Web Appendix to

"Joint Extreme Value-at-Risk and Expected Shortfall Dynamics with a Single Integrated Tail Shape Parameter"

Enzo D'Innocenzo[a], *Andre Lucas*[b], *Bernd Schwaab*[c], *and Xin Zhang*[d]

[a]: *University of Bologna*

[b]: *Vrije Universiteit Amsterdam* and *Tinbergen Institute*

[c]: *European Central Bank, Financial Research*

[d]: *Sveriges Riksbank*

Contents

A Proofs

NIC for the EVT-based VaR

Let the conditional exceedance probability of τ_t be equal to κ, and let $0 < \kappa < \gamma$ denote our extreme tail probability. Consider modeling the right-hand extreme tail. For a given values of $f_i = f$ and $\tau_t = \tau$, we have

$$f_{i+1} = \omega + f + \alpha \left(\ln(1 + y_i) - f\right),$$

$$\begin{aligned}
VaR_{i+1} &= \tau \cdot \left(\frac{\gamma}{\kappa}\right)^{-f_{i+1}} = \tau \cdot \left(\frac{\gamma}{\kappa}\right)^{-\omega-(1-\alpha)f-\alpha\ln(1+y_i)} \\
&= \tau \left(\frac{\gamma}{\kappa}\right)^{-\omega-(1-\alpha)f} \cdot \left(\frac{\gamma}{\kappa}\right)^{-\alpha\ln(1+y_i)} = C \cdot \left(\frac{\gamma}{\kappa}\right)^{-\alpha\ln(1+y_i)} \\
&= C \cdot \exp\left(\ln\left(\frac{\gamma}{\kappa}\right)\right)^{-\alpha\ln(1+y_i)} = C \cdot \exp\left(-\alpha\ln(1+y_i)\ln\left(\frac{\gamma}{\kappa}\right)\right) \\
&= C \cdot \exp\left(\ln(1+y_i)\right)^{-\alpha\ln\left(\kappa/\gamma\right)} = C \cdot (1+y_i)^{-\alpha\ln\left(\kappa/\gamma\right)} \\
&= C \cdot \left(X_t/\tau\right)^{\alpha\ln\left(\kappa/\gamma\right)} = \tilde{C} \cdot X_t^{-\alpha\ln\left(\kappa/\gamma\right)},
\end{aligned}$$

for X_t exceeding the threshold τ, i.e., $X_t > \tau > 0$. The shape of the news impact curve for the VaR based on the EVT approach is thus concave as long as $\alpha\ln(\kappa/\gamma) < 1$. Note that for the plots in Section 2.3 we have re-cast our EVT approach to the extreme left-hand tail to make it directly comparable to the approach of Patton et al. (2019).

Preliminary results

Lemma A.1. Under Assumptions 1 and 2, the inequality

$$\mathbb{E}\left[\ln\left|1 + \alpha\left(\epsilon - 1\right)\right|\right] < 0,$$

is always satisfied.

Proof. For $b = \alpha/(1-\alpha) > 0$, we have

$$\begin{aligned}
\mathbb{E}\left[\ln|1+\alpha(\epsilon-1)|\right] &= \ln(1-\alpha) + \int_0^\infty \ln(1+bx)e^{-x}\mathrm{d}x \\
&= \ln(1-\alpha) - \left[\ln(1+bx)e^{-x}\right]_0^\infty + \int_0^\infty \frac{b}{1+bx}e^{-x}\mathrm{d}x \\
&= \ln(1-\alpha) + \int_0^\infty \frac{b}{1+bx}e^{-x}\mathrm{d}x \\
&= \ln(1-\alpha) + e^{1/b}\int_{1/b}^\infty \frac{e^{-x}}{x}\mathrm{d}x \\
&= \ln(1-\alpha) - e^{\alpha^{-1}-1}\ \mathrm{Ei}\left(1-\alpha^{-1}\right) \\
&< 0,
\end{aligned}$$

for $0 < \alpha < 1$, where $\mathrm{Ei}(z) = -\int_{-z}^\infty t^{-1}\, e^{-t}\, \mathrm{d}t$ denotes the exponential integral. ∎

Proof of Theorem 1

Since $\ln(1+y_i) = f_i(\theta_0)\epsilon_i$, we can write the score-driven filter as

$$\begin{aligned}
\hat{f}_{i+1}(\theta) =& \omega + \hat{f}_i(\theta) + \alpha\left(\ln(1+y_i) - \hat{f}_i(\theta)\right) \\
=& \omega + \hat{f}_i(\theta) + \alpha\left(f_i(\theta_0)\epsilon_i - \hat{f}_i(\theta)\right) \\
=& \omega + (1-\alpha)\,\hat{f}_i(\theta) + \alpha\epsilon_i f_i(\theta_0).
\end{aligned}$$

When evaluating the process above at the true parameter vector θ_0, we note that the unobserved process $\{f_{i+1}(\theta_0)\}_{i\in\mathbb{Z}}$ satisfies

$$f_{i+1}(\theta_0) = \omega_0 + (1-\alpha_0+\alpha_0\epsilon_i)\, f_i(\theta_0).$$

Note that both $\hat{f}_i(\theta)$ and $f_i(\theta_0)$ are embedded in the stochastic recurrence equations (SREs) of the form $\hat{f}_{i+1}(\theta) = \hat{\phi}(\hat{f}_i(\theta), y_i, \theta)$ and $f_{i+1}(\theta_0) = \phi(f_i(\theta_0), \epsilon_i, \theta_0)$, respectively.

Part (i): To prove stationarity and ergodicity (SE) of $f_i(\theta_0)$, we apply Theorem 3.1 of Bougerol (1993). We first check the log-moment condition, which is easily satisfied since

$$\begin{aligned}\mathbb{E}\left[\ln^+\left|\phi(\bar{f}_1,\epsilon_i,\theta_0)\right|\right] &\leq \mathbb{E}\left[\ln\omega_0 + \ln\left(1+\frac{(1-\alpha_0+\alpha_0\epsilon_i)}{\omega_0}\bar{f}_1\right)\right] \\ &= \ln\omega_0 + \mathbb{E}\left[\frac{(1-\alpha_0+\alpha_0\epsilon_i)}{\omega_0}\bar{f}_1\right] = \ln\omega_0 + \frac{1}{\omega_0}\bar{f}_1 < \infty,\end{aligned}$$

for all $\bar{f}_1 \in (0,\infty)$, where we have used the fact that ϵ_i is *IID* exponentially distributed with unit mean following Assumption 1. The contraction condition of Bougerol (1993) follows directly, as

$$\mathbb{E}\left[\sup_{\bar{f}}\ln\left|\frac{\partial\phi(\bar{f},\epsilon_i,\theta_0)}{\partial\bar{f}}\right|\right] = \mathbb{E}\left[\ln\left|1+\alpha_0(\epsilon_i-1)\right|\right] < 0,$$

for $\alpha_0 \in (0,1)$ using Lemma A.1 above. Hence, all the conditions of Theorem 3.1 of Bougerol (1993) are satisfied and we conclude that an SE solution $f_i(\theta_0)$ exists and that any initialized sequence converges exponentially fast almost surely ($e.a.s.$) to this unique SE limit. Given $y_i = \exp(f_i(\theta_0)\epsilon_i) - 1$, it follows immediately that y_i is SE by Proposition 4.3 of Krengel (1985).

The existence of moments follows from Lemma 2.4 of Straumann and Mikosch (2006). The almost sure SE representation of $f_i(\theta_0)$ equals

$$f_i(\theta_0) = \omega_0\sum_{i=0}^{\infty}\prod_{j=0}^{i-1}\big(1+\alpha_0(\epsilon_{i-j}-1)\big) > 0. \tag{A.1}$$

Note that $\mathbb{E}\left[(1+\alpha_0(\epsilon_i-1))^q\right] < \infty$ for any finite $q > 0$ given that ϵ_i has a unit exponential distribution. Following to Lemma 2.4 of Straumann and Mikosch (2006), there exists an $0 < \eta < 1$ and a sufficiently small $0 < r \leq q$ such that $\mathbb{E}[(1+\alpha_0(\epsilon_i-1))^r] = \eta$ and thus

$\mathbb{E}[\prod_{j=0}^{i-1}(1+\alpha_0(\epsilon_{i-j}-1))^r] = \eta^i$. Using this, we obtain

$$\mathbb{E}[f_i(\theta_0)^r] = \omega_0^r \sum_{i=0}^{\infty} \mathbb{E}\left[\prod_{j=0}^{i-1}\big(1+\alpha_0(\epsilon_{i-j}-1)\big)^r\right] = \omega_0^r \sum_{i=0}^{\infty} \eta^i < \infty.$$

As $\ln(1+y_i) = f_i(\theta_0)\epsilon_i$, this also directly establishes the existence of a log-moment for $\ln(1+y_i)$ and thus proves the first part of the theorem.

Part (ii): To prove that the filter $\hat{f}_i(\theta)$ is SE, we again apply Theorem 3.1 of Bougerol (1993). The existence of a log-moment is ensured because

$$\begin{aligned}\mathbb{E}\left[\ln^+ \sup_{\theta\in\Theta}\left|\hat{\phi}(\bar{f}_1, y_i, \theta)\right|\right] &\leq C + \ln^+ \sup_{\theta\in\Theta} \omega + \ln^+ \sup_{\theta\in\Theta}(1-\alpha) + \ln^+ \bar{f}_1 + \sup_{\theta\in\Theta} \alpha \mathbb{E}\left[\ln^+ \ln(1+y_i)\right] \\ &< \infty,\end{aligned}$$

for any $\bar{f}_1 \in (0,\infty)$, and where C is a finite constant. The last inequality follows from the assumed $\log^+$ moment for $\ln(1+y_i)$ and is automatically satisfied via part (i) of the theorem if the model is correctly specified.

