks/n}}{q_{1-k/n}}\right]^{\rho/\gamma} - \frac{1}{1-\gamma}\right]\right) \\
+ \;& \mathrm{o}\left(A(n/k)\left[\frac{\xi_{1-ks/n}}{q_{1-k/n}}\right]^{1-(1-\rho)/\gamma+\kappa}\right)
\end{aligned}$$

uniformly in $s \in (0,1]$. According to (E.12),

$$\sup_{0<s\leqslant 1}\left[\frac{\xi_{1-ks/n}}{q_{1-k/n}}\right]^{\rho/\gamma+\kappa} \leqslant 2(\gamma^{-1}-1)^{-\rho-\kappa\gamma} \sup_{0<s\leqslant 1} s^{-\rho-\kappa\gamma} = 2(\gamma^{-1}-1)^{-\rho-\kappa\gamma} < \infty$$

for $\kappa$ small enough (recall that $\rho < 0$) and $n$ large enough. Therefore, by (E.12) again:

$$\begin{aligned}
&\frac{s}{1-2ks/n}\left(\frac{\xi_{1-ks/n}}{q_{1-k/n}} - \frac{\mathbb{E}(Y)}{q_{1-k/n}}\right) \\
= \;& \frac{\gamma}{1-\gamma}\left[\frac{\xi_{1-ks/n}}{q_{1-k/n}}\right]^{1-1/\gamma}\left(1 + A(n/k)\frac{\gamma^{-1}-1}{\rho}\left[\frac{(\gamma^{-1}-1)^{-\rho}s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma}\right] + \mathrm{o}\left(\frac{1}{\sqrt{k}}\right)\right)
\end{aligned}$$

uniformly in $s \in (0,1]$. Divide (E.8) by this expansion and use once again (E.12) to get:

$$\begin{aligned}
&\frac{\widetilde{\xi}_{1-ks/n} - \overline{Y}_n}{\xi_{1-ks/n} - \mathbb{E}(Y)} \\
= \;& \left[\frac{\gamma}{1-\gamma}\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{1-1/\gamma} + \frac{1}{\sqrt{k}}(\gamma^{-1}-1)^{1/2-\gamma}\gamma\int_0^s W_n(t)\,t^{-\gamma-1}\,dt\right. \\
+ \;& \left. A(n/k)\frac{(\gamma^{-1}-1)^{1-\gamma}s^{1-\gamma}}{\rho}\left[\frac{(\gamma^{-1}-1)^{-\rho}s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma}\right] + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-\gamma+1/2-\varepsilon}}{\sqrt{k}}\right)\right] \\
\times \;& \frac{1-\gamma}{\gamma}\left[\frac{q_{1-k/n}}{\xi_{1-ks/n}}\right]^{1-1/\gamma}\left(1 - A(n/k)\frac{\gamma^{-1}-1}{\rho}\left[\frac{(\gamma^{-1}-1)^{-\rho}s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma}\right] + \mathrm{o}\left(\frac{1}{\sqrt{k}}\right)\right) \\
= \;& \left[\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}\right]^{1-1/\gamma} + \frac{1}{\sqrt{k}}\gamma\sqrt{\gamma^{-1}-1}\,s^{\gamma-1}\int_0^s W_n(t)\,t^{-\gamma-1}\,dt \\
+ \;& A(n/k)\frac{\gamma^{-1}-1}{\rho}\left[\frac{(\gamma^{-1}-1)^{-\rho}s^{-\rho}}{1-\gamma-\rho} - \frac{1}{1-\gamma}\right]\left(1 - \left[\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}\right]^{1-1/\gamma}\right) + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k}}\right)
\end{aligned}$$

uniformly in $s \in [k^{-1+\delta}, 1]$ and with arbitrarily large probability (here, as in the proof of Theorem 2, $\delta$ is a sufficiently small positive number to be chosen later). Define a random process $s \mapsto R_n(s)$ by the equality

$$\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}} = 1 + R_n(s).$$

We know, by a combination of convergence (E.12) and Lemma 5, that $R_n(s) \xrightarrow{\mathbb{P}} 0$ uniformly in $s \in [k^{-1+\delta}, 1]$. Recalling that $\sqrt{k}A(n/k) = \mathrm{O}(1)$, the above expansion then reads

$$\begin{aligned}
&\frac{\widetilde{\xi}_{1-ks/n} - \overline{Y}_n}{\xi_{1-ks/n} - \mathbb{E}(Y)} \\
= \;& \left[\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}\right]^{1-1/\gamma} + \frac{1}{\sqrt{k}}\gamma\sqrt{\gamma^{-1}-1}\,s^{\gamma-1}\int_0^s W_n(t)\,t^{-\gamma-1}\,dt + \mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k}}\right). \quad \text{(E.13)}
\end{aligned}$$

We now work on the left-hand side of this identity:

$$\begin{aligned}\frac{\widetilde{\xi}_{1-ks/n}-\overline{Y}_n}{\xi_{1-ks/n}-\mathbb{E}(Y)} &= \left(\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}-1\right)\left(1+\frac{\mathbb{E}(Y)}{\xi_{1-ks/n}-\mathbb{E}(Y)}\right)+\frac{\xi_{1-ks/n}-\overline{Y}_n}{\xi_{1-ks/n}-\mathbb{E}(Y)} \\ &= 1+\left(\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-k/n}}-1\right)(1+\mathrm{o}(1))+\mathrm{O}_{\mathbb{P}}\left(\frac{1}{q_{1-k/n}\sqrt{n}}\right)\end{aligned}$$

by asymptotic proportionality of $q_{1-k/n}$ and $\xi_{1-k/n}$ and the central limit theorem. Since moreover $\gamma<1$, we obtain

$$\frac{\widetilde{\xi}_{1-ks/n}-\overline{Y}_n}{\xi_{1-ks/n}-\mathbb{E}(Y)} = 1+R_n(s)\,(1+\mathrm{o}_{\mathbb{P}}(1))+\mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).$$

Because, uniformly in $s\in[k^{-1+\delta},1]$,

$$\left[\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}\right]^{1-1/\gamma} = (1+R_n(s))^{1-1/\gamma} = \left(1+\left[1-\frac{1}{\gamma}\right]R_n(s)(1+\mathrm{o}_{\mathbb{P}}(1))\right)$$

we obtain, using (E.13) and solving for $R_n(s)$, that:

$$R_n(s) = \frac{1}{\sqrt{k}}\gamma^2\sqrt{\gamma^{-1}-1}\,s^{\gamma-1}\int_0^s W_n(t)\,t^{-\gamma-1}\,dt+\mathrm{o}_{\mathbb{P}}\left(\frac{s^{-1/2-\varepsilon}}{\sqrt{k}}\right).$$

This is the desired result in the restricted case $s\in[k^{-1+\delta},1]$.

We conclude the proof by focusing on the case $s\in(0,k^{-1+\delta})$. The idea is very similar to that of the final stages of the proof of Theorem 2. Choose $\delta\in(0,\varepsilon/(2\varepsilon+1+2\gamma))$: it is enough to show that

$$\sqrt{k}\sup_{0<s<k^{-1+\delta}} s^{1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}-1\right| \xrightarrow{\mathbb{P}} 0.$$

Write then

$$\sqrt{k}\sup_{0<s<k^{-1+\delta}} s^{1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}-1\right| \leqslant \sqrt{k}\sup_{0<s<k^{-1+\delta}} s^{1/2+\varepsilon}\left\{\frac{Y_{n,n}}{\xi_{1-ks/n}}\right\}+\mathrm{o}(1).$$

Using (E.12) again, we obtain

$$\sqrt{k}\sup_{0<s<k^{-1+\delta}} s^{1/2+\varepsilon}\left|\frac{\widetilde{\xi}_{1-ks/n}}{\xi_{1-ks/n}}-1\right| = \mathrm{O}\left(\sqrt{k}\sup_{0<s<k^{-1+\delta}} s^{\gamma+1/2+\varepsilon}\left\{\frac{Y_{n,n}}{q_{1-k/n}}\right\}\right)+\mathrm{o}(1).$$

Argue then as in the end of the proof of Theorem 2 to conclude the present proof. ∎

**Proof of Theorem 4.** To show (i), the main idea is to combine Theorem 2 with a Taylor expansion of the logarithm function. This is not quite as straightforward as one might expect, because the error term in the approximation of the tail empirical expectile process given by Theorem 2 does not converge to 0 uniformly in $s$. The trick we use here is to split the integral defining $\breve{\gamma}_k$ in two parts, corresponding to "low" and "high" values of $s$ respectively; we then show directly that the first part is asymptotically negligible, and we analyse the second part using the aforementioned Taylor expansion. A similar argument is used in *e.g.* page 113 of de Haan and Ferreira (2006) and El Methni and Stupfler (2017a, 2017b). Let us finally mention that to use Theorem 2, we should work with a suitable version of the tail expectile process that allows us to write its Gaussian approximation; we can of course do so since this operation leaves the distribution of the estimator $\breve{\gamma}_k$ unchanged. A similar idea will be used, without further mention, in the proofs of Theorems 5 and 8.

Set then $s_n = k^{-(1-\varepsilon)/(1+2\varepsilon)}$ for some $\varepsilon > 0$ sufficiently small (and in particular less than $1/4$), and write

$$\breve{\gamma}_k = \int_0^{s_n} \log\left(\frac{\widetilde{\xi}_{1-ks/n}}{\widetilde{\xi}_{1-k/n}}\right) ds + \int_{s_n}^1 \log\left(\frac{\widetilde{\xi}_{1-ks/n}}{\widetilde{\xi}_{1-k/n}}\right) ds =: I_{n,1} + I_{n,2}. \tag{E.14}$$

We start by controlling directly $I_{n,1}$. This is done by writing

$$|I_{n,1}| \leqslant s_n \log\left(\frac{\widetilde{\xi}_1}{\widetilde{\xi}_{1-k/n}}\right).$$

Recall that $\widetilde{\xi}_1 = Y_{n,n}$ and use a combination of convergence (E.4), Lemma 3(i) and Lemma 4 to find that

$$\log\left(\frac{\widetilde{\xi}_1}{\widetilde{\xi}_{1-k/n}}\right) = \log\left(\frac{Y_{n,n}}{\widehat{q}_{1-k/n}}\right) + \mathrm{O}_{\mathbb{P}}(1).$$

Using further the heavy-tailed assumption on the distribution on $Y$, it follows from Theorem 1.1.6, Theorem 1.2.1 and Lemma 1.2.9 in de Haan and Ferreira (2006) that

$$\frac{Y_{n,n}}{U(n)} \xrightarrow{d} 1 + \gamma G_\gamma$$

where $G_\gamma$ has distribution function $x \mapsto \exp(-(1+\gamma x)^{-1/\gamma})$, for $x > -1/\gamma$. It follows that the limiting variable $1 + \gamma G_\gamma$ is positive and thus $\log(Y_{n,n}/U(n)) = \mathrm{O}_{\mathbb{P}}(1)$ by the continuous mapping theorem. Besides, $\widehat{q}_{1-k/n}/U(n/k) = \widehat{q}_{1-k/n}/q_{1-k/n} \xrightarrow{\mathbb{P}} 1$, by Lemma 3(i) again. Therefore

$$\log\left(\frac{\widetilde{\xi}_1}{\widetilde{\xi}_{1-k/n}}\right) = \log\left(\frac{U(n)}{U(n/k)}\right) + \mathrm{O}_{\mathbb{P}}(1).$$

Potter bounds (see *e.g.* Proposition B.1.9.5 in de Haan and Ferreira, 2006) then yield

$$\log\left(\frac{\widetilde{\xi}_1}{\widetilde{\xi}_{1-k/n}}\right) = \mathrm{O}_{\mathbb{P}}(\log k). \tag{E.15}$$

Recalling that $s_n = k^{-(1-\varepsilon)/(1+2\varepsilon)}$ with $\varepsilon < 1/4$, it is now straightforward to get

$$\sqrt{k}|I_{n,1}| = \mathrm{O}_{\mathbb{P}}\left(s_n \times \sqrt{k}\log k\right) = \mathrm{o}_{\mathbb{P}}(1). \tag{E.16}$$

We now work on $I_{n,2}$. Note that for $s \in [s_n, 1]$, $s^{-1/2-\varepsilon}/\sqrt{k} \leqslant s_n^{-1/2-\varepsilon}/\sqrt{k} = k^{-\varepsilon/2} \to 0$; use then Theorem 2 and a Taylor expansion of the logarithm function to obtain

$$\begin{aligned}
I_{n,2} &= -\gamma\int_{s_n}^1 \log(s)\,ds \\
&+ \frac{1}{\sqrt{k}}\gamma^2\sqrt{\gamma^{-1}-1}\left(\int_{s_n}^1 s^{\gamma-1}\left[\int_0^s W_n(t)\,t^{-\gamma-1}\,dt\right]ds - (1-s_n)\int_0^1 W_n(t)\,t^{-\gamma-1}\,dt\right) \\
&+ \frac{\gamma(\gamma^{-1}-1)^{\gamma}}{q_{1-k/n}}\left(\mathbb{E}(Y)\int_{s_n}^1 (s^{\gamma}-1)\,ds + \mathrm{o}_{\mathbb{P}}(1)\right) \\
&+ \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho}\left(\int_{s_n}^1 \frac{s^{-\rho}-1}{\rho}\,ds\right)A(n/k) + \mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).
\end{aligned}$$

