Estimation risk in conditional expectiles

Marcelo Fernandes

Sao Paulo School of Economics, FGV

Víctor Henriques

Argus Media

Duda Mendes

Sao Paulo School of Economics, FGV

Abstract: We establish the consistency and asymptotic normality of a two-step estimator of conditional expectiles in the context of location-scale models. We first estimate the parameters of the conditional mean and variance by quasi-maximum likelihood and then compute the unconditional expectile of the innovations using the empirical quantiles of the standardized residuals. We show how replacing true innovations with standardized residuals affects the asymptotic variance of the expectile estimator. In addition, we also obtain asymptotic-valid bootstrap-based confidence intervals. Finally, our empirical analysis reveals that conditional expectiles are very interesting alternatives to assess tail risk in cryptomarkets, relative to traditional quantile-based risk measures, such as value at risk and expected shortfall.

Keywords: Asymmetric least squares, bootstrap prediction intervals, quantiles, tail risk.

Acknowledgments: We are indebted to Silvia Gonçalves, Giuseppe Cavaliere, Alan De Genaro, and Edu Horta, as well as to seminar participants at the UBRI Connect 2024 (Zurich, Switzerland), 2024 High Voltage Econometrics Workshop (Lecce, Italy), University of Illinois at Urbana-Champaign, and University of São Paulo. Fernandes acknowledges financial support not only from FAPESP (2023/01728-0) and CNPq (302278/2018-4 and 302153/2022-5), but also from the Silicon Valley Community Foundation through the University Blockchain Research Initiative (UBRI, 2022/199610). The usual disclaimer applies.

1 Introduction

Quantile-based measures, such as value at risk (VaR) and expected shortfall (ES), reign in risk assessments. The Basel II regulatory framework (`https://www.bis.org/publ/bcbs24.pdf`) calculates minimum capital requirements using a plain VaR approach. However, value at risk is not a satisfactory risk measure for two reasons. First, it does not adequately account for diversification gains because it violates the subadditivity property that characterizes coherent risk measures (Artzner, Delbaen, Eber and Heath, 1999; Föllmer and Schied, 2002). Second, as the value at risk corresponds to a quantile of the distribution, it defines what constitutes a tail realization, but without assessing expected severity. Accordingly, the Basel III regulatory framework (`https://www.bis.org/publ/bcbs265.pdf`) suggests gauging risk by expected shortfall. By measuring the expected loss in excess of the value at risk, ES effectively accounts not only for diversification gains, but also for the severity of the tail realization.

The quest for better risk measures is not over yet, though. Gneiting (2011) demonstrates that ES is not elicitable, so that there is no natural backtesting procedure to assess the performance of ES forecasting models. In addition, quantile-based risk measure estimates consider only the relative frequency of observations above or below their corresponding predictions (Daouia, Girard and Stupfler, 2018; Daouia, Gijbels and Stupfler, 2019a), depending heavily on the tail of the loss distribution (Kuan, Yeh and Hsu, 2009; Daouia et al., 2019a). More specifically, ES is too conservative because it restricts attention to a given tail event, whereas VaR is too lenient for it ignores severity. As such, risk measures based on a given quantile arguably either underestimate or overestimate risk exposures.

The alternative class of risk measures based on Newey and Powell's (1987) asymmetric least squares (ALS) has been gaining traction in the literature. In particular, the expectile is a least-squares analogue of quantiles, offering the only law-invariant, coherent risk measure that simultaneously accounts for diversification gains and allows for straightforward backtesting (Ziegel, 2016; Daouia, Stupfler and Usseglio-Carleve, 2023). Interestingly, depends on both the probability and severity of the tail event, making them particularly suitable for actuarial and portfolio-allocation problems.

This paper establishes the consistency and asymptotic normality of a two-step estimator of conditional expectiles in the context of location-scale models. As in Francq and Zakoïan (2015), we first estimate the parameters of the conditional mean and variance by Gaussian quasi-maximum likelihood (QML) and then compute the unconditional expectile of the innovations using the empirical quantiles of the standardized residuals. Naturally, to conduct inference, we must account for the estimation error in the standardized residuals given that we do not observe the true innovations. To do so, we show how substituting standardized residuals for true innovations affects the estimation of the unconditional expectile in the second step. We find that, under the correct specification of the conditional mean and volatility, the impact is only on the asymptotic variance of the ALS estimator. Apart from deriving a closed-form expression for the asymptotic variance taking estimation risk into account, we also show how to obtain asymptotic-valid confidence intervals based on a fixed-design residual bootstrap algorithm.