To establish the contraction property, note that

$$\mathbb{E}\left[\sup_{\theta\in\Theta}\sup_{\bar{f}} \ln\left|\frac{\partial\hat{\phi}(\bar{f}, y_i, \theta)}{\partial \bar{f}}\right|\right] = \mathbb{E}\left[\sup_{\theta\in\Theta}\ln(1-\alpha)\right] = \sup_{\theta\in\Theta}\ln(1-\alpha) < 0,$$

as $0 < \underline{\alpha} \leq \alpha \leq \bar{\alpha} < 1$. We can now use Theorem 3.1 of Bougerol (1993) and conclude that $\hat{f}_i(\theta)$ is asymptotically SE, and converges *e.a.s.* to a unique SE limit $f_i(\theta)$, i.e., $\sup_{\theta\in\Theta}|\hat{f}_i(\theta) - f_i(\theta)| \xrightarrow{e.a.s.} \to 0$.

This establishes the second part of the theorem.

Part (iii): First note that $\hat{f}_i(\theta) \geq \underline{\omega}_f$, and thus

$$\sup_{\theta\in\Theta}\left|\frac{1}{\hat{f}_i(\theta)} - \frac{1}{f_i(\theta)}\right| \quad \leq \quad \underline{\omega}_f^2 \;\cdot\; \sup_{\theta\in\Theta}\left|\hat{f}_i(\theta) - f_i(\theta)\right| \xrightarrow{e.a.s.} 0.$$

It then follows directly from Lemma 2.1 of Straumann and Mikosch (2006) that $\hat{z}_i^f(\theta) \xrightarrow{e.a.s.} z_i^f(\theta)$ uniformly over $\theta \in \Theta$ if $\mathbb{E}[\ln^+ f_i(\theta_0)] < \infty$. The latter follows immediately from Part (i) above.

The boundedness of the moments follows along the same lines as Lemma A3 of Francq and Zakoïan (2012) by replacing their $a(\eta_t) = \beta_{FZ} + \alpha_{FZ}\eta_t^2$ for *IID* η_t with zero mean, unit variance, and $P(\eta_t^2 = 1) < 1$, by our $1 - \alpha + \alpha\epsilon_i$ for *IID* unit exponential ϵ_i, such that $0 < \beta_{FZ} = 1 - \alpha < 0$ and $\alpha_{FZ} = \alpha$, where β_{FZ} and α_{FZ} denote the parameters in the parameterization of Francq and Zakoïan (2012). Similarly, the boundedness of the inverse moment follows directly along the lines of Lemma 6 of Lee and Hansen (1994).

Proof of Theorem 3

Consistency: We show consistency by verifying the conditions in Theorem 3.4 of White (1994) with respect to the sequence $\{\hat{\mathcal{Q}}_{n_T}(\theta)\}_{n\in\mathbb{N}}$ as defined in (9). Specifically: (i) The parameter space Θ is compact; (ii) $\{\hat{\mathcal{Q}}_{n_T}(\theta)\}_{n\in\mathbb{N}}$ is a sequence of random functions continuous on Θ almost surely; (iii) $\hat{\mathcal{Q}}_{n_T}(\theta) = n^{-1}\sum_{i=1}^{n_T}\hat{Q}_i(\theta) \to \bar{\mathcal{Q}}(\theta) := \mathbb{E}\left[Q_i(\theta)\right]$ as $n_T \to \infty$ almost surely; and (iv) $\{\bar{\mathcal{Q}}(\theta) : \Theta \mapsto \mathbb{R}\}$ has an identifiably unique maximizer $\theta_0 \in \Theta$, that is, $\bar{\mathcal{Q}}(\theta_0) > \bar{\mathcal{Q}}(\theta)\ \forall\theta \neq \theta_0$.

Condition (i) holds by assumption, whereas (ii) trivially follows by continuity of $\{\hat{z}_i^f(\theta)\}_{i\in\mathbb{Z}}$ and $\{z_i^f(\theta)\}_{i\in\mathbb{Z}}$. Furthermore, from Theorem 1 we obtain that $\mathbb{E}\left[\sup_{\theta\in\Theta}\left|\hat{Q}_i(\theta)\right|\right] < \infty$ and $\mathbb{E}\left[\sup_{\theta\in\Theta}\left|Q_i(\theta)\right|\right] < \infty$. Theorem 1 also ensures that the process $\{\hat{z}_i^f(\theta)\}_{i\in\mathbb{Z}}$ converges e.a.s. to its stationary and ergodic limit $\{z_i^f(\theta)\}_{i\in\mathbb{Z}}$. We thus have

$$\sup_{\theta\in\Theta}\left|\hat{Q}_i(\theta) - Q_i(\theta)\right| \leq \sup_{\theta\in\Theta}\left|\ln\hat{z}_i^f(\theta) - \ln z_i^f(\theta)\right| - \epsilon_i \cdot \sup_{\theta\in\Theta}\left|\hat{z}_i^f(\theta) - z_i^f(\theta)\right|.$$

By the mean value theorem, there exist an intermediate point $\hat{f}_i^\star(\theta)$ between $\hat{f}_i(\theta)$ and $f_i(\theta)$ such that, using Lemma 2.1 of Straumann and Mikosch (2006), we obtain that

$$\sup_{\theta\in\Theta}\left|\ln \hat{z}_i^f(\theta) - \ln z_i^f(\theta)\right| = \sup_{\theta\in\Theta}\left|\ln \hat{f}_i\theta) - \ln f_i(\theta)\right| = \sup_{\theta\in\Theta}\left|\frac{1}{f_i^\star(\theta)}\right| \sup_{\theta\in\Theta}\left|\hat{f}_i(\theta) - f_i(\theta)\right|$$
$$\leq \frac{1}{\underline{\omega}_f}\sup_{\theta\in\Theta}\left|\hat{f}_i(\theta) - f_i(\theta)\right| \xrightarrow{e.a.s.} 0.$$

Since $\mathbb{E}\left[\epsilon_i\right] = 1$ by Assumption 1, we can again apply Lemma 2.1 of Straumann and Mikosch (2006) to get

$$\epsilon_i \cdot \sup_{\theta\in\Theta}\left|\hat{z}_i^f(\theta) - z_i^f(\theta)\right| \xrightarrow{e.a.s.} 0.$$

It thus follows that

$$\sup_{\theta\in\Theta}\left|\hat{Q}_i(\theta) - Q_i(\theta)\right| \xrightarrow{e.a.s.} 0, \tag{A.2}$$

with $\mathbb{E}\left[\sup_{\theta\in\Theta}\left|Q_i(\theta)\right|\right] < \infty$. Let $\mathcal{Q}_{n_T}(\theta) = \sum_{i=1}^{n_T} \ln z_i^f(\theta) - \epsilon_i\left(z_i^f(\theta) - 1\right)$ be the SE limit of $\hat{\mathcal{Q}}_{n_T}(\theta)$. Now, from the triangle inequality

$$\sup_{\theta\in\Theta}\left|\frac{1}{n_T}\hat{\mathcal{Q}}_{n_T}(\theta) - \bar{\mathcal{Q}}(\theta)\right| \leq \sup_{\theta\in\Theta}\left|\frac{1}{n_T}\hat{\mathcal{Q}}_{n_T}(\theta) - \mathcal{Q}_{n_T}(\theta)\right| + \sup_{\theta\in\Theta}\left|\frac{1}{n_T}\mathcal{Q}_{n_T}(\theta) - \bar{\mathcal{Q}}(\theta)\right|.$$

The first term on the RHS vanishes almost surely using Lemma 2.1 of Straumann and Mikosch (2006) and (A.2). For the second term, we can apply the ULLN for stationary and ergodic sequences of Rao (1962). As a result, we have

$$\lim_{n_T\to\infty}\frac{1}{n_T}\sum_{i=1}^{n_T}\hat{Q}_i(\theta) = \bar{\mathcal{Q}}(\theta) = 1 + \mathbb{E}\left[\ln z_i^f(\theta) - z_i^f(\theta)\right], \tag{A.3}$$

almost surely. For the last equality we have used the fact that ϵ_i is independent of $z_i^f(\theta)$ and $\mathbb{E}\left[\epsilon_i\right] = 1$, as implied by Assumption 1.

Furthermore, $\bar{\mathcal{Q}}(\theta) \leq 0$ with equality if and only if $z_i^f(\theta) = 1$ almost surely, because

$\log(z) - z + 1 \leq 0$ for any $z \in \mathbb{R}^+$, with equality only for $z = 1$. Note that $z_i^f(\theta_0) = 1$. This in turn implies that $\bar{\mathcal{Q}}(\theta_0) = 0$. We conclude the consistency proof by showing that if $z_i^f(\theta) = z_i^f(\theta_0) = 1$ almost surely for every i, then it must be that $\omega = \omega_0$ and $\alpha_0 = \alpha$. To show this, note that it suffices to show that the implication holds for $f_i(\theta) = f_i(\theta_0)$. Therefore, let $f_i(\theta) = f_i(\theta_0)$ almost surely for every i. We then have that

$$
\begin{aligned}
0 &= f_{i+1}(\theta) - f_{i+1}(\theta_0) \\
&= (\omega - \omega_0) + \big(f_i(\theta) - f_i(\theta_0)\big) - \alpha f_i(\theta) + \alpha_0 \, f_i(\theta_0) + (\alpha - \alpha_0)\, \epsilon_i f_i(\theta_0) \\
&= (\omega - \omega_0) + (\alpha - \alpha_0)\ (\epsilon_i - 1)\ f_i(\theta_0),
\end{aligned}
$$

almost surely. Obviously, from Assumption 1, ϵ_i is an $\mathcal{F}_i$-measurable random variable with a non-degenerate distribution, and from Theorem 1 $f_i(\theta_0)$ also has a non-degenerate distribution. As a result, the equality only holds almost surely if both $\alpha = \alpha_0$ and $\omega = \omega_0$.

The strong consistency of the MLE $\hat{\theta}_{n_T}$ in (3) is then guaranteed by noting that all the conditions of Theorem 3.4 in White (1994) are satisfied.