Since $s_n = k^{-(1-\varepsilon)/(1+2\varepsilon)}$, we find that

$$\sqrt{k}\left|\int_0^1 \log(s)\,ds - \int_{s_n}^1 \log(s)\,ds\right| = \mathrm{O}\left(k^{(-1/2+2\varepsilon)/(1+2\varepsilon)}\log k\right) \to 0.$$

Using again the fact that $s_n \to 0$, along with the conditions $1/q_{1-k/n} = \mathrm{O}_{\mathbb{P}}(1/\sqrt{k})$ and $A(n/k) = \mathrm{O}_{\mathbb{P}}(1/\sqrt{k})$, we get

$$\begin{aligned}
I_{n,2} &= \gamma + \frac{1}{\sqrt{k}}\gamma^2\sqrt{\gamma^{-1}-1}\left(\int_0^1 s^{\gamma-1}\left[\int_0^s W_n(t)\,t^{-\gamma-1}\,dt\right]ds - \int_0^1 W_n(t)\,t^{-\gamma-1}\,dt\right) \\
&- \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\frac{1}{q_{1-k/n}} + \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}A(n/k) + \mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).
\end{aligned}$$

By an integration by parts (with the inner integral being differentiated as a function of $s$), we obtain

$$\int_0^1 s^{\gamma-1}\left[\int_0^s W_n(t)\,t^{-\gamma-1}\,dt\right]ds = \frac{1}{\gamma}\int_0^1 \frac{W_n(s)}{s}(s^{-\gamma}-1)ds$$

and therefore, denoting by

$$Z = \gamma\sqrt{\gamma^{-1}-1}\int_0^1 \frac{W(s)}{s}([1-\gamma]s^{-\gamma}-1)ds$$

where $W$ is a standard Brownian motion, we find that

$$\sqrt{k}(I_{n,2}-\gamma) \xrightarrow{d} \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1 - \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2 + Z.$$

It is now enough to compute the variance of $Z$, which is

$$\mathrm{Var}(Z) = \gamma(1-\gamma)\int_0^1\int_0^1 \frac{\min(s,t)}{st}([1-\gamma]s^{-\gamma}-1)([1-\gamma]t^{-\gamma}-1)\,ds\,dt.$$

It then follows from straightforward but lengthy computations that $\mathrm{Var}(Z) = 2\gamma^3/(1-2\gamma)$; we omit the details. Consequently

$$\sqrt{k}(I_{n,2}-\gamma) \xrightarrow{d} \mathcal{N}\left(\frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1 - \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2, \frac{2\gamma^3}{1-2\gamma}\right). \tag{E.17}$$

Combining (E.14), (E.16) and (E.17) completes the proof of (i).

To show (ii), it suffices to prove that

$$|\breve{\gamma}_k - \widetilde{\gamma}_{k,l}| = \mathrm{O}_{\mathbb{P}}\left(\frac{\log(k)}{l}\right). \tag{E.18}$$

Write then

$$|\breve{\gamma}_k - \widetilde{\gamma}_{k,l}| = \left|\sum_{i=1}^{l}\int_{(i-1)/l}^{i/l}\left[\log\left(\frac{\widetilde{\xi}_{1-ks/n}}{\widetilde{\xi}_{1-k/n}}\right) - \log\left(\frac{\widetilde{\xi}_{1-(i-1)k/(ln)}}{\widetilde{\xi}_{1-k/n}}\right)\right]ds\right|$$

and use the sample-wise monotonicity of the random function $s \mapsto \widetilde{\xi}_{1-ks/n}$ to get

$$|\breve{\gamma}_k - \widetilde{\gamma}_{k,l}| \leqslant \frac{1}{l}\sum_{i=1}^{l}\log\left(\frac{\widetilde{\xi}_{1-(i-1)k/(ln)}}{\widetilde{\xi}_{1-ik/(ln)}}\right) = \frac{1}{l}\log\left(\frac{\widetilde{\xi}_1}{\widetilde{\xi}_{1-k/n}}\right).$$

Conclude then using (E.15), which shows (E.18) and completes the proof. ∎

**Proof of Theorem 5.** Since

$$\overline{\gamma}_k(\alpha) = \alpha\widehat{\gamma}_k + (1-\alpha)\widetilde{\gamma}_k$$

it is sufficient to analyse the joint asymptotic behaviour of $(\widehat{\gamma}_k, \widetilde{\gamma}_k, \widehat{q}_{1-k/n}, \widetilde{\xi}_{1-k/n})$. Let us then start by remarking that

$$\widehat{\gamma}_k = \int_0^1 \log\left(\frac{\widehat{q}_{1-\lfloor k\rfloor s/n}}{q_{1-\lfloor k\rfloor/n}}\right)ds - \log\left(\frac{\widehat{q}_{1-\lfloor k\rfloor/n}}{q_{1-\lfloor k\rfloor/n}}\right).$$

Note that, in Theorem 2, the sequence of Brownian motions is left unchanged if $k$ is changed into $\lfloor k\rfloor$ or $\lceil k\rceil$; this is indeed the fundamental argument behind the proof of Lemma 3(i). Arguing as in the proof of Theorem 4 (*i.e.* splitting the domain $s \in (0,1]$ into low and high values of $s$ and using a Taylor expansion), and using the asymptotic equivalences $\sqrt{\lfloor k\rfloor} \sim \sqrt{k}$

and $A(n/\lfloor k \rfloor) \sim A(n/k)$ (the latter due to the regular variation of $|A|$), we get by Theorem 2 that:

$$\sqrt{k}(\widehat{\gamma}_k - \gamma) = \frac{\lambda_1}{1-\rho} + \gamma\sqrt{\gamma^{-1}-1}\left(\int_0^1 \frac{1}{s} W_n\left(\frac{s}{\gamma^{-1}-1}\right) ds - W_n\left(\frac{1}{\gamma^{-1}-1}\right)\right) + \mathrm{o}_{\mathbb{P}}(1). \tag{E.19}$$

Besides, an inspection of the proof of Theorem 4 shows that

$$\begin{aligned}\sqrt{k}(\widetilde{\gamma}_k - \gamma) &= \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1 - \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2 \\ &+ \gamma\sqrt{\gamma^{-1}-1}\int_0^1 \frac{W_n(s)}{s}([1-\gamma]s^{-\gamma}-1)ds + \mathrm{o}_{\mathbb{P}}(1)\end{aligned} \tag{E.20}$$

where $W_n$ is the sequence of Brownian motions appearing in Theorem 2. Using Theorem 2 twice more, we can also write

$$\sqrt{k}\left(\frac{\widehat{q}_{1-k/n}}{q_{1-k/n}} - 1\right) = \gamma\sqrt{\gamma^{-1}-1}\, W_n\left(\frac{1}{\gamma^{-1}-1}\right) + \mathrm{o}_{\mathbb{P}}(1) \tag{E.21}$$

as well as

$$\sqrt{k}\left(\frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1\right) = \gamma^2\sqrt{\gamma^{-1}-1}\int_0^1 W_n(t)t^{-\gamma-1}dt + \mathrm{o}_{\mathbb{P}}(1). \tag{E.22}$$

As a consequence, the random vector

$$\sqrt{k}\left(\widehat{\gamma}_k - \gamma, \widetilde{\gamma}_k - \gamma, \frac{\widehat{q}_{1-k/n}}{q_{1-k/n}} - 1, \frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1\right)$$

is asymptotically four-variate Gaussian, and as such

$$\sqrt{k}\left(\overline{\gamma}_k(\alpha) - \gamma, \frac{\widehat{q}_{1-k/n}}{q_{1-k/n}} - 1, \frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1\right)$$

is asymptotically trivariate Gaussian. To complete the proof, we analyse the marginal asymptotic behaviour of each of the three components in this vector, as well as their pairwise asymptotic covariance structure.

*Marginal asymptotic behaviour of $\overline{\gamma}_k(\alpha)$:* This is determined by the joint convergence of $\sqrt{k}\,(\widehat{\gamma}_k - \gamma, \widetilde{\gamma}_k - \gamma)$, to what we already know to be a bivariate Gaussian distribution. We also know from Theorem 3.2.5 in de Haan and Ferreira (2006) that

$$\sqrt{k}(\widehat{\gamma}_k - \gamma) \xrightarrow{d} \mathcal{N}\left(\frac{\lambda_1}{1-\rho}, \gamma^2\right).$$

This is of course also a corollary of (E.19). Meanwhile, Theorem 4(ii) gives

$$\sqrt{k}(\widetilde{\gamma}_k - \gamma) \xrightarrow{d} \mathcal{N}\left(\frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1 - \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2, \frac{2\gamma^3}{1-2\gamma}\right).$$

It therefore only remains to calculate the limiting covariance of $\sqrt{k}(\widehat{\gamma}_k - \gamma, \widetilde{\gamma}_k - \gamma)$. This is obtained by computing the expectation of the product of the centred Gaussian terms appearing in the two asymptotic expansions (E.19) and (E.20). In other words, the limiting covariance is

$$\begin{aligned} \text{cov} &= \text{cov}_1 - \text{cov}_2 \\ \text{with } \text{cov}_1 &:= \gamma(1-\gamma)\int_0^1\int_0^1 \frac{\min(s,(\gamma^{-1}-1)^{-1}t)}{st}([1-\gamma]s^{-\gamma}-1)ds\,dt \\ \text{and } \text{cov}_2 &:= \gamma(1-\gamma)\int_0^1 \frac{\min(s,(\gamma^{-1}-1)^{-1})}{s}([1-\gamma]s^{-\gamma}-1)ds. \end{aligned}$$

Recalling that $\gamma^{-1}-1>1$, straightforward computations entail

$$\begin{aligned} \text{cov}_1 &= (\gamma^{-1}-1)^{\gamma-1} - \gamma[\gamma+1+\gamma\log(\gamma^{-1}-1)] \\ \text{and } \text{cov}_2 &= \gamma[(\gamma^{-1}-1)^{\gamma} - 1 - \gamma\log(\gamma^{-1}-1)]. \end{aligned} \tag{E.23}$$

This results in

$$\text{cov} = \gamma[(\gamma^{-1}-1)^{\gamma-1}-\gamma] = \gamma^2\left(\frac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma}-1\right).$$

Wrapping up, we obtain

$$\sqrt{k}\,(\widehat{\gamma}_k - \gamma, \widetilde{\gamma}_k - \gamma) \xrightarrow{d} \mathcal{N}(\boldsymbol{m}, \boldsymbol{V}) \tag{E.24}$$

where $\boldsymbol{m}$ is the $2\times 1$ vector

$$\boldsymbol{m} := \left(\frac{\lambda_1}{1-\rho}, \frac{(1-\gamma)(\gamma^{-1}-1)^{-\rho}}{(1-\rho)(1-\gamma-\rho)}\lambda_1 - \mathbb{E}(Y)\frac{\gamma^2(\gamma^{-1}-1)^{\gamma}}{\gamma+1}\lambda_2\right)$$

and $\boldsymbol{V}$ is the $2\times 2$ matrix

$$\boldsymbol{V} := \begin{pmatrix} \gamma^2 & \gamma^2\left(\dfrac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma}-1\right) \\ \gamma^2\left(\dfrac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma}-1\right) & \dfrac{2\gamma^3}{1-2\gamma} \end{pmatrix}.$$

After some more straightforward computations, we conclude that

$$\sqrt{k}\,(\overline{\gamma}_k(\alpha) - \gamma) \xrightarrow{d} \mathcal{N}\left(b_\alpha, \mathfrak{V}_\alpha(1,1)\right)$$

with the notation of the statement of Theorem 5.

*Marginal asymptotic behaviour of $\widehat{q}_{1-k/n}$:* It is a straightforward byproduct of Equation (E.21) that

$$\sqrt{k}\left(\frac{\widehat{q}_{1-k/n}}{q_{1-k/n}}-1\right) \xrightarrow{d} \mathcal{N}(0,\gamma^2).$$

*Marginal asymptotic behaviour of* $\widetilde{\xi}_{1-k/n}$*:* It is a direct consequence of Equation (E.22) that

$$\sqrt{k}\left(\frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}}-1\right)\xrightarrow{d}\mathcal{N}\left(0,\frac{2\gamma^3}{1-2\gamma}\right).$$

See also the discussion below Theorem 2 in the main paper.