We corroborate our asymptotic results with Monte Carlo experiments. We follow the setup in Christoffersen and Gonçalves's (2005), in order to reproduce the main stylized facts in asset returns. We find that the fixed-design residual bootstrap performs very well, resulting in distributions that are very close to those based on true innovations, even when the volatility is highly persistent. Both bias and root mean squared error of the two-step estimator decrease sharply with the sample size. More importantly, a backtesting analysis reveals that the coverage rates of the bootstrap-based prediction intervals are reasonably close to nominal levels, ensuring a proper assessment of estimation risk for conditional expectiles. In particular, the prediction intervals of the conditional expectiles are more precise and much tighter than those for the conditional expected shortfall, regardless of the sample size, innovation distribution, and volatility persistence.

Our approach is well in line with the literature. Gao and Song (2008) establish consistency and asymptotic normality of the value-at-risk and expected shortfall based on GARCH standardized residuals. Francq and Zakoïan (2015) extend the asymptotic theory to cover several GARCH-type specifications, including the exponential GARCH (Nelson, 1991), asymmetric power ARCH (Ding, Granger and Engle, 1993), and GJR-GARCH (Glosten, Jagannathan and Runkle, 1993). As for resampling methods, Christoffersen and Gonçalves (2005) and Spierdijk (2016) investigate their large-sample behavior for quantile-based risk measures through Monte

Carlo simulations. The former employs the *m-out-of-n* bootstrap algorithm put forth by Sherman and Carlstein (2004) for ARMA–GARCH model, whereas the latter rests on Pascual, Romo and Ruiz's (2006) recursive-design bootstrap scheme. More recently, Heinemann and Telg (2018) and Beutner, Heinemann and Smeekes (2024) establish the consistency of Cavaliere, Pedersen and Rahbek's (2018) fixed-design residual bootstrap approach for conditional expected-shortfall and value-at-risk measures, respectively.

To the best of our knowledge, this is the first study to assess how the first-step estimation error affects inference on conditional expectiles. Holzmann and Klar (2016) and Krätschmer and Zähle (2017) develop the asymptotic theory for the estimation of unconditional expectiles at a fixed level τ using independent and identically distributed (iid) data. Daouia et al. (2018, 2020) derive the asymptotic distribution of a weighted ALS estimator for extreme expectiles at level τ_T, with $\tau_T \to 1$ as the sample size T grows. Girard, Stupfler and Usseglio-Carleve (2021) extend their analysis to entertain the estimation of conditional extreme expectiles in heavy-tailed heteroskedastic regressions using a two-step approach. In particular, they show that estimating standardized residuals does not affect the asymptotic distribution of extreme quantile estimators because the latter converges at the slower rate $\sqrt{T(1-\tau_T)}$ than the first-step estimation error shrinks to zero. Unfortunately, this is not the case here. The estimation of expectiles at a fixed level τ converges at the same $\sqrt{T}$-rate as the estimation of the conditional mean and variance parameters, and hence the latter affects the asymptotic distribution of the former.

Finally, we empirically assess the performance of conditional expectiles relative to quantile-based risk measures in cryptocurrency markets (for an overview, see Makarov and Schoar, 2022). Our motivation is twofold. First, crypto assets are mainly appealing to investors because of their low correlation with traditional asset classes (see, among others, Bianchi and Babiak, 2022). This suggests that we should not gauge tail risk by looking only at the value at risk for we would miss out any diversification benefit. Second, extreme tail events are quite frequent in crypto markets (Gkillas and Katsiampa, 2018; Scaillet, Treccani and Trevisan, 2018; Borri, 2019; Nguyen, Chevapatrakul and Yao, 2020). This casts doubt on the suitability of expected shortfall as a robust risk measure, given it is very hard to estimate it in a precise and accurate manner under heavy tails. Accordingly, crypto markets make fertile ground for the application of conditional expectiles in the stipulation of minimum capital requirements.