Asymptotic normality: Next, by strong consistency of the MLE $\hat{\theta}_{n_T}$, we obtain that, for large enough n_T the following Taylor expansion is allowed:

$$
\nabla^{\theta} \hat{\mathcal{Q}}_{n_T}(\hat{\theta}_{n_T}) = \nabla^{\theta} \hat{\mathcal{Q}}_{n_T}(\theta_0) + \nabla^{\theta\theta} \hat{\mathcal{Q}}_{n_T}(\theta^{\star}) \left(\hat{\theta}_{n_T} - \theta_0\right), \tag{A.4}
$$

where $\hat{\mathcal{Q}}_{n_T}(\theta) = \sum_{i=1}^{n_T} \hat{Q}_i(\theta)$ and $|\theta^{\star} - \theta_0| < \left|\hat{\theta}_{n_T} - \theta_0\right|$. It is easy to see that since the MLE $\hat{\theta}_{n_T}$ is the maximizer of $\hat{\mathcal{Q}}_{n_T}(\theta)$ and $\theta_0 \in int(\Theta)$ by Assumption 2, we have $\nabla^{\theta} \hat{\mathcal{Q}}_{n_T}(\hat{\theta}_{n_T}) = \mathbf{0}_2$, and hence we can rewrite (A.4) as

$$
\frac{1}{n_T} \nabla^{\theta\theta} \hat{\mathcal{Q}}_{n_T}(\theta^{\star}) \left(\hat{\theta}_{n_T} - \theta_0\right) = -\frac{1}{n_T} \nabla^{\theta} \hat{\mathcal{Q}}_{n_T}(\theta_0). \tag{A.5}
$$

To prove the asymptotic normality of the MLE $\hat{\theta}_{n_T}$ we verify the conditions given in Theorem 6.2 of White (1994). In particular, we let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space, and verify that: (i) The parameter space Θ is a compact subset of $\mathbb{R}^2$ with non-empty interior, (ii) the random function $\hat{\mathcal{Q}}_{n_T}(\theta) : \Omega \times \Theta \mapsto \mathbb{R}$ is continuously differentiable of order 2 on Θ almost surely, (iii) The MLE $\hat{\theta}_{n_T} : \Omega \mapsto \Theta$ is $\mathcal{F}$-measurable and strongly consistent, i.e. $\hat{\theta}_{n_T} \xrightarrow{a.s.} \theta_0$ where $\theta_0 \in int(\Theta)$; (iv) the score vector satisfies $n_T^{-1/2} \nabla^\theta \mathcal{Q}_{n_T}(\theta_0) \Rightarrow \mathcal{N}\left(\mathbf{0}_2, \mathbb{E}\left[\nabla^\theta Q_i(\theta_0) \nabla^\theta Q_i(\theta_0)^\top\right]\right)$; (v) the uniform stochastic convergence of the Hessian matrix, that is, $\sup_{\theta\in\Theta} \left\| \frac{1}{n_T} \nabla^{\theta\theta} \hat{\mathcal{Q}}_{n_T}(\theta) - \nabla^{\theta\theta} \bar{\mathcal{Q}}(\theta) \right\| \xrightarrow{a.s.} 0$, where $\nabla^{\theta\theta} \bar{\mathcal{Q}}(\theta) = \mathbb{E}\left[\nabla^{\theta\theta} Q_i(\theta)\right]$ is finite; (vi) the limit $\nabla^{\theta\theta} \bar{\mathcal{Q}}(\theta)$ evaluated at the true parameter vector θ_0 satisfies $-\bar{\mathcal{Q}}(\theta_0) = -\mathbb{E}\left[\nabla^{\theta\theta} Q_i(\theta_0)\right] = \boldsymbol{\mathcal{I}}(\theta_0)$, where $\boldsymbol{\mathcal{I}}(\theta_0)$ is the Fisher's information matrix.

Obviously, (i)–(iii) are directly implied by Assumptions 1 and 2.

For (iv) it suffices to prove that $\{\nabla^\theta Q_i(\theta_0)\}_{i\in\mathbb{N}}$ is a stationary and ergodic zero-mean martingale difference process with respect to the filtration $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$ with $\mathcal{F}_i = \sigma\{\epsilon_i, \epsilon_{i-1}, \epsilon_{i-2}, \dots\}$. In fact, note that

$$\mathbb{E}\left[\nabla^\theta Q_i(\theta_0) \;\middle|\; \mathcal{F}_{i-1}\right] = \nabla^\theta z_i^f(\theta_0) \left(\frac{1}{z_i^f(\theta_0)} - \mathbb{E}\left[\epsilon_i | \mathcal{F}_{i-1}\right] \right) = \mathbf{0}_2, \tag{A.6}$$

which clearly follows from Assumption 1, the fact that $\nabla^\theta z_i^f(\theta)$ are $\mathcal{F}_{i-1}$-measurable, and that $z_i^f(\theta_0) = 1$ for all i.

Moreover, we can also prove that $\nabla^\theta Q_i(\theta_0)$ is square-integrable since we clearly have $z_i^f(\theta_0) = 1$, and therefore $\nabla^\theta z_i^f(\theta_0) = -\nabla^\theta(1/z_i^f(\theta_0))$ and

$$\begin{aligned}
\mathbb{E}\left[\nabla^\theta Q_i(\theta_0) \nabla^\theta Q_i(\theta_0)^\top\right] &= \mathbb{E}\left[\nabla^\theta z_i^f(\theta_0) \nabla^\theta z_i^f(\theta_0)^\top \left(\frac{1}{z_i^f(\theta_0)} - \epsilon_i \right)^2 \right] \\
&= \mathbb{E}\left[\nabla^\theta z_i^f(\theta_0) \nabla^\theta z_i^f(\theta_0)^\top\right] \\
&= \mathbb{E}\left[\nabla^\theta \left(\frac{1}{z_i^f(\theta_0)} \right) \nabla^\theta \left(\frac{1}{z_i^f(\theta_0)} \right)^\top \right] < \infty,
\end{aligned}$$

as implied by Assumption 1 together with Lemma B.2. Therefore, we are allowed to apply the CLT for square-integrable martingales of Billingsley (1961) to obtain

$$n_T^{-1/2}\,\nabla^\theta \mathcal{Q}_{n_T}(\theta_0) \Rightarrow \mathcal{N}\left(\mathbf{0}_2, \mathbb{E}\left[\nabla^\theta Q_i(\theta_0)\nabla^\theta Q_i(\theta_0)^\top\right]\right),$$

where $\Rightarrow$ denotes convergence in distribution. Next, we focus on (v) and prove the uniform stochastic convergence of the Hessian matrix. From the triangle inequality

$$\begin{aligned}\sup_{\theta\in\Theta}\left\|\frac{1}{n_T}\nabla^{\theta\theta}\hat{\mathcal{Q}}_{n_T}(\theta)-\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta)\right\| \le & \sup_{\theta\in\Theta}\left\|\frac{1}{n_T}\nabla^{\theta\theta}\hat{\mathcal{Q}}_{n_T}(\theta)-\frac{1}{n_T}\nabla^{\theta\theta}\mathcal{Q}_i(\theta)\right\| \\ & +\sup_{\theta\in\Theta}\left\|\frac{1}{n_T}\nabla^{\theta\theta}\mathcal{Q}_{n_T}(\theta)-\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta)\right\|, \qquad \text{(A.7)}\end{aligned}$$

where $\{\nabla^{\theta\theta}Q_i(\theta)\}_{i\in\mathbb{Z}}$ is stationary and ergodic and $\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta) = \mathbb{E}\left[\nabla^{\theta\theta}Q_i(\theta)\right]$ where, by Lemma B.3, $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta}Q_i(\theta)\right\|\right]$ exists.

Hence, from the ULLN of Rao (1962) for stationary and ergodic sequences,

$$\sup_{\theta\in\Theta}\left\|\frac{1}{n_T}\nabla^{\theta\theta}\mathcal{Q}_{n_T}(\theta)-\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta)\right\| \xrightarrow{a.s.} 0.$$

Now, from Theorem 1 and Lemma B.1 together with continuity arguments, we obtain

$$\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta}\hat{Q}_i(\theta)-\nabla^{\theta\theta}Q_i(\theta)\right\| \xrightarrow{e.a.s.} 0.$$

Combining these results, we conclude that (A.7) vanishes almost surely, that is

$$\sup_{\theta\in\Theta}\left\|\frac{1}{n_T}\nabla^{\theta\theta}\hat{\mathcal{Q}}_{n_T}(\theta)-\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta)\right\| \xrightarrow{a.s.} 0.$$

Moreover, by the strong consistency of the MLE, and the fact that $\theta\mapsto\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta)$ is continuous, to complete the proof, we only need to verify (vi) and show that $\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta_0)$ is non-singular.

By applying the law of iterated expectations and Assumption 1, we get that

$$\begin{aligned}
\mathbb{E}\left[\nabla^{\theta\theta}Q_i(\theta_0)\right] &= \mathbb{E}\left[\nabla^{\theta\theta}z_i^f(\theta_0)\left(\frac{1}{z_i^f(\theta_0)}-\epsilon_i\right)-\nabla^{\theta}z_i^f(\theta_0)\nabla^{\theta}z_i^f(\theta_0)^\top\frac{1}{z_i^f(\theta_0)^2}\right]\\
&= \mathbb{E}\left[\mathbb{E}\left[\nabla^{\theta\theta}z_i^f(\theta_0)\left(1-\epsilon_i\right)-\nabla^{\theta}z_i^f(\theta_0)\nabla^{\theta}z_i^f(\theta_0)^\top\;\middle|\;\mathcal{F}_{i-1}\right]\right]\\
&= -\mathbb{E}\left[\nabla^{\theta}z_i^f(\theta_0)\nabla^{\theta}z_i^f(\theta_0)^\top\right],
\end{aligned}$$

since $z_i^f(\theta_0)$, $\nabla^{\theta}z_i^f(\theta_0)$ and $\nabla^{\theta\theta}z_i^f(\theta_0)$ are $\mathcal{F}_{i-1}$-measurable, $z_i^f(\theta_0)=1$ for all i, and $\mathbb{E}\left[\epsilon_i\right]=1$.