*Asymptotic covariance structure of* $(\overline{\gamma}_k(\alpha),\widehat{q}_{1-k/n})$*:* For this, we remark first that $\widehat{\gamma}_k-\gamma$ and $\widehat{q}_{1-k/n}/q_{1-k/n}-1$ are asymptotically independent: this is a consequence of the asymptotic representation of $\widehat{\gamma}_k-\gamma$ obtained in the proof of Theorem 3.2.5 in de Haan and Ferreira (2006) together with Lemma 3.2.3 therein. Besides, the limiting covariance structure of $\sqrt{k}(\widetilde{\gamma}_k-\gamma,\widehat{q}_{1-k/n}/q_{1-k/n}-1)$ is obtained by computing the expectation of the product of the centred Gaussian terms appearing in the asymptotic expansions (E.20) and (E.21). By (E.23) above, this limiting covariance is:

$$\mathrm{cov}_2=\gamma[(\gamma^{-1}-1)^{\gamma}-1-\gamma\log(\gamma^{-1}-1)]$$

with the notation of (E.23). The limiting covariance of $\sqrt{k}(\overline{\gamma}_k(\alpha)-\gamma,\widehat{q}_{1-k/n}/q_{1-k/n}-1)$ is then

$$(1-\alpha)\gamma[(\gamma^{-1}-1)^{\gamma}-1-\gamma\log(\gamma^{-1}-1)]=\mathfrak{V}_{\alpha}(1,2).$$

*Asymptotic covariance structure of* $(\overline{\gamma}_k(\alpha),\widetilde{\xi}_{1-k/n})$*:* It follows from Equations (E.19), (E.20) and (E.22) that the limiting covariance of $\sqrt{k}(\overline{\gamma}_k(\alpha)-\gamma,\widetilde{\xi}_{1-k/n}/\xi_{1-k/n}-1)$ is

$$\mathcal{COV}=\alpha\mathcal{COV}_1+(1-\alpha)\mathcal{COV}_2$$

with

$$\mathcal{COV}_1=\gamma^2(1-\gamma)\left[\int_0^1\int_0^1\frac{\min((\gamma^{-1}-1)^{-1}s,t)}{s}t^{-\gamma-1}ds\,dt-\int_0^1\min((\gamma^{-1}-1)^{-1},t)t^{-\gamma-1}dt\right]$$

and

$$\mathcal{COV}_2=\gamma^2(1-\gamma)\int_0^1\int_0^1\frac{\min(s,t)}{s}([1-\gamma]s^{-\gamma}-1)t^{-\gamma-1}ds\,dt.$$

Direct computations yield

$$\mathcal{COV}_1=\gamma^2\left(\frac{(\gamma^{-1}-1)^{\gamma}}{(1-\gamma)^2}-1\right)-\gamma^2\left(\frac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma}-1\right)=\frac{\gamma^3(\gamma^{-1}-1)^{\gamma}}{(1-\gamma)^2}$$

and

$$\mathcal{COV}_2=\frac{\gamma^3}{(1-\gamma)(1-2\gamma)}.$$

Consequently

$$\mathcal{COV}=\frac{\gamma^3}{(1-\gamma)^2}\left[\alpha(\gamma^{-1}-1)^{\gamma}+(1-\alpha)\frac{1-\gamma}{1-2\gamma}\right]=\mathfrak{V}_{\alpha}(1,3).$$

*Asymptotic covariance structure of* $(\widehat{q}_{1-k/n}, \widetilde{\xi}_{1-k/n})$*:* Combining Equations (E.21) and (E.22), we find that the limiting covariance of $\sqrt{k}(\widehat{q}_{1-k/n}/q_{1-k/n} - 1, \widetilde{\xi}_{1-k/n}/\xi_{1-k/n} - 1)$ is

$$\gamma^2(1-\gamma)\int_0^1 \min(t, (\gamma^{-1}-1)^{-1})t^{-\gamma-1}dt = \gamma^2\left(\frac{(\gamma^{-1}-1)^{\gamma}}{1-\gamma} - 1\right) = \mathfrak{V}_\alpha(2,3)$$

after some straightforward calculations.

Combining these arguments on marginal convergence and asymptotic covariance structure, we get

$$\sqrt{k}\left(\overline{\gamma}_k(\alpha) - \gamma, \frac{\widehat{q}_{1-k/n}}{q_{1-k/n}} - 1, \frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1\right) \xrightarrow{d} \mathcal{N}(\mathfrak{m}_\alpha, \mathfrak{V}_\alpha) \tag{E.25}$$

with $\mathfrak{m}_\alpha$ and $\mathfrak{V}_\alpha$ as in the statement of Theorem 5. This concludes the proof. ∎

**Proof of Theorem 6.** Applying Theorem 5 and arguing as in the proof of Theorem 1 in Daouia *et al.* (2018), we get the joint convergence

$$\sqrt{k}\left(\frac{\widehat{\xi}_{1-k/n}(\alpha)}{\xi_{1-k/n}} - 1, \frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} - 1\right) \xrightarrow{d} \left([(1-\gamma)^{-1} - \log(\gamma^{-1}-1)]\Gamma_\alpha + \Theta - \lambda,\, \Xi\right)$$

where $(\Gamma_\alpha, \Theta, \Xi)$ is the limiting vector in Theorem 5, and

$$\lambda := \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho} - 1}{\rho}\right)\lambda_1 + \gamma(\gamma^{-1}-1)^{\gamma}\mathbb{E}(Y)\lambda_2.$$

Then clearly

$$\sqrt{k}\left(\frac{\overline{\xi}_{1-k/n}(\alpha,\beta)}{\xi_{1-k/n}} - 1\right) \xrightarrow{d} [(1-\gamma)^{-1} - \log(\gamma^{-1}-1)]\beta\Gamma_\alpha + \beta\Theta + (1-\beta)\Xi - \beta\lambda.$$

Set $\Psi_\alpha = \Gamma_\alpha - b_\alpha$ and rearrange the bias component to complete the proof. ∎

**Proof of Theorem 7.** Define $p_n = 1 - \tau_n'$ and we note that

$$\begin{aligned}\log\left(\frac{\overline{\xi}^{\star}_{1-p_n}(\alpha,\beta)}{\xi_{1-p_n}}\right) &= (\overline{\gamma}_{1-k/n}(\alpha) - \gamma)\log\left(\frac{k}{np_n}\right) + \log\left(\frac{\overline{\xi}_{1-k/n}(\alpha,\beta)}{\xi_{1-k/n}}\right) \\ &\quad - \log\left(\left[\frac{np_n}{k}\right]^{\gamma}\frac{\xi_{1-p_n}}{\xi_{1-k/n}}\right).\end{aligned}$$

The convergence $\log[k/(np_n)] \to \infty$ yields

$$\frac{\sqrt{k}}{\log[k/(np_n)]} \log\left(\frac{\overline{\xi}_{1-k/n}(\alpha,\beta)}{\xi_{1-k/n}}\right) = \mathrm{O}_{\mathbb{P}}\left(1/\log[k/(np_n)]\right) = \mathrm{o}_{\mathbb{P}}(1) \tag{E.26}$$

$$\begin{aligned}
\text{and} \quad & \frac{\sqrt{k}}{\log[k/(np_n)]} \log\left(\left[\frac{np_n}{k}\right]^{\gamma} \frac{\xi_{1-p_n}}{\xi_{1-k/n}}\right) \\
&= \frac{\sqrt{k}}{\log[k/(np_n)]} \left(\log\left(\frac{\xi_{1-p_n}}{q_{1-p_n}}\right) - \log\left(\frac{\xi_{1-k/n}}{q_{1-k/n}}\right) + \log\left(\left[\frac{np_n}{k}\right]^{\gamma} \frac{q_{1-p_n}}{q_{1-k/n}}\right)\right) \\
&= \mathrm{O}\left(\frac{\sqrt{k}}{\log[k/(np_n)]} \left[\frac{1}{q_{1-k/n}} + |A(n/k)| + \frac{1}{q_{1-p_n}} + |A(1/p_n)|\right]\right) \\
&= \mathrm{O}\left(\frac{\sqrt{k}}{\log[k/(np_n)]} \left[\frac{1}{q_{1-k/n}} + |A(n/k)|\right]\right) \\
&= \mathrm{o}(1).
\end{aligned} \tag{E.27}$$

Here, convergence (E.26) is a consequence of Theorem 6. Convergence (E.27) follows from a combination of Proposition 1, Theorem 2.3.9 in de Haan and Ferreira (2006) and the regular variation of $|A|$. Combining these convergences and using the delta-method leads to the desired result. ∎

**Proof of Proposition 2.** Statement (i) is a clear consequence of the fact that the expectile-based ES at level $\tau$ is an increasing linear functional of the restriction of the expectile function on the interval $[\tau, 1]$, in the sense that

$$\xi_t^{(1)} \leqslant \xi_t^{(2)} \ \forall t \in [\tau,1] \ \Rightarrow \ \mathrm{XES}_\tau^{(1)} := \frac{1}{1-\tau}\int_\tau^1 \xi_t^{(1)} dt \leqslant \frac{1}{1-\tau}\int_\tau^1 \xi_t^{(2)} dt =: \mathrm{XES}_\tau^{(2)}.$$

To show statement (ii), note that, for $\tau \geqslant 1/2$, $\mathrm{XTCE}_\tau$ is clearly translation invariant and positive homogeneous (because so are expectiles above level $\tau \geqslant 1/2$, and conditional expectations). A simple counter-example to monotonicity and subadditivity is the following: set $\tau = 1/2$, so that

$$\mathrm{XTCE}_{1/2}(Z) = \mathbb{E}(Z \,|\, Z > \xi_{1/2}(Z)) = \mathbb{E}(Z \,|\, Z > \mathbb{E}(Z)).$$

We then actually show that $\mathrm{XTCE}_{1/2}$ is neither monotonic nor subadditive. For this, we consider a uniform random variable $U$ on $[0,1]$ and we set

$$X = 2\,\mathbb{1}_{\{5/6 \leqslant U < 1\}} \ \text{ and } \ Y = \mathbb{1}_{\{1/2 \leqslant U < 5/6\}} + 2\,\mathbb{1}_{\{5/6 \leqslant U < 1\}}.$$

Then clearly $X \leqslant Y$ with probability 1, and $X$ and $Y$ are discrete variables taking values in the set $\{0,1,2\}$, with $\mathbb{E}(X) = \mathbb{E}(X\mathbb{1}_{\{X>0\}}) = 1/3$ and $\mathbb{E}(Y) = \mathbb{E}(Y\mathbb{1}_{\{Y>0\}}) = 2/3$. As such

$$\mathbb{E}(X \,|\, X > \mathbb{E}(X)) = \mathbb{E}(X \,|\, X > 0) = 2$$

$$\text{and} \qquad \mathbb{E}(Y \,|\, Y > \mathbb{E}(Y)) = \mathbb{E}(Y \,|\, Y > 0) = \frac{2/3}{1/2} = \frac{4}{3}.$$

This establishes that $\mathbb{E}(Y \,|\, Y > \mathbb{E}(Y)) < \mathbb{E}(X \,|\, X > \mathbb{E}(X))$: $\mathrm{XTCE}_{1/2}$ is not a monotonic risk measure. Besides,

$$X + Y = \mathbb{1}_{\{1/2 \leqslant U < 5/6\}} + 4\,\mathbb{1}_{\{5/6 \leqslant U < 1\}}$$

so that $\mathbb{E}(X+Y) = \mathbb{E}([X+Y]\mathbb{1}_{\{X+Y>0\}}) = 1$ and then

$$\mathbb{E}(X+Y \,|\, X+Y > \mathbb{E}(X+Y)) = \mathbb{E}(X+Y \,|\, X+Y > 1) = 4.$$

This shows that $\mathbb{E}(X+Y \,|\, X+Y > \mathbb{E}(X+Y)) > \mathbb{E}(X \,|\, X > \mathbb{E}(X)) + \mathbb{E}(Y \,|\, Y > \mathbb{E}(Y))$, proving that $\mathrm{XTCE}_{1/2}$ is not a subadditive risk measure either. ■

**Proof of Proposition 3.** It follows from the asymptotic proportionality relationship $\xi_\tau/q_\tau \sim (\gamma^{-1}-1)^{-\gamma}$ as $\tau \to 1$ (see Bellini and Di Bernardino, 2017) that

$$\mathrm{XES}_\tau = \frac{1}{1-\tau}\int_\tau^1 \xi_\alpha d\alpha = (\gamma^{-1}-1)^{-\gamma}\left\{\frac{1}{1-\tau}\int_\tau^1 q_\alpha(1+r(\alpha))d\alpha\right\}$$

where $r(\alpha) \to 0$ as $\alpha \to 1$. It is then clear that

$$\mathrm{XES}_\tau \sim (\gamma^{-1}-1)^{-\gamma}\left\{\frac{1}{1-\tau}\int_\tau^1 q_\alpha d\alpha\right\} = (\gamma^{-1}-1)^{-\gamma}\mathrm{QES}_\tau \quad \text{as} \quad \tau \to 1.$$

This proves that

$$\frac{\mathrm{XES}_\tau}{\mathrm{QES}_\tau} \sim (\gamma^{-1}-1)^{-\gamma} \sim \frac{\xi_\tau}{q_\tau} \quad \text{as} \quad \tau \to 1,$$

by asymptotic proportionality again. Besides, the equality $q_\alpha = U((1-\alpha)^{-1})$ and a change of variables entail

$$\frac{\mathrm{QES}_\tau}{q_\tau} = \frac{1}{1-\tau}\int_\tau^1 \frac{q_\alpha}{q_\tau}d\alpha = \int_1^\infty y^{-1}\frac{U((1-\tau)^{-1}y)}{U((1-\tau)^{-1})}\frac{dy}{y}.$$

The condition $\gamma < 1$ and a uniform convergence theorem such as Proposition B.1.10 in de Haan and Ferreira (2006, p.360) entail

$$\frac{\mathrm{QES}_\tau}{q_\tau} \to \int_1^\infty y^{\gamma-2}dy = \frac{1}{1-\gamma} \quad \text{as} \quad \tau \to 1.$$

Consequently

$$\frac{\mathrm{XES}_\tau}{\xi_\tau} \sim \frac{\mathrm{QES}_\tau}{q_\tau} \to \frac{1}{1-\gamma} \quad \text{as} \quad \tau \to 1.$$