We find that one-step-ahead prediction intervals of the conditional expectiles are more precise and tighter than those of expected shortfall, yielding much more reasonable capital requirements. They indeed seem to avoid both the permissiveness of the value-at-risk measure and conservatism of the expected shortfall. In addition, we explore the one-to-one mapping between quantiles and expectiles to assess how cryptocurrency risks evolve over time through the lens of the gain-loss ratio. Altogether, we believe our empirical analyses contribute to a better understanding of cryptocurrency markets, complementing previous studies in the literature. For instance, Zhang, Li, Xiong and Wang (2021) examine how downside risk affects the cross-section of cryptocurrency returns, whereas Makarov and Schoar (2019, 2020) investigate price discovery and arbitrage opportunities in cryptomarkets, respectively.

The remainder of this paper proceeds as follows. Section 2 discusses the main aspects of the class of ALS-based risk measures, paying special attention to the expectile. Section 2.2 entertains the conditional location-scale model we employ to estimate conditional expectiles, as well as the assumptions we impose to ensure consistency and asymptotic normality of our two-step estimator. Section 4 describes the fixed-design residual bootstrap algorithm we use to construct prediction intervals. Section 5 reports Monte Carlo simulations that validate our inference procedures. Section 6 assesses tail risk in cryptomarkets, whereas Section 7 offers concluding remarks.

2 Expectiles

Let the loss $L \in \mathbb{R}$ be a square integrable random variable with loss distribution function F_L. Newey and Powell (1987) define the τ-th expectile XP_τ^L as

$$\mathrm{XP}_\tau^L = \underset{\theta\in\mathbb{R}}{\arg\min} \int_{\mathbb{R}} |\tau - \mathbf{1}(\ell < \theta)|(\ell-\theta)^2 \mathrm{d}F_L(\ell) = \underset{\theta\in\mathbb{R}}{\arg\min}\, \mathbb{E}\left[\rho_\tau(L-\theta)\right] \tag{1}$$

where $\rho_\tau(\ell) = |\tau - \mathbf{1}(\ell \leq 0)|\ell^2$ is the expectile check function and $\mathbf{1}(A)$ denotes the indicator function that takes value one if A is true, zero otherwise. Apart from coinciding with the mean for $\tau = 1/2$, the expectile in (1) is well defined and unique for any $\tau \in (0,1)$ for any integrable random variable (Newey and Powell, 1987; Abdous and Remillard, 1995).

Replacing the absolute deviation in the quantile check function by a quadratic deviation facilitates optimization in a substantial manner. The quantile objective function of quantile regressions is not continuously differentiable at zero, whereby numerical implementation occasionally leads to quantile-crossing functions. In addition, (conditional) quantile estimators consider only the relative frequency of observations above or below their corresponding predictions (Daouia et al., 2018; Daouia et al., 2019a), with asymptotic distributions that strongly depend on the density function and smoothing parameters (Cheng and Parzen, 1997). In contrast, the objective function of the expectile regression is continuously differentiable almost everywhere, with a straightforward, efficient implementation by iterative reweighted least squares that employs all available information (Daouia et al., 2018). Asymptotic normality of the resulting ALS estimator requires only finite second moments (Holzmann and Klar, 2016).

The first-order condition of (1) also indicates that expectiles depend both on the probability and magnitude of tail realizations:

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{\mathbb{R}} |\mathbf{1}(\theta \geq y) - \tau| (\theta - y)^2 \,\mathrm{d}F_L(\ell) \bigg|_{\theta = \mathrm{XP}_\tau^L} \\
&= \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{-\infty}^{\theta} (1-\tau)(-y)^2 \,\mathrm{d}F_L(\ell) \bigg|_{\theta = \mathrm{XP}_\tau^L} + \frac{\mathrm{d}}{\mathrm{d}\theta} \int_{\boldsymbol{\theta}}^{\infty} \tau (\theta - y)^2 \,\mathrm{d}F_L(\ell) \bigg|_{\theta = \mathrm{XP}_\tau^L} \\
&= 2(1-\tau) \int_{-\infty}^{\mathrm{XP}_\tau^L} (\mathrm{XP}_\tau^L - y) \,\mathrm{d}F_L(\ell) + 2\tau \int_{\mathrm{XP}_\tau^L}^{\infty} (\mathrm{XP}_\tau^L - y) \,\mathrm{d}F_L(\ell) \\
&= 2(1-\tau)\, \mathbb{E}(L - \mathrm{XP}_\tau^L)^- - 2\tau\, \mathbb{E}(L - \mathrm{XP}_\tau^L)^+,
\end{aligned}
\tag{2}
$$