Note that the process $\{\nabla^{\theta}z_i^f(\theta_0)\}_{i\in\mathbb{Z}}$ can be written as

$$\begin{aligned}
\nabla^{\theta}z_{i+1}^f(\theta) &= \nabla^{\theta}\left(\frac{f_{i+1}(\theta_0)}{f_{i+1}(\theta)}\right) = -\,\frac{f_{i+1}(\theta_0)}{f_{i+1}(\theta)^2}\nabla^{\theta}f_{i+1}(\theta)\\
&= -\,\frac{f_{i+1}(\theta_0)}{f_{i+1}(\theta)^2}\left(\nabla^{\theta}\omega+\left(\nabla^{\theta}\alpha\right)\left(f_i(\theta_0)\epsilon_i-f_i(\theta)\right)+(1-\alpha)\nabla^{\theta}f_i(\theta)\right)\\
&= \left(\frac{f_{i+1}(\theta_0)}{f_{i+1}(\theta)}\right)^2\left(\frac{f_i(\theta_0)}{f_{i+1}(\theta_0)}\right)\left(w_t(\theta)+(1-\alpha)\nabla^{\theta}z_i^f(\theta)\right), \qquad \text{(A.8)}\\
w_t(\theta) &= \left(-1/f_i(\theta_0)\quad,\quad(1/z_i^f(\theta))-\epsilon_i\right)^\top,
\end{aligned}$$

where $f_i(\theta_0)/f_{i+1}(\theta_0)=1/(1+\alpha_0(\epsilon_i-1)+\omega_0/f_i(\theta_0))$. Since $\{\nabla^{\theta}z_i(\theta_0)\}_{i\in\mathbb{Z}}$ are stationary and ergodic, if $\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta)$ were singular, then $\exists\boldsymbol{\lambda}\in\mathbb{R}^2\backslash\{\mathbf{0}_2\}$ such that $\boldsymbol{\lambda}^\top\nabla^{\theta}z_i^f(\theta_0)=\mathbf{0}_2$ almost surely $\forall i\in\mathbb{N}$. This is obviously ruled out by the functional form of (A.8) and the unit exponential distributional form of ϵ_i and, therefore, it must be that $\boldsymbol{\lambda}^\top\nabla^{\theta}z_i^f(\theta_0)=\mathbf{0}_2\iff\boldsymbol{\lambda}=\mathbf{0}_2$ and thus, $\nabla^{\theta\theta}\bar{\mathcal{Q}}(\theta_0)$ is non-singular. In conclusion, we note that the Fisher's information equality $\mathbb{E}\left[\nabla^{\theta}Q_i(\theta_0)\nabla^{\theta}Q_i(\theta_0)^\top\right]=-\mathbb{E}\left[\nabla^{\theta\theta}Q_i(\theta_0)\right]=\boldsymbol{\mathcal{I}}(\theta_0)$ follows by standard arguments.

Proof of Theorem 4

We prove this Theorem by verifying conditions (C.1)–(C.4) in Lemma 1 of Cavaliere et al. (2025). First, we note that, under our current set of assumptions, $n_T\xrightarrow{a.s.}\infty$ as $T\to\infty$; see the discussion at the start of Section 3.

Now, for condition (C.1), we already proved in Theorem 3, point (iv), that $\{\nabla^{\theta} Q_i(\theta_0)\}_{i\in\mathbb{N}}$ is a stationary and ergodic zero-mean martingale difference process with respect to the filtration $\{\mathcal{F}_i\}_{i\in\mathbb{N}}$. Thus, by the functional CLT for stationary and ergodic martingale differences (see, e.g., Billingsley (1999), Theorem 18.3) together with an application of Theorem 2.1 of Gut (2009), we have

$$\boldsymbol{\mathcal{J}}(\theta_0)^{-1/2} \frac{1}{\sqrt{n_T}} \sum_{i=1}^{[n_T \cdot r]} \nabla^{\theta} Q_i(\theta_0) \Rightarrow W(r) \qquad (\text{for } n_T \to \infty),$$

where $W(r)$ denote a Brownian motion with covariance matrix $\boldsymbol{I}_2$. Therefore, condition (C.1) holds. Moreover, condition (C.2) also holds, as implied by the result obtained in point (iv) in the proof of Theorem 3, and with another application of Theorem 2.1 of Gut (2009).

Condition (C.3) requires that the third-order derivatives of the log-likelihood difference $\nabla^{\theta\theta\theta} Q_i(\theta)$ are uniformly bounded by a stationary and ergodic process over a closed neighborhood of the true parameter vector θ_0, say $\{w_i\}_{i\in\mathbb{Z}}$, that has a bounded first moment, i.e., $\mathbb{E}[w_i] = \bar{w} < \infty$. This is proved in Lemma B.4 below, where we also show that $\frac{1}{n_T}\sum_{i=1}^{n_T} w_i \xrightarrow{a.s.} \bar{w}$, which directly follows by the ergodic theorem and Theorem 2.1 of Gut (2009).

Finally, it remains to check if condition (C.4) is satisfied, which requires that $n_T/T \xrightarrow{a.s.} c$ for some $c \in (0, \infty)$. However, this clearly hold by means of the arguments discussed in Section 3, but with $c \in (0, 1]$ since the number of POTs is always positive, and cannot exceed T, the number of observation.

Then, by Lemma 1 of Cavaliere et al. (2025), we obtain the desired consistency and asymptotic normality of the MLE under random number of POTs.

B Technical lemmas

We define the operators $\nabla^{\theta} = \frac{\partial}{\partial\theta}$, $\nabla^{\theta\theta} = \frac{\partial^2}{\partial\theta\partial\theta^\top}$ and $\nabla^{\theta\theta\theta} = \frac{\partial\,\mathrm{vec}}{\partial\theta^\top}\left(\frac{\partial^2}{\partial\theta\partial\theta^\top}\right)$. In addition, we denote the score vector by $\nabla^{\theta}Q_i(\theta) = \left(\nabla^{\omega}Q_i(\theta), \nabla^{\alpha}Q_i(\theta)\right)^\top \in \mathbb{R}^2$, the Hessian matrix

$$\nabla^{\theta\theta}Q_i(\theta) = \begin{pmatrix} \nabla^{\omega\omega}Q_i(\theta) & \nabla^{\omega\alpha}Q_i(\theta) \\ \nabla^{\omega\alpha}Q_i(\theta) & \nabla^{\alpha\alpha}Q_i(\theta) \end{pmatrix} \in \mathbb{R}^{2\times 2},$$

and the third derivative matrix

$$\nabla^{\theta\theta\theta}Q_i(\theta) = \begin{pmatrix} \nabla^{\omega\omega\omega}Q_i(\theta) & \nabla^{\omega\omega\alpha}Q_i(\theta) \\ \nabla^{\omega\alpha\omega}Q_i(\theta) & \nabla^{\omega\alpha\alpha}Q_i(\theta) \\ \nabla^{\omega\alpha\omega}Q_i(\theta) & \nabla^{\omega\alpha\alpha}Q_i(\theta) \\ \nabla^{\alpha\alpha\omega}Q_i(\theta) & \nabla^{\alpha\alpha\alpha}Q_i(\theta) \end{pmatrix} \in \mathbb{R}^{4\times 2}.$$

It is important to note that differentiating the log-likelihood difference $Q_i(\theta)$ is equivalent to differentiating $\hat{\ell}_i(\theta)$ as defined in (3) since $\ell_i(\theta_0)$ does not depend on θ. Define $z_i^{f,\mathrm{inv}}(\theta) = 1/z_i^f(\theta)$. The elements of the score vector are given by

$$\nabla^{\theta}Q_i(\theta) = \left(z_i^f(\theta)^{-1} - \epsilon_i\right)\nabla^{\theta}z_i^f(\theta) = \left(\epsilon_i - z_i^{f,\mathrm{inv}}(\theta)\right)\frac{\nabla^{\theta}z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^2}, \tag{B.1}$$

the Hessian matrix

$$\begin{aligned}\nabla^{\theta\theta}Q_i(\theta) &= \left(z_i^f(\theta)^{-1} - \epsilon_i\right)\nabla^{\theta\theta}z_i^f(\theta) - \frac{\nabla^{\theta}z_i^f(\theta)\nabla^{\theta}z_i^f(\theta)^\top}{z_i^f(\theta)^2} \\ &= \left(\epsilon_i - z_i^{f,\mathrm{inv}}(\theta)\right)\left(\frac{\nabla^{\theta\theta}z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^2} - 2\frac{\nabla^{\theta}z_i^{f,\mathrm{inv}}(\theta)\nabla^{\theta}z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^3}\right) \\ &\quad - \frac{\nabla^{\theta}z_i^{f,\mathrm{inv}}(\theta)\nabla^{\theta}z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^2},\end{aligned} \tag{B.2}$$