Let us now turn to the terms $\mathrm{XTCE}_\tau/\mathrm{QTCE}_\tau$ and $\mathrm{XTCE}_\tau/\xi_\tau$. On the one hand, we have

$$\mathrm{XTCE}_\tau = \frac{\mathbb{E}\left[Y\mathbb{1}(Y>\xi_\tau)\right]}{\overline{F}(\xi_\tau)} = \frac{\mathbb{E}\left[(Y-\xi_\tau)_+\right]}{\overline{F}(\xi_\tau)} + \xi_\tau,$$

where $y_+ = \max(y,0)$. On the other hand, it follows from the proof of Theorem 11 in Bellini *et al.* (2014) that

$$\frac{\mathbb{E}\left[(Y-\xi_\tau)_+\right]}{\overline{F}(\xi_\tau)} \sim \frac{\xi_\tau}{\gamma^{-1}-1} \quad \text{as} \quad \tau \to 1.$$

Therefore

$$\frac{\mathrm{XTCE}_\tau}{\xi_\tau} \sim \frac{1}{1-\gamma} \quad \text{as} \quad \tau \to 1.$$

Likewise, we have

$$\mathrm{QTCE}_\tau = \frac{\mathbb{E}\left[Y\mathbb{1}(Y>q_\tau)\right]}{\overline{F}(q_\tau)} = \frac{\mathbb{E}\left[(Y-q_\tau)_+\right]}{\overline{F}(q_\tau)} + q_\tau,$$

with

$$\frac{\mathbb{E}\left[(Y-q_\tau)_+\right]}{\overline{F}(q_\tau)} \sim \frac{q_\tau}{\gamma^{-1}-1} \quad \text{as} \quad \tau \to 1.$$

Then

$$\frac{\mathrm{QTCE}_\tau}{q_\tau} \sim \frac{1}{1-\gamma} \quad \text{as} \quad \tau \to 1.$$

Whence

$$\frac{\mathrm{XTCE}_\tau}{\mathrm{QTCE}_\tau} \sim \frac{\xi_\tau}{q_\tau} \quad \text{as} \quad \tau \to 1,$$

which completes the proof. ∎

**Proof of Proposition 4.** The starting point to show the first expansion is Proposition 1(i), which yields

$$\begin{aligned}
\mathrm{XES}_\tau &= \frac{1}{1-\tau}\int_\tau^1 \xi_\alpha d\alpha \\
&= (\gamma^{-1}-1)^{-\gamma}\Bigg(\mathrm{QES}_\tau + \gamma(\gamma^{-1}-1)^\gamma \mathbb{E}(Y)(1+\mathrm{o}(1)) \\
&+ \left\{\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho} + \frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho} + \mathrm{o}(1)\right\}\frac{1}{1-\tau}\int_\tau^1 q_\alpha A((1-\alpha)^{-1})d\alpha\Bigg).
\end{aligned}$$

Use a change of variables to get

$$\frac{1}{1-\tau}\int_\tau^1 q_\alpha A((1-\alpha)^{-1})d\alpha = U((1-\tau)^{-1})A((1-\tau)^{-1})\int_1^\infty y^{-1}\frac{U((1-\tau)^{-1}y)A((1-\tau)^{-1}y)}{U((1-\tau)^{-1})A((1-\tau)^{-1})}\frac{dy}{y}.$$

This entails, using a uniform convergence theorem such as Proposition B.1.10 in de Haan and Ferreira (2006, p.360), that

$$\begin{aligned}\frac{1}{1-\tau}\int_\tau^1 q_\alpha A((1-\alpha)^{-1})d\alpha &\sim U((1-\tau)^{-1})A((1-\tau)^{-1})\int_1^\infty y^{\gamma+\rho-2}dy \text{ as } \tau\to 1\\ &= \frac{q_\tau A((1-\tau)^{-1})}{1-\gamma-\rho}.\end{aligned}$$

Since $\mathrm{QES}_\tau \sim q_\tau/(1-\gamma)$, our earlier expansion yields

$$\begin{aligned}\frac{\mathrm{XES}_\tau}{\mathrm{QES}_\tau} &= (\gamma^{-1}-1)^{-\gamma}\Bigg(1+\frac{\gamma(1-\gamma)(\gamma^{-1}-1)^\gamma\mathbb{E}(Y)}{q_\tau}(1+\mathrm{o}(1))\\ &+ \left\{\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho}+\frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho}+\mathrm{o}(1)\right\}\frac{1-\gamma}{1-\gamma-\rho}A((1-\tau)^{-1})\Bigg). \quad \text{(E.28)}\end{aligned}$$

Furthermore, it is a consequence of a uniform inequality such as Theorem 2.3.9 in de Haan and Ferreira (2006) applied to the function $U$ that

$$\begin{aligned}\frac{\mathrm{QES}_\tau}{q_\tau} &= \int_1^\infty y^{-1}\frac{U((1-\tau)^{-1}y)}{U((1-\tau)^{-1})}\frac{dy}{y}\\ &= \int_1^\infty y^{-1}\left(y^\gamma + A((1-\tau)^{-1})y^\gamma\frac{y^\rho-1}{\rho}(1+\mathrm{o}(1))\right)\frac{dy}{y}\\ &= \int_1^\infty y^{\gamma-2}dy + \frac{A((1-\tau)^{-1})}{\rho}\int_1^\infty\left(y^{\gamma+\rho-2}-y^{\gamma-2}\right)dy(1+\mathrm{o}(1))\\ &= \frac{1}{1-\gamma}\left(1+\frac{1}{1-\gamma-\rho}A((1-\tau)^{-1})(1+\mathrm{o}(1))\right). \quad \text{(E.29)}\end{aligned}$$

Finally, Proposition 1(i) reads

$$\begin{aligned}\frac{q_\tau}{\xi_\tau} &= (\gamma^{-1}-1)^\gamma\Bigg(1-\frac{\gamma(\gamma^{-1}-1)^\gamma\mathbb{E}(Y)}{q_\tau}(1+\mathrm{o}(1))\\ &- \left(\frac{(\gamma^{-1}-1)^{-\rho}}{1-\gamma-\rho}+\frac{(\gamma^{-1}-1)^{-\rho}-1}{\rho}+\mathrm{o}(1)\right)A((1-\tau)^{-1})\Bigg). \quad \text{(E.30)}\end{aligned}$$

A use of the identity

$$\frac{\mathrm{XES}_\tau}{\xi_\tau} = \frac{\mathrm{XES}_\tau}{\mathrm{QES}_\tau}\times\frac{\mathrm{QES}_\tau}{q_\tau}\times\frac{q_\tau}{\xi_\tau}$$

and a combination of (E.28), (E.29) and (E.30) complete the proof after some straightforward computations. ■

**Proof of Theorem 8.** By Theorem 3:

$$\begin{aligned}\frac{\widetilde{\mathrm{XES}}_{1-k/n}}{\mathrm{XES}_{1-k/n}} - 1 &= \frac{1}{\sqrt{k}}\gamma^2\sqrt{\gamma^{-1}-1} \times \frac{\int_0^1 (\int_0^s W_n(t)t^{-\gamma-1}\,dt)s^{\gamma-1}\xi_{1-ks/n}\,ds}{\int_0^1 \xi_{1-ks/n}\,ds} \\ &+ \mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}} \times \frac{\int_0^1 s^{-1/2-\varepsilon}\xi_{1-ks/n}\,ds}{\int_0^1 \xi_{1-ks/n}\,ds}\right).\end{aligned}$$

Using (E.12) and the fact that $\gamma < 1/2$, we obtain:

$$\frac{\widetilde{\mathrm{XES}}_{1-k/n}}{\mathrm{XES}_{1-k/n}} - 1 = \frac{1}{\sqrt{k}}(\gamma[1-\gamma])^{3/2}\int_0^1 \left(\int_0^s W_n(t)t^{-\gamma-1}\,dt\right)\frac{ds}{s} + \mathrm{o}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\right).$$

Denoting by $W$ a standard Brownian motion, we get, using an integration by parts, that:

$$\sqrt{k}\left(\frac{\widetilde{\mathrm{XES}}_{1-k/n}}{\mathrm{XES}_{1-k/n}} - 1\right) \xrightarrow{d} (\gamma[1-\gamma])^{3/2}\int_0^1 W(s)s^{-\gamma-1}\log(s)\,ds.$$

Since the rhs above is a centred Gaussian random variable, it only remains to compute its variance, which is

$$v = \gamma^3(1-\gamma)^3\int_0^1\int_0^1 \min(s,t)s^{-\gamma-1}t^{-\gamma-1}\log(s)\log(t)\,ds\,dt.$$

It then follows from straightforward but lengthy computations that

$$v = \frac{2\gamma^3(1-\gamma)(3-4\gamma)}{(1-2\gamma)^3}$$

as required. ∎

**Proof of Theorem 9.** The proof of this result is entirely similar to that of Theorem 7 (applying Theorem 8 instead of Theorem 6, and Proposition 4 instead of Proposition 1). We omit the details. ∎

**Proof of Theorem 10.** We examine first the convergence of $\overline{\mathrm{XES}}^{\star}_{1-p_n}(\alpha,\beta)$. Define $p_n = 1-\tau_n'$ and write

$$\begin{aligned}\log\left(\frac{\overline{\mathrm{XES}}^{\star}_{1-p_n}(\alpha,\beta)}{\mathrm{XES}_{1-p_n}}\right) &= \log\left(\frac{\overline{\xi}^{\star}_{1-p_n}(\alpha,\beta)}{\xi_{1-p_n}}\right) + \log\left(\frac{[1-\overline{\gamma}_{1-k/n}(\alpha)]^{-1}}{[1-\gamma]^{-1}}\right) \\ &\quad - \log\left(\frac{\mathrm{XES}_{1-p_n}}{[1-\gamma]^{-1}\xi_{1-p_n}}\right).\end{aligned}$$

By Theorem 7 and the delta-method,

$$\frac{\sqrt{k}}{\log[k/(np_n)]}\log\left(\frac{\overline{\xi}^{\star}_{1-p_n}(\alpha,\beta)}{\xi_{1-p_n}}\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha). \tag{E.31}$$

Using then Theorem 5, the delta-method and the convergence $\log[k/(np_n)] \to \infty$, we get

$$\frac{\sqrt{k}}{\log[k/(np_n)]} \log\left(\frac{[1-\overline{\gamma}_{1-k/n}(\alpha)]^{-1}}{[1-\gamma]^{-1}}\right) \xrightarrow{\mathbb{P}} 0. \tag{E.32}$$

Using finally a combination of Proposition 1(i), Proposition 4 and the regular variation of $|A|$ and $t \mapsto q_{1-t^{-1}}$, we obtain

$$\frac{\sqrt{k}}{\log[k/(np_n)]} \log\left(\frac{\mathrm{XES}_{1-p_n}}{[1-\gamma]^{-1}\xi_{1-p_n}}\right) \to 0. \tag{E.33}$$

Combining convergences (E.31), (E.32) and (E.33), it follows that

$$\frac{\sqrt{k}}{\log[k/(np_n)]} \log\left(\frac{\overline{\mathrm{XES}}^{\star}_{1-p_n}(\alpha,\beta)}{\mathrm{XES}_{1-p_n}}\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha).$$

Another use of the delta-method completes the proof of the convergence of $\overline{\mathrm{XES}}^{\star}_{1-p_n}(\alpha,\beta)$.

We now show the convergence of $\widehat{\mathrm{XES}}^{\star}_{1-p_n}(\alpha,\beta)$. For this we write

$$\begin{aligned}\log\left(\frac{\widehat{\mathrm{XES}}^{\star}_{1-p_n}(\alpha,\beta)}{\mathrm{XES}_{1-p_n}}\right) &= \log\left(\frac{\overline{\xi}^{\star}_{1-p_n}(\alpha,\beta)}{\xi_{1-p_n}}\right) + \log\left(\frac{\widehat{\mathrm{QES}}_{1-k/n}}{\widehat{q}_{1-k/n}} \cdot \frac{q_{1-k/n}}{\mathrm{QES}_{1-k/n}}\right) \\ &\quad + \log\left(\frac{\mathrm{QES}_{1-k/n}}{q_{1-k/n}}\right) - \log\left(\frac{\mathrm{XES}_{1-p_n}}{\xi_{1-p_n}}\right)\end{aligned}$$

where we set

$$\widehat{\mathrm{QES}}_{1-k/n} := \frac{1}{\lfloor k \rfloor} \sum_{i=1}^{\lfloor k \rfloor} Y_{n-i+1,n} = \int_0^1 \widehat{q}_{1-\lfloor k \rfloor s/n}\, ds.$$

Remark now that, since $\widehat{q}_{1-\lfloor k \rfloor/n} = Y_{n-\lfloor k \rfloor,n} = \widehat{q}_{1-k/n}$, we have

$$\log\left(\frac{\widehat{\mathrm{QES}}_{1-k/n}}{\widehat{q}_{1-k/n}} \cdot \frac{q_{1-k/n}}{\mathrm{QES}_{1-k/n}}\right) = \log\left(\int_0^1 \frac{\widehat{q}_{1-\lfloor k \rfloor s/n}}{\widehat{q}_{1-\lfloor k \rfloor/n}} ds\right) - \log\left(\frac{\mathrm{QES}_{1-k/n}}{q_{1-k/n}}\right).$$