where $L^+ = \max\{L, 0\}$ and $L^- = \min\{-L, 0\}$. It then follows that

$$
\tau = \frac{\mathbb{E}(L - \mathrm{XP}_\tau^L)^-}{\mathbb{E}(L - \mathrm{XP}_\tau^L)} = \frac{\int_{-\infty}^{\mathrm{XP}_\tau^L} (\mathrm{XP}_\tau^L - y) \,\mathrm{d}F_L(\ell)}{\int_{\mathbb{R}} (\mathrm{XP}_\tau^L - y) \,\mathrm{d}F_L(\ell)}, \tag{3}
$$

implying that the expectile corresponds to the ratio of the average deviation of Y below XP_τ^L to the overall average deviation (Kuan et al., 2009). Accordingly, Remillard (2013), Bellini and Di Bernardino (2017) and Bellini, Klar and Müller (2018) link expectiles to gain-loss ratios. For instance, evaluating Keating and Shadwick's (2002) omega ratio at the expectile yields $\Omega_Y(\mathrm{XP}_\tau^L) = (1-\tau)/\tau$.

At any rate, expectiles relate to quantiles in several ways. First, they both characterize the entire distribution. The difference is that the expectile function $\tau \mapsto \mathrm{XP}_\tau$ is continuous and monotonically increasing on τ for any distribution (Philipps, 2022, Proposition 2). Second, there is a straightforward link between expectiles and quantiles:

$$\tau(\alpha) = \mathrm{XP}_{\tau}^{-1}(q_\alpha) = \frac{\int_{-\infty}^{q_\alpha} |\ell - q_\alpha| \,\mathrm{d}F_L(\ell)}{\int_{\mathbb{R}} |\ell - q_\alpha| \,\mathrm{d}F_L(\ell)}, \tag{4}$$

so that $\mathrm{XP}_{\tau(\alpha)} = q_\alpha$ (Yao and Tong, 1996; Philipps, 2022). Third, one can employ expectiles to estimate quantiles (Waltrup, Sobotka, Kneib and Kauermann, 2015; Daouia, Girard and Stupfler, 2019b) given that both are contained within the convex hull of the distribution's support. This means quantiles are a strict subset of the corresponding expectiles. Conveniently, this remains true in a regression context if the distribution belongs to the location-scale family (Yao and Tong, 1996). Lastly, Philipps (2022) also traces parallels between expectiles and L^p-quantiles (Chen, 1996) and m-quantiles (Breckling and Chambers, 1988).

And yet, the literature on expectiles remains small relative to the literature on quantiles (Waltrup et al., 2015). After a slow start in the 1990s (Breckling and Chambers, 1988; Efron, 1991; Jones, 1994; Abdous and Remillard, 1995; Yao and Tong, 1996), it is now gaining traction (Martin, 2014; Bellini and Di Bernardino, 2017; Krätschmer and Zähle, 2017; Daouia et al., 2018; Daouia, Girard and Stupfler, 2020; Girard et al., 2021). The primary reason is that the expectile is the only coherent risk measure that allows for straightforward backtesting due to elicitability (Bellini and Bignozzi, 2015; Ziegel, 2016; Daouia et al., 2023). And coherent risk measures are desirable because they satisfy the axioms of monotonicity, translation invariance, subadditivity and positive homogeneity. Monotonicity reflects that, if one position is less risky than another, the risk measure should assign a lower risk value. The translation invariance axiom states that adding a constant amount to the position's payout should increase riskiness by exactly that amount. The subadditivity axiom ensures there are diversification gains in pooling risks, whereas positive homogeneity dictates that multiplying the position by a positive constant should scale the risk measure by the same constant. See McNeil, Frey and Embrechts (2015) for more details.