and the third derivative matrix

$$
\begin{aligned}
\nabla^{\theta\theta\theta} Q_i(\theta) = & \left(z_i^f(\theta)^{-1} - \epsilon_i \right) \nabla^{\theta\theta\theta} z_i^f(\theta) - \frac{\left(\operatorname{vec} \nabla^{\theta\theta} z_i^f(\theta) \right) \nabla^{\theta} z_i^f(\theta)^\top}{z_i^f(\theta)^2} \\
& - \frac{\left(\nabla^{\theta\theta} z_i^f(\theta) \otimes \nabla^{\theta} z_i^f(\theta) \right) + \left(\nabla^{\theta} z_i^f(\theta) \otimes \nabla^{\theta\theta} z_i^f(\theta) \right)}{z_i^f(\theta)^2} \\
& + \frac{2 \left(\nabla^{\theta} z_i^f(\theta) \otimes \nabla^{\theta} z_i^f(\theta) \right) \nabla^{\theta} z_i^f(\theta)^\top}{z_i^f(\theta)^3} \\
= & \big(\epsilon_i - z_i^{f,\mathrm{inv}}(\theta) \big) \left(\frac{\nabla^{\theta\theta\theta} z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^2} - 2 \frac{\big(\operatorname{vec} \nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta) \big) \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^3} \right. \\
& \qquad - 2 \frac{\big(\nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta) \otimes \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \big) + \big(\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \otimes \nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta) \big)}{z_i^{f,\mathrm{inv}}(\theta)^3} \\
& \qquad \left. + 6 \frac{\big(\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top \big) \otimes \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^4} \right) \\
& - \operatorname{vec} \left(\frac{\nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^2} - \frac{\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^3} \right) \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top \\
& - \frac{\big(\nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta) \otimes \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \big) + \big(\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \otimes \nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta) \big)}{z_i^{f,\mathrm{inv}}(\theta)^2} \\
& + 2 \frac{\big(\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \otimes \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) \big) \nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^3} .
\end{aligned}
\tag{B.3}
$$

where the first derivative processes $\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) = -z_i^f(\theta)^2 \, \nabla^{\theta} z_i^f(\theta)$ are defined as $\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta) = \left(\nabla^{\omega} z_i^{f,\mathrm{inv}}(\theta), \nabla^{\alpha} z_i^{f,\mathrm{inv}}(\theta) \right)^\top \in \mathbb{R}^2$, where

$$
\nabla^{\theta} z_{i+1}^{f,\mathrm{inv}}(\theta) = \begin{pmatrix} \nabla^{\omega} z_{i+1}^{f,\mathrm{inv}}(\theta) \\ \nabla^{\alpha} z_{i+1}^{f,\mathrm{inv}}(\theta) \end{pmatrix} = \begin{pmatrix} \dfrac{\nabla^{\omega} f_{i+1}(\theta)}{f_{i+1}(\theta_0)} \\ \dfrac{\nabla^{\alpha} f_{i+1}(\theta)}{f_{i+1}(\theta_0)} \end{pmatrix} = \frac{\nabla^{\theta} f_{i+1}(\theta)}{f_{i+1}(\theta_0)}, \tag{B.4}
$$

and

$$\nabla^{\theta} f_{i+1}(\theta) = \phi_i^{\theta}\left(f_i(\theta), \nabla^{\theta} f_i(\theta), \epsilon_i, \theta\right) = \begin{pmatrix} 1 \\ f_i(\theta_0)\epsilon_i - f_i(\theta) \end{pmatrix} + (1-\alpha)\,\nabla^{\theta} f_i(\theta). \tag{B.5}$$

For the second derivative processes $\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)$, we have

$$\nabla^{\theta\theta} z_{i+1}^{f,\text{inv}}(\theta) = \frac{\nabla^{\theta\theta} f_{i+1}(\theta)}{f_{i+1}(\theta_0)} = f_{i+1}(\theta_0)^{-1} \begin{pmatrix} \nabla^{\omega\omega} f_{i+1}(\theta) & \nabla^{\omega\alpha} f_{i+1}(\theta) \\ \nabla^{\omega\alpha} f_{i+1}(\theta) & \nabla^{\alpha\alpha} f_{i+1}(\theta) \end{pmatrix}, \tag{B.6}$$

$$\begin{aligned} \nabla^{\theta\theta} f_{i+1}(\theta) &= \phi_i^{\theta\theta}\left(f_i(\theta), \nabla^{\theta} f_i(\theta), \nabla^{\theta\theta} f_i(\theta), \epsilon_i, \theta\right) \\ &= \nabla^{\theta\top}\left(\begin{pmatrix} 1 \\ f_i(\theta_0)\epsilon_i - f_i(\theta) \end{pmatrix} + (1-\alpha)\,\nabla^{\theta} f_i(\theta)\right) \\ &= -\begin{pmatrix} 0 & \nabla^{\omega} f_i(\theta) \\ \nabla^{\omega} f_i(\theta) & 2\nabla^{\alpha} f_i(\theta) \end{pmatrix} + (1-\alpha)\,\nabla^{\theta\theta} f_i(\theta). \end{aligned} \tag{B.7}$$

Finally, for the third derivative processes $\nabla^{\theta\theta\theta} z_i^{f,\text{inv}}(\theta)$, we have

$$\nabla^{\theta\theta\theta} z_{i+1}^{f,\text{inv}}(\theta) = \frac{\nabla^{\theta\theta\theta} f_{i+1}(\theta)}{f_{i+1}(\theta_0)} = f_{i+1}(\theta_0)^{-1} \begin{pmatrix} \nabla^{\omega\omega\omega} f_i(\theta) & \nabla^{\omega\omega\alpha} f_i(\theta) \\ \nabla^{\omega\alpha\omega} f_i(\theta) & \nabla^{\omega\alpha\alpha} f_i(\theta) \\ \nabla^{\omega\alpha\omega} f_i(\theta) & \nabla^{\omega\alpha\alpha} f_i(\theta) \\ \nabla^{\alpha\alpha\omega} f_i(\theta) & \nabla^{\alpha\alpha\alpha} f_i(\theta) \end{pmatrix}, \tag{B.8}$$

$$\begin{aligned} \nabla^{\theta\theta\theta} f_{i+1}(\theta) &= \phi_i^{\theta\theta\theta}\left(f_i(\theta), \nabla^{\theta} f_i(\theta), \nabla^{\theta\theta} f_i(\theta), \nabla^{\theta\theta\theta} f_i(\theta), \epsilon_i, \theta\right) \\ &= -\begin{pmatrix} 0 & \nabla^{\omega\omega} f_i(\theta) \\ \nabla^{\omega\omega} f_i(\theta) & 2\nabla^{\omega\alpha} f_i(\theta) \\ \nabla^{\omega\omega} f_i(\theta) & 2\nabla^{\omega\alpha} f_i(\theta) \\ 2\nabla^{\omega\alpha} f_i(\theta) & 3\nabla^{\alpha\alpha} f_i(\theta) \end{pmatrix} + (1-\alpha)\nabla^{\theta\theta\theta} f_i(\theta) \end{aligned} \tag{B.9}$$

Similar derivations hold for the initialized counterparts $\hat{z}_{i+1}^{f,\text{inv}}(\theta) = \hat{f}_{i+1}/f_{i+1}(\theta_0)$.

The following Lemma shows that the derivative processes $\nabla^{\theta} z_i^{f,\text{inv}}(\theta)$, $\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)$ and $\nabla^{\theta\theta\theta} z_i^{f,\text{inv}}(\theta)$ of the ratio process $\{z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$ are also asymptotically stationary and ergodic with bounded log-moments.

Lemma B.1. Under the conditions of Theorem 1,

$$\sup_{\theta\in\Theta}\left\|\nabla^{\theta}\hat{z}_i^{f,\text{inv}}(\theta)-\nabla^{\theta} z_i^{f,\text{inv}}(\theta)\right\| \xrightarrow{e.a.s.} 0,$$
$$\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta}\hat{z}_i^{f,\text{inv}}(\theta)-\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)\right\| \xrightarrow{e.a.s.} 0,$$
$$\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta\theta}\hat{z}_i^{f,\text{inv}}(\theta)-\nabla^{\theta\theta\theta} z_i^{f,\text{inv}}(\theta)\right\| \xrightarrow{e.a.s.} 0,$$

for stationary and ergodic derivative processes $\nabla^{\theta} z_i^{f,\text{inv}}(\theta)$, $\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)$ and $\nabla^{\theta\theta\theta} z_i^{f,\text{inv}}(\theta)$.

Proof of Lemma B.1

We note that $\nabla^{\theta}\hat{f}_{i+1}(\theta)$ is a function of both the filter $\hat{f}_i(\theta)$ and its derivative $\nabla^{\theta}\hat{f}_i(\theta)$. To establish the stationarity and ergodicity, we verify the conditions given in Theorem 2.10 of Straumann and Mikosch (2006) for perturbed stochastic recurrence equations (SREs).

It is immediate to see that the conditions S.1 and S.2 stated in Theorem 2.10 of Straumann and Mikosch (2006) are the same as the log-moment and the contraction condition in Theorem 3.1 of Bougerol (1993), and these are clearly implied by Theorem 1, since the mapping function $\phi_i^{\theta}\left(\hat{f}_i(\theta),\nabla^{\theta}\hat{f}_i(\theta),\epsilon_i,\theta\right)$ has finite log-moment and the contraction condition is satisfied because $0<\underline{\alpha}<\alpha<\bar{\alpha}<1$. We then only have to check condition S.3 of Straumann and Mikosch (2006), that ensures that the perturbed and unperturbed SRE converge sufficiently fast for the difference between their asymptotic solutions to vanish exponentially fast.

The condition follows by showing that

$$\sup_{\theta\in\Theta}\left\|\phi_i^{\theta}\left(\hat{f}_i(\theta),\nabla^{\theta}\bar{f}_1(\theta),\epsilon_i,\theta\right)-\phi_i^{\theta}\left(f_i(\theta),\nabla^{\theta}\bar{f}_1(\theta),\epsilon_i,\theta\right)\right\| \xrightarrow{e.a.s.} 0,$$

where $\nabla^{\theta}\bar{f}_1(\theta)$ is some fixed starting point for the derivative recursion. It is straightforward to see that the norm is given by

$$\sup_{\theta\in\Theta}\left\|\begin{pmatrix}0\\ \hat{f}_i(\theta)-f_i(\theta)\end{pmatrix}\right\| \xrightarrow{e.a.s.} 0.$$

As $\nabla^{\theta}\hat{z}_i^{f,\text{inv}}(\theta)=\nabla^{\theta}\hat{f}_i(\theta)/f_i(\theta_0)$, the first result now follows immediately.