Combine then Theorem 2, the delta-method, and (E.29) together with a Taylor expansion to obtain

$$\frac{\sqrt{k}}{\log[k/(np_n)]} \log\left(\frac{\widehat{\mathrm{QES}}_{1-k/n}}{\widehat{q}_{1-k/n}} \cdot \frac{q_{1-k/n}}{\mathrm{QES}_{1-k/n}}\right) = \mathrm{O}_{\mathbb{P}}(1/\log[k/(np_n)]) = \mathrm{o}_{\mathbb{P}}(1). \tag{E.34}$$

Besides, a combination of Equation (E.29) and Proposition 4 with a Taylor expansion yields

$$\begin{aligned}
&\frac{\sqrt{k}}{\log[k/(np_n)]}\left[\log\left(\frac{\mathrm{QES}_{1-k/n}}{q_{1-k/n}}\right)-\log\left(\frac{\mathrm{XES}_{1-p_n}}{\xi_{1-p_n}}\right)\right]\\
&=\mathrm{O}\left(\frac{\sqrt{k}}{\log[k/(np_n)]}\left[\frac{1}{q_{1-k/n}}+|A(n/k)|+\frac{1}{q_{1-p_n}}+|A(1/p_n)|\right]\right)\\
&=\mathrm{O}\left(\frac{\sqrt{k}}{\log[k/(np_n)]}\left[\frac{1}{q_{1-k/n}}+|A(n/k)|\right]\right)\\
&=\mathrm{o}(1). \qquad \text{(E.35)}
\end{aligned}$$

Finally, use together (E.31), (E.34) and (E.35) and the delta-method to complete the proof. ■

**Proof of Theorem 11.** We only show the result for $\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)$ as the proofs of the other convergences are entirely similar. The key point is to write

$$\widetilde{\mathrm{XES}}^{\star}_{\widehat{\tau}'_n(p_n)}(\alpha)=\left(\frac{1-\widehat{\tau}'_n(p_n)}{1-\tau'_n(p_n)}\right)^{-\overline{\gamma}_{\tau_n}(\alpha)}\widetilde{\mathrm{XES}}^{\star}_{\tau'_n(p_n)}(\alpha). \qquad \text{(E.36)}$$

It is, moreover, shown as part of the proof of Theorem 6 in Daouia *et al.* (2018) that

$$\frac{1-\widehat{\tau}'_n(p_n)}{1-\tau'_n(p_n)}=1+\mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{n(1-\tau_n)}}\right)$$

(combine (B.52), (B.53), (B.54) and (B.55) in the supplementary material document of [Daouia *et al.*, 2018], noting that the strict monotonicity of $F_Y$ is not required thanks to Proposition 1(i) in the present paper; this also results in a corrected version of (B.51) in the former paper). Therefore

$$\begin{aligned}
\left(\frac{1-\widehat{\tau}'_n(p_n)}{1-\tau'_n(p_n)}\right)^{-\overline{\gamma}_{\tau_n}(\alpha)} &= \exp\left(-\overline{\gamma}_{\tau_n}(\alpha)\log\left[\frac{1-\widehat{\tau}'_n(p_n)}{1-\tau'_n(p_n)}\right]\right)\\
&= \exp\left(-\left[\gamma+\mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{n(1-\tau_n)}}\right)\right]\times\mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{n(1-\tau_n)}}\right)\right)\\
&= 1+\mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{n(1-\tau_n)}}\right). \qquad \text{(E.37)}
\end{aligned}$$

Furthermore, using Proposition 5, we conclude that the conditions of Theorem 9 are satisfied if the parameter $\tau'_n$ there is set equal to $\tau'_n(p_n)$. By Theorem 9 then:

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-\tau'_n(p_n))]}\left(\frac{\widetilde{\mathrm{XES}}^{\star}_{\tau'_n(p_n)}(\alpha)}{\mathrm{XES}_{\tau'_n(p_n)}}-1\right)\xrightarrow{d}\mathcal{N}(b_\alpha,v_\alpha).$$

Now

$$\log\left[\frac{1-\tau_n}{1-\tau_n'(p_n)}\right] = \log\left[\frac{1-\tau_n}{1-p_n}\right] + \log\left[\frac{1-p_n}{1-\tau_n'(p_n)}\right]$$

and in the right-hand side of this identity, the first term tends to infinity, while the second term converges to a finite constant in view of Proposition 5. As a conclusion

$$\log\left[\frac{1-\tau_n}{1-\tau_n'(p_n)}\right] \sim \log\left[\frac{1-\tau_n}{1-p_n}\right].$$

Hence the convergence

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-p_n)]}\left(\frac{\widetilde{\mathrm{XES}}^{\star}_{\tau_n'(p_n)}(\alpha)}{\mathrm{XES}_{\tau_n'(p_n)}} - 1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha). \tag{E.38}$$

We conclude the proof by writing

$$\mathrm{XES}_{\tau_n'(p_n)} = \mathrm{QES}_{p_n} \times \left\{(1-\gamma)\frac{\mathrm{XES}_{\tau_n'(p_n)}}{\xi_{\tau_n'(p_n)}}\right\} \times \left\{(1-\gamma)\frac{\mathrm{QES}_{p_n}}{q_{p_n}}\right\}^{-1}$$

(since $\xi_{\tau_n'(p_n)} \equiv q_{p_n}$ by definition). By a combination of Propositions 4 and 5 with the regular variation of the functions $|A|$ and $t \mapsto q_{1-t^{-1}}$, one gets

$$(1-\gamma)\frac{\mathrm{XES}_{\tau_n'(p_n)}}{\xi_{\tau_n'(p_n)}} = 1 + \mathrm{o}\left(\frac{\log[(1-\tau_n)/(1-p_n)]}{\sqrt{n(1-\tau_n)}}\right).$$

Similarly and by (E.29),

$$(1-\gamma)\frac{\mathrm{QES}_{p_n}}{q_{p_n}} = 1 + \mathrm{o}\left(\frac{\log[(1-\tau_n)/(1-p_n)]}{\sqrt{n(1-\tau_n)}}\right).$$

Therefore

$$\frac{\mathrm{XES}_{\tau_n'(p_n)}}{\mathrm{QES}_{p_n}} - 1 = \mathrm{o}\left(\frac{\log[(1-\tau_n)/(1-p_n)]}{\sqrt{n(1-\tau_n)}}\right).$$

Together with (E.38), this entails

$$\frac{\sqrt{n(1-\tau_n)}}{\log[(1-\tau_n)/(1-p_n)]}\left(\frac{\widetilde{\mathrm{XES}}^{\star}_{\tau_n'(p_n)}(\alpha)}{\mathrm{QES}_{p_n}} - 1\right) \xrightarrow{d} \mathcal{N}(b_\alpha, v_\alpha). \tag{E.39}$$

Combining (E.36), (E.37) and (E.39) completes the proof. ∎

# Appendix: Preliminary results and their proofs

The first preliminary lemma, which we will use to show Proposition 1, is a technical result on second-order regular variation that seems to be informally known in the literature. We prove it for the sake of completeness.

**Lemma 1.** *Assume that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds. Then we have the following two convergences:*

*(i)* $\lim_{t\to\infty} \dfrac{1}{A(1/\overline{F}(t))}\left(\dfrac{U(1/\overline{F}(t))}{t} - 1\right) = 0;$

*(ii)* $\lim_{t\to\infty} \dfrac{1}{A(t)}\left(\dfrac{1/\overline{F}(U(t))}{t} - 1\right) = 0.$

**Proof of Lemma 1.** The proof of this lemma is based on that of Theorem B.3.19 in de Haan and Ferreira (2006). We only show (i), the proof of (ii) being entirely similar. Recall that

$$U(t) = \inf\{x \,|\, 1/\overline{F}(x) \geqslant t\}$$

so that $U(1/\overline{F}(t)) \leqslant t$. Furthermore, condition $\mathcal{C}_2(\gamma, \rho, A)$ is nothing but second-order extended regular variation in the sense of convergence (B.3.3) in de Haan and Ferreira (2006), which is known to be locally uniform in $x \in (0, \infty)$ (see Remark B.3.8.1 in de Haan and Ferreira, 2006). Pick $\varepsilon \in \mathbb{R}$ arbitrarily close to 0: by using condition $\mathcal{C}_2(\gamma, \rho, A)$ with $t$ replaced by $1/\overline{F}(t)$ and $x = 1 + \varepsilon A(1/\overline{F}(t))$, $t \to \infty$, we get

$$\lim_{t\to\infty} \frac{1}{A(1/\overline{F}(t))}\left[\frac{U([1 + \varepsilon A(1/\overline{F}(t))]/\overline{F}(t))}{U(1/\overline{F}(t))} - (1 + \varepsilon A(1/\overline{F}(t)))^{\gamma}\right] = 0$$

or equivalently

$$\lim_{t\to\infty} \frac{1}{A(1/\overline{F}(t))}\left[\frac{U([1 + \varepsilon A(1/\overline{F}(t))]/\overline{F}(t))}{U(1/\overline{F}(t))} - 1\right] = \gamma\varepsilon.$$

Assume that $A$ is positive and take $\varepsilon > 0$; the proof in the other case is similar by taking $\varepsilon < 0$ instead. Using the definition of $U$ again, we find that $U([1 + \varepsilon A(1/\overline{F}(t))]/\overline{F}(t)) \geqslant t$, and thus

$$0 \leqslant \liminf_{t\to\infty} \frac{1}{A(1/\overline{F}(t))}\left(\frac{t}{U(1/\overline{F}(t))} - 1\right) \leqslant \limsup_{t\to\infty} \frac{1}{A(1/\overline{F}(t))}\left(\frac{t}{U(1/\overline{F}(t))} - 1\right) \leqslant \gamma\varepsilon.$$

Let $\varepsilon \downarrow 0$ to complete the proof. ∎

The second lemma is an equivalent characterisation of sample expectiles, and is the crucial initial step for the proof of our Gaussian approximations.

**Lemma 2.** *We have, for any* $\tau \in (0,1)$,

$$\widetilde{\xi}_\tau - \frac{1}{n}\sum_{i=1}^n Y_i = \frac{2\tau-1}{1-\tau} \times \frac{1}{n}\sum_{i=1}^n (Y_i - \widetilde{\xi}_\tau)\mathbb{1}_{\{Y_i > \widetilde{\xi}_\tau\}}.$$

**Proof of Lemma 2.** The proof is straightforward; we include it for the sake of completeness. Since $\eta_\tau$ is strictly convex and continuously differentiable with derivative $\eta'_\tau(y) = 2|\tau - \mathbb{1}_{\{y\leqslant 0\}}|y$, the sample expectile $\widetilde{\xi}_\tau$ is characterised by

$$\sum_{i=1}^n |\tau - \mathbb{1}_{\{Y_i \leqslant \widetilde{\xi}_\tau\}}|(Y_i - \widetilde{\xi}_\tau) = 0.$$

In other words,

$$(1-\tau)\sum_{i=1}^n (\widetilde{\xi}_\tau - Y_i)\mathbb{1}_{\{Y_i \leqslant \widetilde{\xi}_\tau\}} = \tau \sum_{i=1}^n (Y_i - \widetilde{\xi}_\tau)\mathbb{1}_{\{Y_i > \widetilde{\xi}_\tau\}}$$

and thus

$$(1-\tau)\sum_{i=1}^n (\widetilde{\xi}_\tau - Y_i) = (2\tau-1)\sum_{i=1}^n (Y_i - \widetilde{\xi}_\tau)\mathbb{1}_{\{Y_i > \widetilde{\xi}_\tau\}}.$$

Dividing on each side by $n(1-\tau)$ and rearranging yields the result. ∎

The third lemma is a generalisation of the weighted approximation of the tail empirical quantile process to non-integer sequences $k$ tailored to our purpose. It also gives a representation of the Gaussian term that is of independent interest, for example when evaluating the correlation between two quantiles or expectiles at different orders.

**Lemma 3.** *Suppose that condition* $\mathcal{C}_2(\gamma, \rho, A)$ *holds. Let* $k = k(n) \to \infty$ *be a positive sequence such that* $k/n \to 0$ *and* $\sqrt{k}A(n/k) = \mathrm{O}(1)$. *Then, subject to a potential enlargement of the underlying probability space and to choosing a suitable version of the empirical process* $\widehat{F}_n$, *there exists a sequence* $W_n = W_n^{(k)}$ *of standard Brownian motions such that, for any* $\varepsilon > 0$ *sufficiently small:*

*(i) We have*

$$\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} = s^{-\gamma} + \frac{1}{\sqrt{k}}\left(\gamma s^{-\gamma-1}W_n(s) + \sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho} + s^{-\gamma-1/2-\varepsilon}\,\mathrm{o}_{\mathbb{P}}(1)\right)$$

*uniformly in* $s \in (0,1]$.