Mathematical properties aside, we must always have in mind that, in practice, we have to assess risk measures empirically through estimates and/or forecasts. This is exactly the idea of elicitability, which requires the feasibility of backtesting risk measures through the minimization of expected scores (Gneiting, 2011; Bellini and Bignozzi, 2015; Ziegel, 2016). Recall that a score is a function $S : \mathbb{R} \times \mathbb{R} \to [0, \infty)$ such that $S(u, v) \geq 0$ with $S(u, v) = 0$ if and only if $u = v$;

$S(u,v)$ is increasing for $u > v$ and decreasing for $u < v$; and $S(u,v)$ is continuous in u. Gneiting (2011) shows that strictly consistent, homogeneous scoring functions for expectiles are of the form $S(r,L) = \mathbf{1}\{L > r\}(1-2\tau)\big(\phi(r) - \phi(L) - \phi'(r)(r-L)\big) - (1-\tau)\big(\phi(r) - \phi'(r)(r-L)\big)$, with ϕ denoting a strictly convex and integrable function with subgradient ϕ'.

The natural choice is the two-homogeneous scoring function given by $\phi(r) = r^2$ as in Newey and Powell (1987), though Nolde and Ziegel (2017) also entertains a zero-homogeneous alternative with $\phi(r) = \ln r$:

$$S_{\mathrm{XP},2}(r,L) = -\mathbf{1}\{L > r\}(1-2\tau)(L-r)^2 + (1-\tau)r(r-2L) \tag{5}$$

$$S_{\mathrm{XP},0}(r,L) = -\mathbf{1}\{L > r\}(1-2\tau)\left(\ln\frac{L}{r} + 1 - \frac{L}{r}\right) - (1-\tau)\left(\ln r - 1 + \frac{L}{r}\right), \text{ for } r > 0 \tag{6}$$

respectively. The corresponding scoring functions for the value-at-risk and expected-shortfall measures are respectively

$$S_{\mathrm{VaR},1}(r,L) = \big(1-\alpha-\mathbf{1}\{L > r\}\big)r + \mathbf{1}\{L > r\}L \tag{7}$$

$$S_{\mathrm{VaR},0}(r,L) = \big(1-\alpha-\mathbf{1}\{L > r\}\big)\ln r + \mathbf{1}\{L > r\}\ln L, \text{ for } r > 0 \tag{8}$$

$$S_{\mathrm{VaR,ES},1/2}(r_1,r_2,L) = \mathbf{1}\{x > r_1\}\,\frac{x-r_1}{2\sqrt{r_2}} + (1-v)\frac{r_1+r_2}{2\sqrt{r_2}} \tag{9}$$

$$S_{\mathrm{VaR,ES},0}(r_1,r_2,L) = \mathbf{1}\{x > r_1\}\,\frac{x-r_1}{r_2} + (1-v)\left(\frac{r_1}{r_2} - 1 + \ln r_2\right), \text{ for } r_2 > 0. \tag{10}$$

The scores for the value at risk are respectively 1- and 0-homogeneous functions, whereas those for both value at risk and expected shortfall are 1/2- and 0-homogeneous functions. See, for more details, Thomson (1979), Saerens (2000), Acerbi and Szekely (2014), Fissler and Ziegel (2016), and Nolde and Ziegel (2017).

2.1 Comparison between expectiles and quantile-based risk measures

In this section, we compare expectiles to traditional quantile-based risk measures. We start with the value-at-risk measure VaR_α at level α, which reads

$$\mathrm{VaR}_\alpha(L) = q_\alpha = F_L^{-1}(\alpha) = \inf\{\ell \in \mathbb{R} :\, F_L(\ell) \geq \alpha\}, \tag{11}$$

where q_α is the quantile function at level $\alpha \in (0,1)$. By definition, quantiles automatically satisfy the axioms of monotonicity, translation invariance and positive homogeneity. In addition, they are elicitable for strictly increasing distribution functions (Thomson, 1979; Saerens, 2000).

Unfortunately, the value-at-risk measure fails to account not only for the magnitude of losses beyond the level α, but also for diversification gains in view that it does not satisfy the axiom of subadditivity (Danielsson, Embrechts, Goodhart, Keating, Muennich, Renault, Shin et al., 2001).