The second and third results follow along the same lines, but now using both the SREs defined by $\phi_i^{\theta\theta}(f_i(\theta),\nabla^{\theta}f_i(\theta),\nabla^{\theta\theta}f_i(\theta))$ in (B.7) and $\phi_i^{\theta\theta\theta}\left(f_i(\theta),\nabla^{\theta}f_i(\theta),\nabla^{\theta\theta}f_i(\theta),\nabla^{\theta\theta\theta}f_i(\theta),\epsilon_i,\theta\right)$ in (B.9), and using the e.a.s. convergence of $\hat{f}_i(\theta)$, $\nabla^{\theta}\hat{f}_i(\theta)$ and $\nabla^{\theta\theta}\hat{f}_i(\theta)$ to their SE limits. ∎

Next we introduce another lemma that provides a suitable number of bounded moments for the derivatives of the ratio process $\{z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$, i.e., $\{\nabla^{\theta}z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$, $\{\nabla^{\theta\theta}z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$ and $\{\nabla^{\theta\theta\theta}z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$. As it is clear from equations (B.1), (B.2) and (B.3), this is a necessary step to ensure that the score vector of the log-likelihood is a martingale difference sequence with bounded and constant variance-covariance matrix, the empirical mean of the negative Hessian matrix converges almost surely to a positive-definite constant matrix and, further, that the third derivatives of the log-likelihood function are uniformly bounded.

Lemma B.2. Under the conditions of Theorem 1, the derivatives processes $\{\nabla^{\theta}z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$, $\{\nabla^{\theta\theta}z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$ and $\{\nabla^{\theta\theta\theta}z_i^{f,\text{inv}}(\theta)\}_{i\in\mathbb{Z}}$ have k uniformly bounded moments $\forall k>0$, that is

$$\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta}z_i^{f,\text{inv}}(\theta)\right\|^k\right]<\infty,\ \ \mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta}z_i^{f,\text{inv}}(\theta)\right\|^k\right]<\infty,\ \ \mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta\theta}z_i^{f,\text{inv}}(\theta)\right\|^k\right]<\infty.$$

Proof of Lemma B.2

Consider the SRE (B.4), then we have $\left\|\nabla^{\theta}z_i^{f,\text{inv}}(\theta)\right\|=\left\|\nabla^{\theta}f_i(\theta)/f_i(\theta_0)\right\|$, and

$$\left\|\frac{\nabla^{\theta}f_{i+1}(\theta)}{f_i(\theta_0)}\right\|\leq(1-\alpha)^i\left\|\frac{\nabla^{\theta}\bar{f}_1(\theta)}{\omega_0}\right\|+\sum_{j=0}^{i}(1-\alpha)^j\left\|\begin{pmatrix}\frac{1}{\omega_0}\\ \frac{\epsilon_{i-j}f_{i-j}(\theta_0)}{f_i(\theta_0)}-\frac{f_{i-j}(\theta)}{f_i(\theta_0)}\end{pmatrix}\right\|,$$

so that, for i sufficiently large, we get

$$\left\| \frac{\nabla^{\theta} f_{i+1}(\theta)}{f_i(\theta_0)} \right\| \leq C + \sum_{j=0}^{\infty} (1-\alpha)^j \left\| \frac{\epsilon_{i-j} f_{i-j}(\theta_0)}{f_i(\theta_0)} \right\| + \sum_{j=0}^{\infty} (1-\alpha)^j \left\| \frac{f_{i-j}(\theta)}{f_i(\theta_0)} \right\|.$$

Since Theorem 1 implies that $\mathbb{E}\left[\log^+ \left|\epsilon_i f_i(\theta_0)\right|\right] \leq \log 2 + \mathbb{E}\left[\log^+ |\epsilon_i|\right] + \mathbb{E}\left[\log^+ \left|f_i(\theta_0)\right|\right] < \infty$ and $\mathbb{E}\left[\log^+ \left|f_i(\theta)\right|\right] < \infty$, then by Lemma 2.2 of Berkes et al. (2003) and using the exponential decay of the weights $(1-\alpha)$, it holds that $\sum_{j=0}^{i} (1-\alpha)^j \left\|\epsilon_{i-j} f_{i-j}(\theta_0)\right\| < \infty$ and $\sum_{j=0}^{t} (1-\alpha)^j \left\|f_{i-j}(\theta)\right\| < \infty$ with probability one.

Next, we also note that by Assumption 1 it clearly holds that $\mathbb{E}\left[|\epsilon_i|^r\right] < \infty$ for some sufficiently small $r > 0$, whereas in Theorem 1 we already proved that $\mathbb{E}\left[\left|f_i(\theta_0)\right|^r\right] < \infty$. From this, it follows that $\mathbb{E}\left[\sup_{\theta\in\Theta} \left|f_i(\theta)\right|^r\right] < \infty$, because, for i sufficiently large and the strict stationarity of $\left\{f_i(\theta)\right\}_{i\in\mathbb{Z}}$, we have

$$\begin{aligned} f_{i+1}(\theta) =& \omega + (1-\alpha) f_i(\theta) + \alpha \ln(1+y_i) = \omega + (1-\alpha) f_i(\theta) + \alpha \epsilon_i f_i(\theta_0) \\ =& \frac{\omega}{\alpha} + \sum_{j=0}^{\infty} (1-\alpha)^j \epsilon_{i-j} f_{i-j}(\theta_0), \end{aligned} \tag{B.10}$$

so that, for all $\delta > 0$, an application of the Markov's and Cauchy-Schwartz inequalities yields

$$\mathbb{P}\left(\sup_{\theta\in\Theta} \sum_{j=0}^{\infty} (1-\alpha)^j \epsilon_{i-j} f_{i-j}(\theta_0) > \delta \right) \leq \delta^{-r/2} \mathbb{E}\left[|\epsilon_0|^r\right] \mathbb{E}\left[\left|f_0(\theta_0)\right|^r\right] \sup_{\theta\in\Theta} \sum_{j=0}^{\infty} (1-\alpha)^j < \infty.$$

Moreover, using the almost sure representation in (B.10), we have

$$\left\| \nabla^{\omega} \left(\frac{f_{i+1}(\theta)}{f_{i+1}(\theta_0)} \right) \right\| = \left\| \frac{\nabla^{\omega} f_{i+1}(\theta)}{f_{i+1}(\theta_0)} \right\| = \left\| \left(\alpha\, f_{i+1}(\theta_0)\right)^{-1} \right\| \leq \left\| \underline{\alpha}\, \omega_0 \right\|^{-1}, \tag{B.11}$$

and, using $\epsilon_i f_i(\theta_0) \leq \alpha_0^{-1} f_{i+1}(\theta_0)$,

$$\left\| \frac{\nabla^{\alpha} f_{i+1}(\theta)}{f_{i+1}(\theta_0)} \right\| = \left\| \frac{-\frac{\omega}{\alpha^2} + \sum_{j=0}^{\infty} (1-\alpha + j\,\alpha)(1-\alpha)^{j-1} \epsilon_{i-j} f_{i-j}(\theta_0)}{f_{i+1}(\theta_0)} \right\|$$

$$\leq \frac{\overline{\omega}}{\underline{\alpha}^2\,\omega_0} + \left\| \frac{\sum_{j=0}^{\infty}(1-\alpha+j\,\alpha)\,(1-\alpha)^{j-1}\,\epsilon_{i-j}f_{i-j}(\theta_0)}{\omega_0+\sum_{j=0}^{\infty}(1-\alpha)^j\epsilon_{i-j}f_{i-j}(\theta_0)} \right\|$$

$$= \frac{\overline{\omega}}{\underline{\alpha}^2\,\omega_0} + \left\| \frac{\sum_{j=0}^{\infty}(1-\alpha+j\,\alpha)\,(1-\alpha)^{j-1}\ln(1+y_{i-j})}{\omega_0+\sum_{j=0}^{\infty}(1-\alpha)^j\ln(1+y_{i-j})} \right\|. \tag{B.12}$$

The rest of the proof now follows along the same lines as Lemma 5.2 in Berkes et al. (2003). A similar argument proves the result for the second and third derivative process $\{\nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta)\}_{i\in\mathbb{Z}}$ and $\{\nabla^{\theta\theta\theta} z_i^{f,\mathrm{inv}}(\theta)\}_{i\in\mathbb{Z}}$. ■

Lemma B.3. Under the conditions of Theorem 1, the Hessian processes $\{\nabla^{\theta\theta}Q_i(\theta)\}_{i\in\mathbb{Z}}$ has a uniformly bounded moment, that is

$$\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta}Q_i(\theta)\right\|\right] < \infty.$$

Proof of Lemma B.3

Using equation (B.2), together with a combination of Hölder and Minkowsky inequalities, we obtain

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta}Q_i(\theta)\right\|\right] \\
\leq &\left(\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\epsilon_i - z_i^{f,\mathrm{inv}}(\theta)\right\|^2\right]\right)^{1/2}\left(\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^2} - \frac{\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^3}\right\|^2\right]\right)^{1/2} \\
&+\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^2}\right\|\right] \\
\leq &C\times\left(\left(\mathbb{E}\left[\epsilon_i^2\right]\right)^{1/2} + \left(\mathbb{E}\left[\sup_{\theta\in\Theta}\left|z_i^{f,\mathrm{inv}}(\theta)\right|^2\right]\right)^{1/2}\right) \\
&\times\left(\left(\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\nabla^{\theta\theta} z_i^{f,\mathrm{inv}}(\theta)}{z_i^{f,\mathrm{inv}}(\theta)^2}\right\|^2\right]\right)^{1/2} + \left(\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)\nabla^{\theta} z_i^{f,\mathrm{inv}}(\theta)^\top}{z_i^{f,\mathrm{inv}}(\theta)^3}\right\|^2\right]\right)^{1/2}\right)
\end{aligned}$$

$$+ C \times \mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\frac{\nabla^{\theta} z_i^{f,\text{inv}}(\theta)\nabla^{\theta} z_i^{f,\text{inv}}(\theta)^\top}{z_i^{f,\text{inv}}(\theta)^2}\right\|\right].$$