*(ii) If* $\widehat{\overline{F}}_n(u) = n^{-1}\sum_{i=1}^n \mathbb{1}_{\{Y_i>u\}}$ *is the empirical survival function of the* $Y_i$*, we have*

$$\frac{n}{k}\widehat{\overline{F}}_n\left(xq_{1-k/n}\right) - x^{-1/\gamma} = \frac{1}{\sqrt{k}}\left(W_n(x^{-1/\gamma}) + \sqrt{k}A(n/k)x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho} + x^{(\varepsilon-1/2)/\gamma}\,\mathrm{o}_{\mathbb{P}}(1)\right)$$

*uniformly on half-lines of the form* $[x_0,\infty)$*, for* $x_0 > 0$*.*

*Moreover, the sequence* $W_n$ *can be chosen as* $W_n(s) = W_n^{(k)}(s) = \sqrt{n/k}\,\overline{W}_n(ks/n)$*, where* $\overline{W}_n$ *is a sequence of Brownian motions which is fixed across all possible choices of* $k$*.*

**Proof of Lemma 3.** Note that (i) is exactly Theorem 2.4.8 in de Haan and Ferreira (2006), recalling that the function $A_0$ therein is asymptotically equivalent to $A$, in the case when $k$ is a sequence of integers. If now the sequence $k$ is not a sequence of integers, we write, for $n$ so large that $q_{1-k/n} > 0$,

$$\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \leqslant \frac{\widehat{q}_{1-\lfloor k\rfloor s/n}}{q_{1-\lfloor k\rfloor/n}}\left(\frac{q_{1-\lfloor k\rfloor/n}}{q_{1-k/n}} - 1\right) + \frac{\widehat{q}_{1-\lfloor k\rfloor s/n}}{q_{1-\lfloor k\rfloor/n}}.$$

By local uniformity of condition $\mathcal{C}_2(\gamma,\rho,A)$ (see *e.g.* Theorem 2.3.9 in de Haan and Ferreira, 2006), one gets

$$\begin{aligned}\sqrt{k}\left(\frac{q_{1-\lfloor k\rfloor/n}}{q_{1-k/n}} - 1\right) &= \sqrt{k}\left(\left[\frac{k}{\lfloor k\rfloor}\right]^\gamma - 1\right) + \sqrt{k}A(n/k)\times\left\{\frac{1}{A(n/k)}\left(\frac{U(n/\lfloor k\rfloor)}{U(n/k)} - \left[\frac{k}{\lfloor k\rfloor}\right]^\gamma\right)\right\}\\ &= \mathrm{o}(1).\end{aligned}$$

Therefore

$$\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \leqslant \frac{\widehat{q}_{1-\lfloor k\rfloor s/n}}{q_{1-\lfloor k\rfloor/n}}\left[1 + \mathrm{o}\left(\frac{1}{\sqrt{k}}\right)\right].$$

Bearing in mind that $s \in (0,1]$ and therefore $s^{-\gamma} \leqslant s^{-\gamma-1/2-\varepsilon}$, we may now use the approximation for the sequence of integers $\lfloor k\rfloor$ to get

$$\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} \leqslant s^{-\gamma} + \frac{1}{\sqrt{\lfloor k\rfloor}}\left(\gamma s^{-\gamma-1}W_n(s) + \sqrt{\lfloor k\rfloor}A(n/\lfloor k\rfloor)s^{-\gamma}\frac{s^{-\rho}-1}{\rho} + s^{-\gamma-1/2-\varepsilon}\,\mathrm{o}_{\mathbb{P}}(1)\right)$$

where the $\mathrm{o}_{\mathbb{P}}(1)$ is uniform in $s \in (0,1]$. An analogue lower bound, in terms of the ceiling function $\lceil\cdot\rceil$, applies. We can now use the facts that

$$\sqrt{k}\left(\frac{\sqrt{\lfloor k\rfloor}}{\sqrt{k}} - 1\right) \to 0 \quad\text{and}\quad \frac{A(n/\lfloor k\rfloor)}{A(n/k)} \to 1$$

(the latter being due to the regular variation of $A$) and similarly when the floor function is replaced by the ceiling function, to get statement (i). To prove statement (ii), we note that if

the sequence $k$ is made of integers, there is a sequence $\widetilde{W}_n$ of (potentially different) Brownian motions such that, for a suitable version of the empirical process $\widehat{F}_n$:

$$\frac{n}{k}\widehat{\overline{F}}_n\left(xq_{1-k/n}\right) - x^{-1/\gamma} = \frac{1}{\sqrt{k}}\left(\widetilde{W}_n(x^{-1/\gamma}) + \sqrt{k}A(n/k)x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho} + x^{(\varepsilon-1/2)/\gamma}\,\mathrm{o}_{\mathbb{P}}(1)\right).$$

This follows from Theorem 5.1.4 in de Haan and Ferreira (2006). That $\widetilde{W}_n$ can be taken equal to $W_n$ is hinted at in Remark 5.1.3 in de Haan and Ferreira (2006), and can be shown as follows: work throughout with the above version of $\widehat{F}_n$, and denote by $s \mapsto \widehat{q}_{1-ks/n}$ the related tail quantile process. Our goal is to show that for any $\eta, \delta > 0$, we have, for $n$ large enough,

$$\mathbb{P}\left(\sup_{0<s\leqslant 1} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} - s^{-\gamma}\right) - \gamma s^{-\gamma-1}\widetilde{W}_n(s) - \sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}\right| > \eta\right) < \delta.$$

First note that, for any $a \in (0,1]$,

$$\begin{aligned}
&\sup_{0<s\leqslant a} s^{\gamma+1/2+\varepsilon}\left|\gamma s^{-\gamma-1}\widetilde{W}_n(s) + \sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}\right| \\
\leqslant\ & \gamma \sup_{0<s\leqslant a} s^{-1/2+\varepsilon}\left|\widetilde{W}_n(s)\right| + \left|\sqrt{k}A(n/k)\right| \sup_{0<s\leqslant a} s^{1/2+\varepsilon}\left|\frac{s^{-\rho}-1}{\rho}\right| \\
\overset{d}{=}\ & a^{\varepsilon} \times \gamma \sup_{0<s\leqslant 1} s^{-1/2+\varepsilon}\left|\widetilde{W}_n(s)\right| + \sup_{0<s\leqslant a} s^{1/2+\varepsilon}\left|\frac{s^{-\rho}-1}{\rho}\right| \times \mathrm{O}_{\mathbb{P}}(1)
\end{aligned}$$

by self-similarity of the Brownian motion $\widetilde{W}_n$. One can therefore choose $a > 0$ so small that

$$\mathbb{P}\left(\sup_{0<s\leqslant a} s^{\gamma+1/2+\varepsilon}\left|\gamma s^{-\gamma-1}\widetilde{W}_n(s) + \sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}\right| > \frac{\eta}{4}\right) < \frac{\delta}{4}. \tag{E.40}$$

Using statement (i) together with the triangle inequality, and repeating exactly the same argument, we obtain that we can choose $a > 0$ so small that

$$\mathbb{P}\left(\sup_{0<s\leqslant a} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} - s^{-\gamma}\right)\right| > \frac{\eta}{4}\right) < \frac{\delta}{4}. \tag{E.41}$$

Combining (E.40) and (E.41) results, for this choice of $a > 0$, in the inequality

$$\mathbb{P}\left(\sup_{0<s\leqslant a} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} - s^{-\gamma}\right) - \gamma s^{-\gamma-1}\widetilde{W}_n(s) - \sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}\right| > \frac{\eta}{2}\right) < \frac{\delta}{2}.$$

It is therefore sufficient, for our purpose, to show that for any $a > 0$

$$\sup_{a<s\leqslant 1} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} - s^{-\gamma}\right) - \gamma s^{-\gamma-1}\widetilde{W}_n(s) - \sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}\right| = \mathrm{o}_{\mathbb{P}}(1),$$

and since $s^{\gamma+1/2+\varepsilon} \geqslant a^{\gamma+1/2+\varepsilon} > 0$ on $[a,1]$, it is actually enough to show that

$$\sup_{a\leqslant s\leqslant 1}\left|\sqrt{k}\left(\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}}-s^{-\gamma}\right)-\gamma s^{-\gamma-1}\widetilde{W}_n(s)-\sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}\right| = \mathrm{o}_{\mathbb{P}}(1). \tag{E.42}$$

By statement (i), we have

$$\sqrt{k}\sup_{a\leqslant s\leqslant 1}\left|\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}}-s^{-\gamma}\right| = \mathrm{O}_{\mathbb{P}}(1).$$

Set then $x = x_n(s) = \widehat{q}_{1-ks/n}/q_{1-k/n}$ in the approximation of $\widehat{\overline{F}}_n\left(xq_{1-k/n}\right)$ to get, uniformly in $s \in [a,1]$,

$$\begin{aligned}&\frac{\lfloor ks\rfloor}{k}-[x_n(s)]^{-1/\gamma}\\ =\;&\frac{1}{\sqrt{k}}\left(\widetilde{W}_n([x_n(s)]^{-1/\gamma})+\sqrt{k}A(n/k)[x_n(s)]^{-1/\gamma}\frac{[x_n(s)]^{\rho/\gamma}-1}{\gamma\rho}+[x_n(s)]^{(\varepsilon-1/2)/\gamma}\,\mathrm{o}_{\mathbb{P}}(1)\right).\end{aligned}$$

By the uniform convergence of $x_n(s)$ to $s^{-\gamma}$ on $[a,1]$, as well as the continuity properties of Brownian motion, this entails

$$[x_n(s)]^{-1/\gamma} = s-\frac{1}{\sqrt{k}}\left(\widetilde{W}_n(s)+\sqrt{k}A(n/k)s\frac{s^{-\rho}-1}{\gamma\rho}+\mathrm{o}_{\mathbb{P}}(1)\right)$$

uniformly in $s \in [a,1]$. By a Taylor expansion, we find

$$\frac{\widehat{q}_{1-ks/n}}{q_{1-k/n}} = x_n(s) = s^{-\gamma}+\frac{1}{\sqrt{k}}\left(\gamma s^{-\gamma-1}\widetilde{W}_n(s)+\sqrt{k}A(n/k)s^{-\gamma}\frac{s^{-\rho}-1}{\rho}+\mathrm{o}_{\mathbb{P}}(1)\right)$$

uniformly in $s \in [a,1]$. This is exactly (E.42). The adaptation of (ii) to an arbitrary sequence $k$ (not necessarily of integers) then follows by a direct adaptation of the arguments used to show (i).

That $W_n$ can be chosen as indicated in the final statement is most easily seen by inspecting the proof of Theorem 2.4.8 in de Haan and Ferreira (2006): the centrepiece of the proof is Lemma 2.4.10 therein, which states that there is an independent sequence $(Z_i)_{i\geqslant 1}$ of unit Pareto random variables such that, if $s \mapsto Z_{n-\lfloor ks\rfloor,n}$ is the related tail quantile process, then one can construct a sequence of Brownian motions $W_n$ such that, for any $\gamma > 0$ and any $\varepsilon > 0$ sufficiently small,

$$\sup_{k^{-1}\leqslant s\leqslant 1} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{(\frac{k}{n}Z_{n-\lfloor ks\rfloor,n})^{\gamma}-1}{\gamma}-\frac{s^{-\gamma}-1}{\gamma}\right)-s^{-\gamma-1}W_n(s)\right| = \mathrm{o}_{\mathbb{P}}(1). \tag{E.43}$$

The sequence $W_n$ in Theorem 2.4.8 of de Haan and Ferreira (2006), and therefore in the statement of the Lemma, is exactly the sequence $W_n$ satisfying this relationship for the

sequence $(Y_i = U(Z_i))_{i\geqslant 1}$; this is seen by combining (2.4.23), (2.4.24) and (2.4.25) p.59 of de Haan and Ferreira (2006). Equation (E.43), meanwhile, is shown in the following way: Proposition 2.4.9 in de Haan and Ferreira (or equivalently, Theorem 6.2.1 in Csörgő and Horváth, 1993) yields that, for a suitable choice of an independent sequence $(Z_i)_{i\geqslant 1}$ of unit Pareto random variables, there is a sequence of Brownian bridges $B_n$ such that

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} n^{\varepsilon} t^{\varepsilon-1/2}(1-t)^{\varepsilon-1/2}\left|\sqrt{n}(1-t)^{\gamma+1}\left(\frac{Z^{\gamma}_{\lceil nt\rceil,n}-1}{\gamma}-\frac{(1-t)^{-\gamma}-1}{\gamma}\right)-B_n(t)\right|$$

is stochastically bounded. Setting $t = 1 - ks/n$ and rearranging yields in particular that

$$\sup_{k^{-1}\leqslant s\leqslant 1} k^{\varepsilon}\left|s^{\gamma+1/2+\varepsilon}\sqrt{k}\left(\frac{(\frac{k}{n}Z_{n-\lfloor ks\rfloor,n})^{\gamma}-1}{\gamma}-\frac{s^{-\gamma}-1}{\gamma}\right)-\sqrt{\frac{n}{k}}\,s^{\varepsilon-1/2}B_n(1-ks/n)\right|$$

is stochastically bounded. Since $k^{\varepsilon}\to\infty$, it follows that

$$\sup_{k^{-1}\leqslant s\leqslant 1} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{(\frac{k}{n}Z_{n-\lfloor ks\rfloor,n})^{\gamma}-1}{\gamma}-\frac{s^{-\gamma}-1}{\gamma}\right)-\sqrt{\frac{n}{k}}\,s^{-\gamma-1}B_n(1-ks/n)\right| = \mathrm{o}_{\mathbb{P}}(1).$$

Set now $\overline{B}_n(t) = B_n(1-t)$, which makes $\overline{B}_n$ a sequence of Brownian bridges as well, and let $\overline{W}_n$ be any sequence of Brownian motions such that

$$\overline{B}_n(t) = \overline{W}_n(t) - t\overline{W}_n(1)$$

(for instance, $\overline{W}_n(t) = \overline{B}_n(t) + tV_n$, where for each $n$, $V_n$ is a standard Gaussian random variable independent of the process $\overline{B}_n$). Note that the sequence $\overline{W}_n$ is constructed independently of $k$. We have

$$\sqrt{\frac{n}{k}}\,s^{-\gamma-1}B_n(1-ks/n) = \sqrt{\frac{n}{k}}\,s^{-\gamma-1}\overline{W}_n(ks/n) - \sqrt{\frac{n}{k}}\,s^{-\gamma-1}\times\frac{ks}{n}\overline{W}_n(1)$$

and clearly

$$\sup_{k^{-1}\leqslant s\leqslant 1} s^{\gamma+1/2+\varepsilon}\left|\sqrt{\frac{n}{k}}\,s^{-\gamma-1}\times\frac{ks}{n}\overline{W}_n(1)\right| = \left|\sqrt{\frac{k}{n}}\overline{W}_n(1)\right| = \mathrm{O}_{\mathbb{P}}\left(\sqrt{\frac{k}{n}}\right) = \mathrm{o}_{\mathbb{P}}(1).$$

It follows that

$$\sup_{k^{-1}\leqslant s\leqslant 1} s^{\gamma+1/2+\varepsilon}\left|\sqrt{k}\left(\frac{(\frac{k}{n}Z_{n-\lfloor ks\rfloor,n})^{\gamma}-1}{\gamma}-\frac{s^{-\gamma}-1}{\gamma}\right)-s^{-\gamma-1}\sqrt{n/k}\,\overline{W}_n(ks/n)\right| = \mathrm{o}_{\mathbb{P}}(1).$$

A choice of $W_n$ giving Equation (E.43) is therefore $W_n(s) = W_n^{(k)}(s) = \sqrt{n/k}\,\overline{W}_n(ks/n)$, as claimed in the statement of the result. ∎

The fourth lemma is a preliminary consistency result for intermediate sample expectiles, under a weaker moment condition than that of Theorem 1.