Artzner et al. (1999), Acerbi, Nordio and Sirtori (2001), Rockafellar and Uryasev (2002) argue that the expected shortfall (ES), as defined by the expected loss given that it exceeds the value-at-risk, tackles both weaknesses of the VaR measure. For an integrable loss L with distribution F_L, the ES at level $\alpha \in (0, 1)$ is

$$\mathrm{ES}_\alpha(L) = \mathbb{E}\left[L \mid L > \mathrm{VaR}_\alpha(L)\right]. \tag{12}$$

In particular, ES is a severity-based risk measure that belongs to the class of spectral risk measures (Acerbi, 2002). However, it fails in two accounts. First, ES is not elicitable by itself, and hence backtesting is not straightforward (Gneiting, 2011). Second, it yields very imprecise estimates in finite samples for large values of α (Hull and White, 2014), especially for heavy-tailed loss distributions (Yamai and Yoshiba, 2002).

Both value at risk and expected shortfall depend heavily on the tail shape, though (Kuan et al., 2009; Daouia et al., 2019a). While the ES is too conservative because it is conditional only on the tail event, the VaR is too lenient for it does not account for the severity of the tail event. As such, quantile-based risk measures either underestimate or overestimate the risk exposure of a position.

Bellini and Di Bernardino (2017) interpret the expectile risk measure as the amount of capital that should be added to a position to yield a sufficiently high gain–loss ratio. As the only risk measure that meets the conditions for coherence, law invariance and elicitability (Ziegel, 2016), expectiles are extremely convenient for modeling, forecasting and backtesting purposes. In particular, the key advantage of expectiles over VaR and ES is the fact that ALS estimation uses more efficiently the available data, by exploiting both the severity and probability of tail events (Daouia et al., 2018). Both VaR and ES estimates indeed disregard the shape of the empirical distribution to the right of the corresponding quantile. In addition, the precision of the ES estimates depends heavily on α, sample size and tail thickness. This obviously poses a problem for establishing capital requirements, especially in the case of highly volatile and heavy-tailed asset returns (Yamai and Yoshiba, 2002; Hull and White, 2014).

For $\tau = \alpha > 1/2$, expectiles are always below both value-at-risk and expected-shortfall measures for every sample size, regardless of tail thickness. However, this ignores the equivalence result in (4). For instance, the first percentile $\mathrm{VaR}_{0.01}$ is close to the expectile at $\tau = 0.145\%$ for the Gaussian distribution (Bellini and Di Bernardino, 2017; Nolde and Ziegel, 2017). Chen (2018) shows that, if $\tau(\alpha)$ is such that $\mathrm{XP}_{\tau(\alpha)} = \mathrm{VaR}_\alpha$, then

$$\tau(\alpha) = \frac{\alpha\,(\mathrm{ES}_\alpha - \mathrm{VaR}_\alpha)}{\mathrm{VaR}_\alpha + 2\alpha\,(\mathrm{ES}_\alpha - \mathrm{VaR}_\alpha)} \tag{13}$$

provided that the loss distribution has mean zero. Straightforward manipulations then yield

$$\mathrm{ES}_\alpha = \mathrm{VaR}_\alpha \left(1 + \frac{\tau}{\alpha(1-2\tau)}\right) = \mathrm{VaR}_\alpha \left(1 + \frac{1}{\alpha(\Omega_\alpha - 1)}\right). \tag{14}$$

The first equality corresponds to the expectile-based expected shortfall proposed by Taylor (2008), whereas the second equality makes the connection with the omega ratio (Taylor, 2022). Solving for Ω_α gives way to

$$\Omega_\alpha = 1 + \frac{\mathrm{VaR}_\alpha}{\alpha(\mathrm{ES}_\alpha - \mathrm{VaR}_\alpha)}. \tag{15}$$

This allows us to compute the expected gain-loss ratio as a function of the quantile for some fixed level α. Besides, (15) makes clear that, for a fixed α, we should expect that the gain-loss ratio to increase (decrease) as the gap between expected shortfall and value-at-risk shrinks (enlarges, respectively). As such, we can infer by means of the omega ratio whether a given value-at-risk model is conservative (or permissive).