By Assumption 1 we clearly have that $\mathbb{E}\left[\epsilon_i^2\right] = 1$ whereas by Theorem 1(iii) it holds that $\mathbb{E}\left[\sup_{\theta\in\Theta}\left|z_i^{f,\text{inv}}(\theta)\right|^k\right] = \mathbb{E}\left[\sup_{\theta\in\Theta}\left|1/z_i^f(\theta)\right|^k\right] < \infty$ for any $k > 0$. Furthermore, in Lemma B.2 we proved that the derivative processes satisfy $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta} z_i^{f,\text{inv}}(\theta)\right\|^k\right] < \infty$ and $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)\right\|^k\right] < \infty$ for any $k > 0$, and therefore, by combining all these results, we infer that $\mathbb{E}\left[\sup_{\theta\in\Theta}\left\|\nabla^{\theta\theta} Q_i(\theta)\right\|\right] < \infty$, thus concluding the proof of the Lemma. ■

Lemma B.4. Under the conditions of Theorem 1, there exists a closed neighborhood $N_\delta(\theta_0)$ defined as $N_\delta(\theta_0) = \left\{\theta \in \Theta : \|\theta - \theta_0\| < \delta\right\}$ for some $\delta > 0$ for which the third derivative processes $\{\nabla^{\theta\theta\theta} Q_i(\theta)\}_{i\in\mathbb{Z}}$ satisfies

$$\sup_{\theta\in N_\delta(\theta_0)}\left\|\frac{1}{n_T}\sum_{i=1}^{n_T}\nabla^{\theta\theta\theta} Q_i(\theta)\right\| \leq \frac{1}{n_T}\sum_{i=1}^{n_T} w_i, \tag{B.13}$$

where $\{w_i\}_{i\in\mathbb{Z}}$ is stationary and ergodic and has a bounded moment $\mathbb{E}\left[w_i\right] = \bar{w}$. Furthermore, it holds that $\frac{1}{n_T}\sum_{i=1}^{n_T} w_i \xrightarrow{a.s.} \bar{w}$.

Proof of Lemma B.4

Using the triangular inequality, we have

$$\sup_{\theta\in N_\delta(\theta_0)}\left\|\frac{1}{n_T}\sum_{i=1}^{n_T}\nabla^{\theta\theta\theta} Q_i(\theta)\right\| \leq \frac{1}{n_T}\sum_{i=1}^{n_T}\sup_{\theta\in N_\delta(\theta_0)}\left\|\nabla^{\theta\theta\theta} Q_i(\theta)\right\|$$

Then, by repeated applications of the the Cauchy-Schwarz and the Minkowski inequalities, to the third-order derivative in equation (B.3), we obtain

$$\sup_{\theta\in N_\delta(\theta_0)}\left\|\nabla^{\theta\theta\theta} Q_i(\theta)\right\|$$
$$\leq\left(\sup_{\theta\in N_\delta(\theta_0)}\left\|\epsilon_i - z_i^{f,\text{inv}}(\theta)\right\|^2\right)^{1/2}\left(\sup_{\theta\in N_\delta(\theta_0)}\left\|\frac{\nabla^{\theta\theta\theta} z_i^{f,\text{inv}}(\theta)}{z_i^{f,\text{inv}}(\theta)^2} - 2\frac{\left(\text{vec}\,\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)\right)\nabla^{\theta} z_i^{f,\text{inv}}(\theta)^\top}{z_i^{f,\text{inv}}(\theta)^3}\right.\right.$$

$$
\begin{aligned}
&\qquad - 2\frac{\left(\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta) \otimes \nabla^{\theta} z_i^{f,\text{inv}}(\theta)\right) + \left(\nabla^{\theta} z_i^{f,\text{inv}}(\theta) \otimes \nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)\right)}{z_i^{f,\text{inv}}(\theta)^3} \\
&\qquad + 6\frac{\left(\nabla^{\theta} z_i^{f,\text{inv}}(\theta)\nabla^{\theta} z_i^{f,\text{inv}}(\theta)^\top\right) \otimes \nabla^{\theta} z_i^{f,\text{inv}}(\theta)}{z_i^{f,\text{inv}}(\theta)^4}\Bigg\|^2\Bigg)^{1/2} \\
&+ \left(\sup_{\theta\in N_\delta(\theta_0)}\left\|\frac{\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)}{z_i^{f,\text{inv}}(\theta)^2} - \frac{\nabla^{\theta} z_i^{f,\text{inv}}(\theta)\nabla^{\theta} z_i^{f,\text{inv}}(\theta)^\top}{z_i^{f,\text{inv}}(\theta)^3}\right\|^2\right)^{1/2} \left(\sup_{\theta\in N_\delta(\theta_0)}\left\|\nabla^{\theta} z_i^{f,\text{inv}}(\theta)^\top\right\|^2\right)^{1/2} \\
&\sup_{\theta\in N_\delta(\theta_0)}\left\|\frac{\left(\nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta) \otimes \nabla^{\theta} z_i^{f,\text{inv}}(\theta)\right) + \left(\nabla^{\theta} z_i^{f,\text{inv}}(\theta) \otimes \nabla^{\theta\theta} z_i^{f,\text{inv}}(\theta)\right)}{z_i^{f,\text{inv}}(\theta)^2}\right\| \\
&+ \sup_{\theta\in N_\delta(\theta_0)}\left\|2\frac{\left(\nabla^{\theta} z_i^{f,\text{inv}}(\theta) \otimes \nabla^{\theta} z_i^{f,\text{inv}}(\theta)\right)\nabla^{\theta} z_i^{f,\text{inv}}(\theta)^\top}{z_i^{f,\text{inv}}(\theta)^3}\right\|.
\end{aligned}
$$

Therefore, by using similar arguments as those employed in the proof of Lemma B.3, and some tedious calculations allow us to establish (B.13). Hence, $\sup_{\theta\in N_\delta(\theta_0)}\left\|\nabla^{\theta\theta\theta} Q_i(\theta)\right\| \leq w_i$ where $\{w_t\}_{i\in\mathbb{Z}}$ is stationary and ergodic as desired, and the claimed almost sure convergence directly follows by the ergodic theorem, which ends the proof of the Lemma. ∎

C Derivation of market risk measures

To derive the one-step-ahead VaR, we note that

$$\begin{aligned}\overline{G}(X_t) &= 1 - G(X_t) = \mathbb{P}(X_t > X_t) = \mathbb{P}(X_t > \tau_t)\mathbb{P}(X_t > X_t | X_t > \tau_t) \\ &= \mathbb{P}(X_t > \tau_t)\mathbb{P}(X_t > X_t | (X_t - \tau_t)/\tau_t > 0) = \overline{G}(\tau_t)\overline{F}(y_i),\end{aligned}$$

where the third equality sign uses a standard conditioning argument, and $y_i = (X_{t_i} - \tau_{t_i})/\tau_{t_i}$. We can use this result to obtain $\text{VaR}^{1-\gamma}(X_t \mid \mathcal{F}_{t-1}, \theta) = q_t^{1-\gamma}(X_t)$ by setting

$$\begin{aligned}&\overline{G}(X_{t_i}) = \overline{G}(\tau_{t_i})\overline{F}(y_i) = \gamma \\ \Longleftrightarrow\ &\frac{n_{t_i}}{t_i}(1 + y_i)^{-1/f_i} = \gamma \\ \Longleftrightarrow\ &1 + \tau_{t_i}^{-1}(q_{t_i}^{1-\gamma}(X_{t_i}) - \tau_{t_i}) = \left(\frac{\gamma}{n_{t_i}/t_i}\right)^{-f_i} \\ \Longleftrightarrow\ &q_{t_i}^{1-\gamma}(X_{t_i}) = \tau_{t_i}\left(\frac{\gamma}{n_{t_i}/t_i}\right)^{-f_i},\end{aligned} \tag{C.1}$$

where n_t/t serves as an estimator of $\overline{G}(\tau_t)$. This expression coincides with the expression given in the main text.

The Expected Shortfall $\text{ES}^{1-\gamma}(X_t)$ is given by

$$\begin{aligned}\text{ES}^{1-\gamma}(X_t) &= \frac{1}{\gamma}\int_{1-\gamma}^{1} q_t^s(X_t)\mathrm{d}s \\ &= \frac{\text{VaR}^{1-\gamma}(X_t \mid \mathcal{F}_{t-1}, \theta)}{1 - f_i},\end{aligned} \tag{C.2}$$

which is derived by moving constant terms in front of the integral and noting that

$$\int_{1-\gamma}^{1} (1 - s)^{-f_i}\mathrm{d}s = \frac{\gamma^{1-f_i}}{1 - f_i}$$

for $f_i < 1$.

For completeness, we note that in D'Innocenzo et al. (2024) the market risk measures are given by

$$VaR^{1-\gamma}(X_{t_i}) = q_{t_i}^{1-\gamma}(X_{t_i}) = \tau_{t_i} + \frac{\delta_{t_i}}{f_i} \left[\left(\frac{\gamma}{n_{t_i}/t_i} \right)^{-f_i} - 1 \right], \tag{C.3}$$

$$ES^{1-\gamma}(X_{t_i}) = \frac{\mathrm{VaR}^{1-\gamma}(X_{t_i} \mid \mathcal{F}_{t-1}, \theta)}{1 - f_i} + \frac{\delta_{t_i} - f_i \tau_{t_i}}{1 - f_i}, \tag{C.4}$$

for a tail-scale parameter δ_t. It is easily verified that these expressions collapse to (C.1) and (C.2) if we set $\delta_{t_i} = f_i \tau_{t_i}$ in line with the limiting result of de Haan and Ferreira (2006).