**Lemma 4.** *Let $k = k(n) \to \infty$ be a positive sequence such that $k/n \to 0$. Suppose further that the distribution of $Y$ is heavy-tailed with tail index $\gamma \in (0, 1/2)$, and assume that $\mathbb{E}|Y_-|^2 < \infty$. Then*

$$\frac{\widetilde{\xi}_{1-k/n}}{\xi_{1-k/n}} \xrightarrow{\mathbb{P}} 1 \quad as \quad n \to \infty.$$

**Proof of Lemma 4.** The idea of the proof follows closely that of Theorem 2 in Daouia *et al.* (2018), which was an asymptotic normality result formulated using the parametrisation $\tau_n = 1 - k/n$, where $\tau_n \to 1$ is such that $n(1-\tau_n) \to \infty$. To make it easier for the reader to relate the present proof with the one of Daouia *et al.* (2018), we adopt this parametrisation here. We shall therefore show that $\widetilde{\xi}_{\tau_n}/\xi_{\tau_n} \xrightarrow{\mathbb{P}} 1$, and we will actually prove the stronger statement

$$v_n \left( \frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) \xrightarrow{\mathbb{P}} 0 \quad \text{provided} \quad v_n \to \infty \quad \text{and} \quad v_n = \mathrm{o}\left( \sqrt{n(1-\tau_n)} \right).$$

Note that

$$v_n \left( \frac{\widetilde{\xi}_{\tau_n}}{\xi_{\tau_n}} - 1 \right) = \arg\min_{u \in \mathbb{R}} \psi_n(u) \tag{E.44}$$

$$\text{with} \qquad \psi_n(u) := \frac{v_n^2}{n(1-\tau_n)} \sum_{i=1}^{n} \frac{1}{2\xi_{\tau_n}^2} \left[ \eta_{\tau_n}\left( Y_i - \xi_{\tau_n} - \frac{u\xi_{\tau_n}}{v_n} \right) - \eta_{\tau_n}(Y_i - \xi_{\tau_n}) \right].$$

Denoting the derivative of $y \mapsto \eta_\tau(y)/2$ by $\varphi_\tau(y) := |\tau - \mathbb{1}_{\{y \leqslant 0\}}| y$, it is straightforward to get (*e.g.* using Lemma 2 in Daouia *et al.*, 2018):

$$\begin{aligned} \psi_n(u) &= -uT_{1,n} + T_{2,n}(u) \qquad (E.45) \\ \text{with} \quad T_{1,n} &:= \frac{v_n}{n(1-\tau_n)} \sum_{i=1}^{n} \frac{1}{\xi_{\tau_n}} \varphi_{\tau_n}(Y_i - \xi_{\tau_n}) =: \sum_{i=1}^{n} S_{n,i} \\ \text{and} \quad T_{2,n}(u) &:= -\frac{v_n^2}{n(1-\tau_n)\xi_{\tau_n}^2} \sum_{i=1}^{n} \int_0^{u\xi_{\tau_n}/v_n} (\varphi_{\tau_n}(Y_i - \xi_{\tau_n} - t) - \varphi_{\tau_n}(Y_i - \xi_{\tau_n}))dt. \end{aligned}$$

The random variables $S_{n,i}$ are independent, identically distributed, and centred since

$$\xi_{\tau_n} = \arg\min_{u \in \mathbb{R}} \mathbb{E}(\eta_{\tau_n}(Y_i - u) - \eta_{\tau_n}(Y_i)) \Rightarrow \mathbb{E}(\varphi_{\tau_n}(Y_i - \xi_{\tau_n})) = 0,$$

by differentiating under the expectation sign. Now note that

$$\mathbb{E}(|\varphi_\tau(Y - \xi_\tau)|^2) = \mathrm{O}\left(\xi_\tau^2(1-\tau)\right) \quad \text{as} \quad \tau \uparrow 1,$$

by Lemma 4 in Daouia *et al.* (2018). Therefore

$$\mathrm{Var}(T_{1,n}) = \mathrm{O}\left(\frac{v_n^2}{n(1-\tau_n)}\right) \to 0.$$

Because $\mathbb{E}(T_{1,n}) = 0$, Chebyshev's inequality then yields

$$T_{1,n} \xrightarrow{\mathbb{P}} 0. \tag{E.46}$$

Now

$$T_{2,n}(u) = T_{3,n}(u) - \frac{v_n^2}{(1-\tau_n)\xi_{\tau_n}^2}\int_0^{u\xi_{\tau_n}/v_n} [\mathbb{E}(\varphi_{\tau_n}(Y-\xi_{\tau_n}-t)) - \mathbb{E}(\varphi_{\tau_n}(Y-\xi_{\tau_n}))]dt. \tag{E.47}$$

Note that the integral on the right-hand side of (E.47) is indeed well-defined because of Fubini's theorem and the fact that $Y$ has a finite absolute first moment. The random term $T_{3,n}(u)$, meanwhile, is a sum of independent, identically distributed and centred random variables, which we shall examine after having controlled this nonrandom integral. Note then that, by Lemma 3 in Daouia *et al.* (2018),

$$\begin{aligned} & \mathbb{E}(\varphi_{\tau_n}(Y-\xi_{\tau_n}-t)) - \mathbb{E}(\varphi_{\tau_n}(Y-\xi_{\tau_n})) \\ = \;& (1-2\tau_n)\mathbb{E}((Y-\xi_{\tau_n}-t)(\mathbb{1}_{\{Y\leqslant \xi_{\tau_n}+t\}} - \mathbb{1}_{\{Y\leqslant\xi_{\tau_n}\}})) - t\mathbb{E}(|\tau_n - \mathbb{1}_{\{Y\leqslant\xi_{\tau_n}\}}|). \end{aligned} \tag{E.48}$$

Clearly

$$\mathbb{E}(|\tau_n - \mathbb{1}_{\{Y\leqslant\xi_{\tau_n}\}}|) = \tau_n\overline{F}(\xi_{\tau_n}) + (1-\tau_n)F(\xi_{\tau_n}).$$

It therefore follows from (E.4) that

$$\mathbb{E}(|\tau_n - \mathbb{1}_{\{Y\leqslant\xi_{\tau_n}\}}|) = \gamma^{-1}(1-\tau_n)(1+\mathrm{o}(1)) \tag{E.49}$$

as $n\to\infty$. Let further $\psi(t) := \mathbb{E}((Y-t)\mathbb{1}_{\{Y>t\}})$ and observe that

$$\begin{aligned} \mathbb{E}((Y-\xi_{\tau_n}-t)(\mathbb{1}_{\{Y\leqslant\xi_{\tau_n}+t\}} - \mathbb{1}_{\{Y\leqslant\xi_{\tau_n}\}})) &= \mathbb{E}((Y-\xi_{\tau_n}-t)(\mathbb{1}_{\{Y>\xi_{\tau_n}\}} - \mathbb{1}_{\{Y>\xi_{\tau_n}+t\}})) \\ &= \psi(\xi_{\tau_n}) - \psi(\xi_{\tau_n}+t) - t\overline{F}(\xi_{\tau_n}). \end{aligned}$$

Integrating by parts entails

$$\psi(\xi_{\tau_n}) - \psi(\xi_{\tau_n}+t) = \int_{\xi_{\tau_n}}^{\xi_{\tau_n}+t}\overline{F}(x)dx = \xi_{\tau_n}\overline{F}(\xi_{\tau_n})\int_1^{1+t/\xi_{\tau_n}}\frac{\overline{F}(\xi_{\tau_n}v)}{\overline{F}(\xi_{\tau_n})}dv$$

from which we deduce that

$$\mathbb{E}((Y-\xi_{\tau_n}-t)(\mathbb{1}_{\{Y\leqslant\xi_{\tau_n}+t\}} - \mathbb{1}_{\{Y\leqslant\xi_{\tau_n}\}})) = \overline{F}(\xi_{\tau_n})\left(\xi_{\tau_n}\int_1^{1+t/\xi_{\tau_n}}\frac{\overline{F}(\xi_{\tau_n}v)}{\overline{F}(\xi_{\tau_n})}dv - t\right).$$

We now control the term between brackets as follows: let $I_n(u) = (0, |u|\xi_{\tau_n}/v_n)$, note that $t/\xi_{\tau_n} \to 0$ uniformly in $t$ such that $|t| \in I_n(u)$, and therefore

$$\begin{aligned} & \sup_{|t|\in I_n(u)} \left| \frac{\xi_{\tau_n}}{t} \int_1^{1+t/\xi_{\tau_n}} \frac{\overline{F}(\xi_{\tau_n} v)}{\overline{F}(\xi_{\tau_n})} dv - 1 \right| \\ \leqslant & \sup_{|t|\in I_n(u)} \frac{\xi_{\tau_n}}{|t|} \left| \int_1^{1+t/\xi_{\tau_n}} \left[ \frac{\overline{F}(\xi_{\tau_n} v)}{\overline{F}(\xi_{\tau_n})} - v^{-1/\gamma} \right] dv \right| + \sup_{|t|\in I_n(u)} \frac{\xi_{\tau_n}}{|t|} \left| \int_1^{1+t/\xi_{\tau_n}} \left( v^{-1/\gamma} - 1 \right) dv \right| \\ \to & \ 0 \end{aligned}$$

by the uniform convergence theorem for regularly varying functions (see Theorem 1.5.2 in Bingham *et al.*, 1987, p.22) and the convergence $v_n \to \infty$. Consequently, by (E.4):

$$\mathbb{E}((Y - \xi_{\tau_n} - t)(\mathbb{1}_{\{Y \leqslant \xi_{\tau_n} + t\}} - \mathbb{1}_{\{Y \leqslant \xi_{\tau_n}\}})) = t(1 - \tau_n) r_n(t) \tag{E.50}$$

with $r_n(t) \to 0$ uniformly in $t$ such that $|t| \in I_n(u)$. Combine (E.47), (E.48), (E.49) and (E.50) to get

$$\begin{aligned} T_{2,n}(u) &= \frac{u^2}{2\gamma}(1 + \mathrm{o}(1)) + T_{3,n}(u), \qquad (E.51) \\ \text{with } T_{3,n}(u) &:= -\frac{v_n^2}{n(1-\tau_n)\xi_{\tau_n}^2} \sum_{i=1}^n \int_0^{u\xi_{\tau_n}/v_n} [\mathcal{S}_{n,i}(\xi_{\tau_n} + t) - \mathcal{S}_{n,i}(\xi_{\tau_n})] dt \end{aligned}$$

where the $\mathcal{S}_{n,i}(v) := \varphi_{\tau_n}(Y_i - v) - \mathbb{E}(\varphi_{\tau_n}(Y_i - v))$ are independent copies of $\mathcal{S}_n(v) := \varphi_{\tau_n}(Y - v) - \mathbb{E}(\varphi_{\tau_n}(Y - v))$. Thus

$$\mathrm{Var}(T_{3,n}(u)) = \frac{v_n^4}{n(1-\tau_n)^2 \xi_{\tau_n}^4} \mathrm{Var}\left( \int_0^{u\xi_{\tau_n}/v_n} [\mathcal{S}_n(\xi_{\tau_n} + t) - \mathcal{S}_n(\xi_{\tau_n})] dt \right).$$

We now notice that for any $v$, $\mathcal{S}_n(v)$ is centred and thus

$$\mathrm{Var}(T_{3,n}(u)) = \frac{v_n^4}{n(1-\tau_n)^2 \xi_{\tau_n}^4} \int_{[0,\, u\xi_{\tau_n}/v_n]^2} \mathbb{E}([\mathcal{S}_n(\xi_{\tau_n} + s) - \mathcal{S}_n(\xi_{\tau_n})][\mathcal{S}_n(\xi_{\tau_n} + t) - \mathcal{S}_n(\xi_{\tau_n})]) ds\, dt$$