D Additional simulation results

This appendix presents additional simulation results that illustrate the estimation of $\hat{\omega}$ and the difficulty in empirically identifying it. Figure D.1's top panels show that the shape of the log-likelihood is such that there are regularly peaks near the correct value, but also many instances where $\hat{\omega}$ is estimated too low (and thus $-\log(\hat{\omega})$ too high). Figure D.1's bottom panels show that this phenomenon is unrelated to the estimation of $\hat{\alpha}$, which has a similar dispersion around each of the modes around its true value $\alpha = 0.01$. For very large sample sizes, the bimodality disappears and the limiting result of asymptotic normality also applies to $\hat{\omega}$; see Figure D.2.

Figure D.3 presents simulation results that are discussed in the main text.

Figure D.1: Simulation results for $\hat{\omega}$

Upper panels: kernel density estimates of the distribution of the MLE for $-\log(\hat{\omega})$ for two (correctly specified) scenarios (1: τ_t known; 2: $\hat{\tau}_t$ estimated). The kernel density estimates are based on $S = 100$ simulations. Lower panels: scater diagrams of $\hat{\alpha}$ (horizontal) and $-\log\hat{\omega}$ (vertical) for $T = 50,000$.

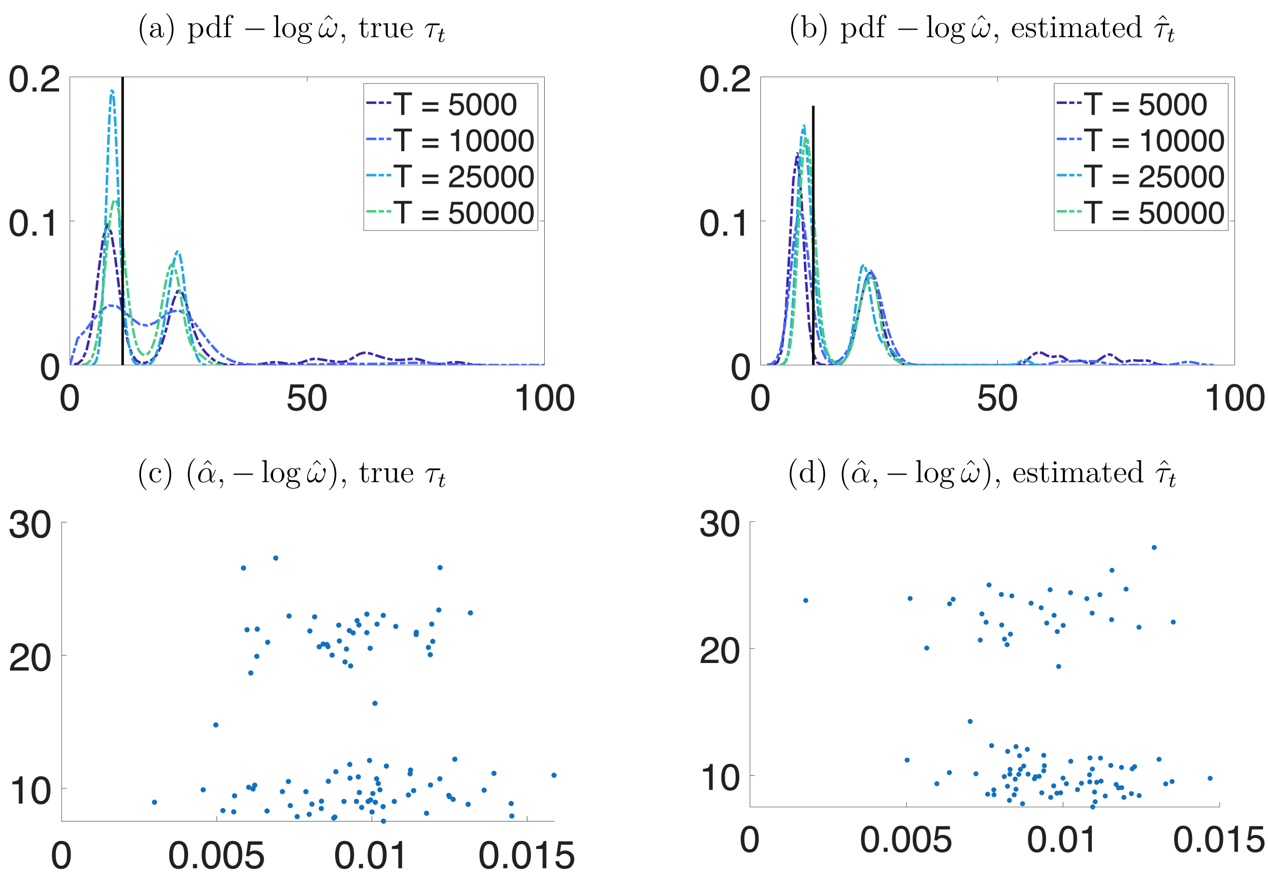


Figure D.2: Asymptotic normality for very large sample sizes, $T = 500,000$

Kernel density plot for $-\log(\hat{\omega})$, vertical line at true value $-\log(1.5 \times 10^{-5})$. Based on 1,000 simulations using true τ_t and $\kappa = 10\%$.

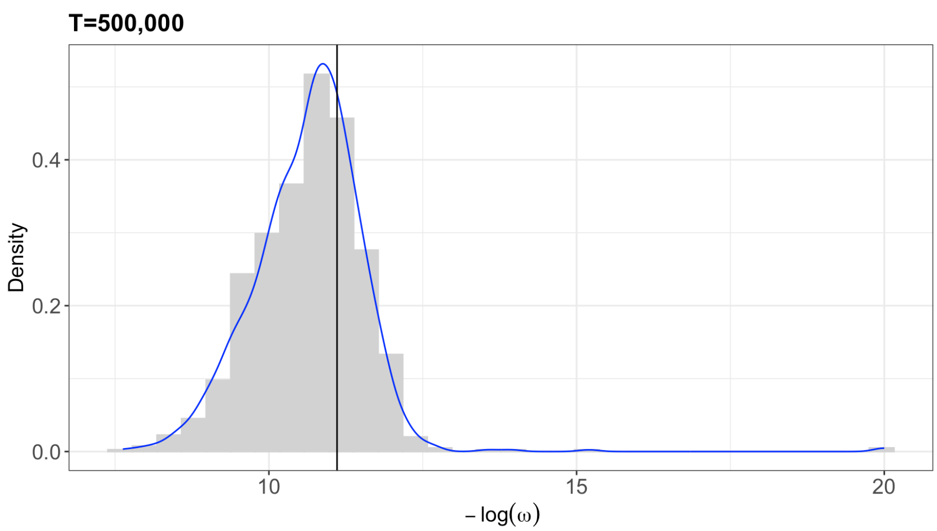
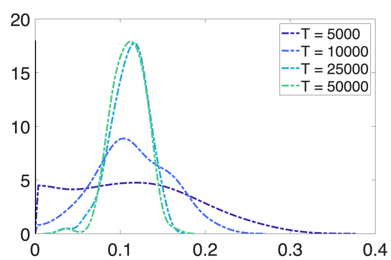


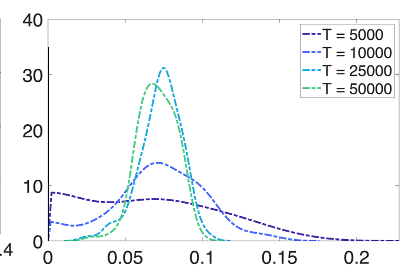
Figure D.3: Simulation results scenario 3

Top panels show kernel density estimates of the distribution of the MLE for $\hat{\alpha}$ for scenario 3 (i.e., a Gaussian AR(1) for the true $\log(f_i)$, such that the model's tail shape dynamics are mis-specified), using estimated thresholds $\hat{\tau}_t$. The kernel density estimates are based on $S = 100$ simulations and are for $\hat{\omega}$ estimated freely or $\hat{\omega}$ fixed at $\bar{\omega} = 10^{-7}$ during estimation. Lower panels show the fit of the filtered $\hat{f}_i(\hat{\theta})$ to the true $f_i(\theta_0)$ in a typical simulation run for two of the sample sizes. Note that the number of POTs is about $\kappa = 10\%$ of the sample size given the mixture setup of the DGP.

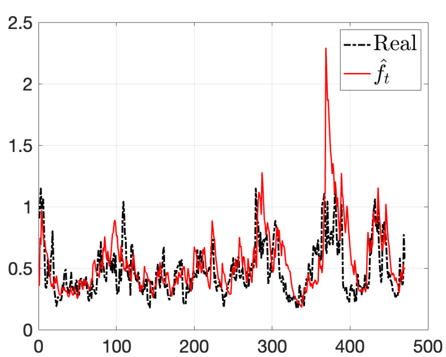
(a) $\hat{\alpha}$ ($\hat{\omega}$ free, $\hat{\tau}_t$)



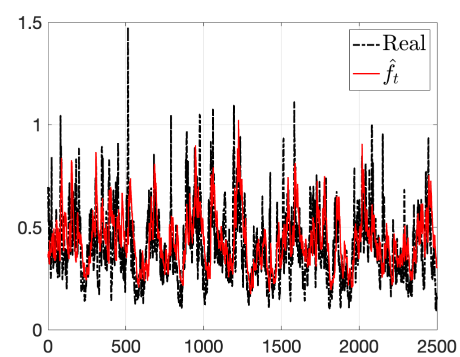
(b) $\hat{\alpha}$ ($\hat{\omega} = \bar{\omega}$, $\hat{\tau}_t$)



(c) fit of filter ($T = 5000$)



(d) fit of filter ($T = 25000$)



References in Online Appendix

Berkes, I., L. Horváth, and P. Kokoszka (2003). GARCH processes: structure and estimation. *Bernoulli 9*(2), 201 – 227.

Patton, A. J., J. F. Ziegel, and R. Chen (2019). Dynamic semiparametric models for Expected Shortfall (and Value-at-Risk). *Journal of Econometrics 211*(2), 388–413.