(here the Fubini theorem was used to switch integrals and expectation, based on the fact that $Y$ has a finite variance). By the Cauchy-Schwarz inequality,

$$\mathrm{Var}(T_{3,n}(u)) \leqslant \frac{v_n^4}{n(1-\tau_n)^2 \xi_{\tau_n}^4} \left( \int_0^{u\xi_{\tau_n}/v_n} \sqrt{\mathbb{E}(|\mathcal{S}_n(\xi_{\tau_n} + t) - \mathcal{S}_n(\xi_{\tau_n})|^2)}\, dt \right)^2. \tag{E.52}$$

Applying Lemma 3 in Daouia *et al.* (2018), we get for any $t$

$$|\mathcal{S}_n(\xi_{\tau_n} + t) - \mathcal{S}_n(\xi_{\tau_n})| \leqslant 2|t|[1 - \tau_n + \mathbb{1}_{\{Y > \xi_{\tau_n} + \min(t,0)\}} + \overline{F}(\xi_{\tau_n} + \min(t, 0))].$$

Using the inequality $|a+b+c|^2 \leqslant 3(a^2+b^2+c^2)$ yields

$$\mathbb{E}(|\mathcal{S}_n(\xi_{\tau_n}+t)-\mathcal{S}_n(\xi_{\tau_n})|^2) \leqslant 12t^2[(1-\tau_n)^2+\overline{F}(\xi_{\tau_n}+\min(t,0))(1+\overline{F}(\xi_{\tau_n}+\min(t,0)))]. \quad \text{(E.53)}$$

The regular variation property of $\overline{F}$ and the convergence $n(1-\tau_n) \to \infty$ entail:

$$\sup_{|s|\in I_n(u)} |\overline{F}(\xi_{\tau_n}+s)-\overline{F}(\xi_{\tau_n})| = \overline{F}(\xi_{\tau_n}) \sup_{|s|\in I_n(u)} \left| \frac{\overline{F}(\xi_{\tau_n}+s)}{\overline{F}(\xi_{\tau_n})} - 1 \right| = \mathrm{o}(\overline{F}(\xi_{\tau_n})) = \mathrm{o}(1-\tau_n) \quad \text{(E.54)}$$

in view of (E.4). Finally, using (E.4) once again and combining (E.52), (E.53) and (E.54) yields

$$\mathrm{Var}(T_{3,n}(u)) = \mathrm{O}\left( \frac{v_n^4}{n(1-\tau_n)\xi_{\tau_n}^4} \left| \int_0^{u\xi_{\tau_n}/v_n} |t|\, dt \right|^2 \right) = \mathrm{O}\left( \frac{1}{n(1-\tau_n)} \right) \to 0$$

as $n \to \infty$. Whence the convergence $T_{3,n}(u) \xrightarrow{\mathbb{P}} 0$; combining (E.45), (E.46) and (E.51) entails

$$\forall u \in \mathbb{R},\ \psi_n(u) \xrightarrow{\mathbb{P}} \frac{u^2}{2\gamma} \quad \text{as} \quad n \to \infty.$$

We conclude by noting that $(\psi_n)$ is a random sequence of continuous convex functions and its pointwise limit defines a nonrandom continuous convex function of $u$ which has a unique minimum at $u^\star = 0$. Applying Theorem 5 in Knight (1999) completes the proof. ∎

The fifth lemma is the key to the computation of the various terms appearing in the implicit relationship linking the tail expectile process to the tail parameters.

**Lemma 5.** *Suppose that $\mathbb{E}|Y_-| < \infty$. Assume further that condition $\mathcal{C}_2(\gamma,\rho,A)$ holds for some $0 < \gamma < 1/2$. Let $k = k(n) \to \infty$ be such that $k/n \to 0$ and $\sqrt{k}A(n/k) = \mathrm{O}(1)$. Then we have, for any $\delta > 0$ sufficiently small:*

$$\sup_{k^{-1+\delta} \leqslant s \leqslant 1} s^\gamma \left| \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} - (\gamma^{-1}-1)^{-\gamma} s^{-\gamma} \right| \xrightarrow{\mathbb{P}} 0.$$

**Proof of Lemma 5.** All the $\mathrm{o}_\mathbb{P}$ and $\mathrm{O}_\mathbb{P}$ terms in the present proof should be understood as uniform in $s \in [k^{-1+\delta}, 1]$; moreover, we work throughout this proof with the version of the tail expectile process induced by the version of the empirical process $\widehat{F}_n$ leading to (E.6). Recall that any Brownian motion $W$ satisfies, for any $\eta > 0$:

$$\forall c > 0,\ \sup_{0<t\leqslant c} t^{-1/2+\eta}|W(t)| < \infty \quad \text{almost surely.}$$

It then comes as a consequence of (E.5) that

$$\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} W_n(x^{-1/\gamma})\, dx = \mathrm{O}_\mathbb{P}\left( \int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{(\eta-1/2)/\gamma}\, dx \right).$$

Moreover, since $\sqrt{k}A(n/k)$ remains bounded:

$$A(n/k)\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{-1/\gamma}\frac{x^{\rho/\gamma}-1}{\gamma\rho}\,dx = \mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{(\eta-1/2)/\gamma}\,dx\right).$$

All in all, combining these two bounds with (E.6) gives:

$$\begin{aligned}&\int_{\widetilde{\xi}_{1-ks/n}}^{\infty}\widehat{\overline{F}}_n(u)du\\ =\;& \frac{k}{n}q_{1-k/n}\left(\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{-1/\gamma}\,dx + \mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\int_{\widetilde{\xi}_{1-ks/n}/q_{1-k/n}}^{\infty} x^{(\eta-1/2)/\gamma}\,dx\right)\right)\end{aligned}$$

or equivalently

$$\int_{\widetilde{\xi}_{1-ks/n}}^{\infty}\widehat{\overline{F}}_n(u)du = \frac{k}{n}q_{1-k/n}\times\frac{\gamma}{1-\gamma}\left(\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{1-1/\gamma} + \mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{1-(1/2-\eta)/\gamma}\right)\right).$$

Plugging this back into (E.3) entails

$$\frac{s}{1-2ks/n}\left(\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} - \frac{\overline{Y}_n}{q_{1-k/n}}\right) = \frac{\gamma}{1-\gamma}\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{1-1/\gamma} + \mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{1-(1/2-\eta)/\gamma}\right).$$

Note that $\overline{Y}_n \xrightarrow{\mathbb{P}} \mathbb{E}(Y) < \infty$ by the law of large numbers, and $\widetilde{\xi}_{1-ks/n} \geqslant \widetilde{\xi}_{1-k/n} \xrightarrow{\mathbb{P}} +\infty$ by Lemma 4. Therefore

$$s\,(1+\mathrm{o}_{\mathbb{P}}(1)) = \frac{\gamma}{1-\gamma}\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{-1/\gamma} + \mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}\left[\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}}\right]^{-(1/2-\eta)/\gamma}\right).$$

Define now a random process $s \mapsto R_n(s)$ by the equality

$$\frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} = (\gamma^{-1}-1)^{-\gamma}s^{-\gamma}(1+R_n(s)).$$

In particular, $1+R_n(s) > 0$ for any $s \in (0,1]$, and

$$1+\mathrm{o}_{\mathbb{P}}(1) = (1+R_n(s))^{-1/\gamma} + \mathrm{O}_{\mathbb{P}}\left(\frac{1}{\sqrt{k}}s^{-1/2-\eta}(1+R_n(s))^{-(1/2-\eta)/\gamma}\right).$$

We infer from this equality that, uniformly in $s \in [k^{-1+\delta},1]$ for $\delta = \delta(\eta) = 4\eta/(4\eta+1) > 0$,

$$1+\mathrm{o}_{\mathbb{P}}(1) = (1+R_n(s))^{-1/\gamma} + \mathrm{o}_{\mathbb{P}}\left((1+R_n(s))^{-(1/2-\eta)/\gamma}\right).$$

It directly follows from this last identity, whose left-hand side should remain bounded uniformly in $s$, that $1+R_n(s)$ must remain bounded away from 0, uniformly in $s \in [k^{-1+\delta},1]$

with arbitrarily large probability as $n \to \infty$. The fact that the left-hand side converges in probability to 1 uniformly in $s$ now entails that $1 + R_n(s)$ should do so as well, which yields

$$\sup_{k^{-1+\delta} \leqslant s \leqslant 1} |R_n(s)| \xrightarrow{\mathbb{P}} 0 \quad \text{as} \quad n \to \infty.$$

Equivalently

$$\sup_{k^{-1+\delta} \leqslant s \leqslant 1} s^{\gamma} \left| \frac{\widetilde{\xi}_{1-ks/n}}{q_{1-k/n}} - (\gamma^{-1} - 1)^{-\gamma} s^{-\gamma} \right| \xrightarrow{\mathbb{P}} 0. \tag{E.55}$$

And since $\eta$ was arbitrarily small, $\delta = 4\eta/(4\eta + 1)$ was arbitrarily small as well, concluding the proof. ∎

The final lemma is a technical result on second-order regular variation which will be used several times in the proofs of Theorems 2 and 3.

**Lemma 6.** *Assume that condition $\mathcal{C}_2(\gamma, \rho, A)$ holds with $\gamma > 0$. Then one can find a function $B$, asymptotically equivalent to $t \mapsto A(1/\overline{F}(t))$ in a neighbourhood of infinity, satisfying the following: for any $\varepsilon$, $\delta > 0$ there exists $t_0 = t_0(\varepsilon, \delta) > 0$ such that for $t$, $tx \geqslant t_0$,*

$$\left| \frac{1}{B(t)} \left( \frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\gamma} \right) - x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\gamma\rho} \right| \leqslant \varepsilon x^{-(1-\rho)/\gamma} \max\left(x^{-\delta}, x^{\delta}\right).$$

**Proof of Lemma 6.** Note that, according to Theorem 2.3.9 in de Haan and Ferreira (2006), condition $\mathcal{C}_2(\gamma, \rho, A)$ is equivalent to

$$\forall x > 0, \ \lim_{t \to \infty} \frac{1}{A(1/\overline{F}(t))} \left( \frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\gamma} \right) = x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\gamma\rho}.$$

Define $f(x) = x^{1/\gamma} \overline{F}(x)$; it is straightforward to show that this condition entails

$$\forall x > 0, \ \lim_{t \to \infty} \frac{f(tx) - f(t)}{\gamma^{-2} f(t) A(1/\overline{F}(t))} = \frac{x^{\rho/\gamma} - 1}{\rho/\gamma}.$$

Then, by Theorem B.2.18 in de Haan and Ferreira (2006), one can find a function $t \mapsto a_0(t)$, equivalent to $t \mapsto \gamma^{-2} f(t) A(1/\overline{F}(t))$ in a neighbourhood of infinity, such that for any $\varepsilon$, $\delta > 0$ there exists $t_0 = t_0(\varepsilon, \delta) > 0$ with

$$t, tx \geqslant t_0 \Rightarrow \left| \frac{f(tx) - f(t)}{a_0(t)} - \frac{x^{\rho/\gamma} - 1}{\rho/\gamma} \right| \leqslant \gamma^2 \varepsilon x^{\rho/\gamma} \max\left(x^{-\delta}, x^{\delta}\right).$$

Multiplying through by $\gamma^{-2} x^{-1/\gamma}$ and recalling that $f(x) = x^{1/\gamma} \overline{F}(x)$, we get

$$t, tx \geqslant t_0 \Rightarrow \left| \frac{1}{\gamma^2 a_0(t)/f(t)} \left( \frac{\overline{F}(tx)}{\overline{F}(t)} - x^{-1/\gamma} \right) - x^{-1/\gamma} \frac{x^{\rho/\gamma} - 1}{\gamma\rho} \right| \leqslant \varepsilon x^{-(1-\rho)/\gamma} \max\left(x^{-\delta}, x^{\delta}\right).$$

Setting $B(t) = \gamma^2 a_0(t)/f(t)$ gives precisely the desired inequality. ∎

# References


Bellini, F. and Di Bernardino, E. (2017). Risk Management with Expectiles, *The European Journal of Finance* **23**: 487–506.

Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987). *Regular Variation*, Cambridge University Press.

Cai, J.J., Einmahl, J.H.J., de Haan, L. and Zhou, C. (2015). Estimation of the marginal expected shortfall: the mean when a related variable is extreme, *Journal of the Royal Statistical Society: Series B* **77**: 417–442.

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*, John Wiley & Sons, Chichester.

Daouia, A., Girard, S. and Stupfler, G. (2017). Extreme M-quantiles as risk measures: From $L^1$ to $L^p$ optimization, *Bernoulli* (to appear). Available at `https://hal.inria.fr/hal-01585215`.

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

El Methni, J. and Stupfler, G. (2017a). Extreme versions of Wang risk measures and their estimation for heavy-tailed distributions, *Statistica Sinica* **27**: 907–930.

El Methni, J. and Stupfler, G. (2017b). Improved estimators of extreme Wang distortion risk measures for very heavy-tailed distributions, *Econometrics and Statistics* (to appear).

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*, Springer.

Knight, K. (1999). Epi-convergence in distribution and stochastic equi-semicontinuity, technical report, University of Toronto.