To sum up, expectiles offer a different perspective of the loss distribution relative to quantile-based risk measures. The literature unfortunately deals mostly with ALS estimation under random sampling, though. There are only a few studies that address the estimation of conditional expectiles in a time-series context (Taylor, 2008; Kuan et al., 2009; Bellini and Di Bernardino, 2017; Girard et al., 2021). In the next section, we discuss a two-step estimator of the conditional expectile following a strategy similar to Francq and Zakoïan (2015). In particular, due to translation invariance and positive homogeneity, it is straightforward to compute conditional expectiles in location-scale models. Accordingly, we first estimate the conditional mean and volatility of asset returns and then compute the unconditional expectiles of their standardized residuals.

2.2 Conditional expectiles

As expectiles are coherent risk measures, they are stable under affine transformations. This means that, in a location-scale model, the conditional expectile at a fixed level τ depends exclusively on the conditional mean and volatility, and of the τ-th expectile of the innovation. As it turns out, asset returns typically exhibit mean close to zero, but a highly persistent time-varying volatility with leverage effects (Bollerslev, Engle and Nelson, 1994) that leads to skewness and heavy tails in asset returns (Fama, 1965). Accordingly, we entertain a GARCH-type approach to model continuously compounded returns $\{y_t\}$:

$$y_{t+1} = \sigma_t \, \eta_{t+1}, \qquad \text{with} \;\; \sigma_t(\boldsymbol{\theta}_0) = \sigma(y_t, y_{t-1}, \ldots; \boldsymbol{\theta}_0) \;\; \text{for} \;\; t \in \mathbb{Z}, \tag{16}$$

where $\{\eta_t\}$ is a sequence of iid random variables with zero mean and unit variance, independent of past returns (i.e., $y_s \perp\!\!\!\perp \eta_t$ for $s < t$), and $\boldsymbol{\theta}_0 \in \mathbb{R}^m$ is a vector of model parametes. The sequence $\{y_t\}$ is a strictly stationary and ergodic solution to model (16).

This setting is very general, with (16) nesting the most popular GARCH-type models in the literature. For instance, Bollerslev's (1986) GARCH(1,1) model is such that $\sigma_t^2 = \omega_0 + \alpha_0 \, y_t^2 + \beta_0 \, \sigma_{t-1}^2$, where $\boldsymbol{\theta}_0 = (\omega_0, \alpha_0, \beta_0)'$.

Given our interest in expectiles, we henceforth assume that log-returns at time $t+1$ are integrable and follow a strictly stationary process adapted to a filtration $\{\mathcal{F}_t = \sigma(y_t, y_{t-1}, ...)\}$. The latter essentially boils down to restrictions in the parameter space. The conditional expectile at level τ then reads

$$\mathrm{XP}_\tau(y_{t+1}|\mathcal{F}_t) = -\sigma_t(\boldsymbol{\theta}_0) \, \mathrm{XP}_\tau^\eta \tag{17}$$

where $\mathrm{XP}_\tau^\eta < 0$ is the τ-th unconditional expectile of η_t for a small value of $\tau \in (0,1)$.

3 Estimation of conditional expectiles

To estimate the parameters in (17), we rely on a two-step approach as in Francq and Zakoïan (2015). We first estimate the conditional volatility parameters $\boldsymbol{\theta}_0$ using quasi-maximum likelihood (QML), and then compute the expectile XP_τ^η using standardized residuals $\hat{\eta}_{t+1} \equiv \eta_{t+1}(\widehat{\boldsymbol{\theta}}_n) = y_{t+1}/\sigma_t(\widehat{\boldsymbol{\theta}}_T)$, where $\hat{\boldsymbol{\theta}}_n$ is the QML estimator of the volatility parameters based on observations $(y_1, \ldots, y_n)$.

If the population parameters $\boldsymbol{\theta}_0$ were available, or equivalently, the true innovations were available, it would be straightforward to recover the expectile. Given a set of innovations $(\eta_1, ..., \eta_n)$ the empirical expectile solves

$$\widehat{\mathrm{XP}}_\tau^\eta = \arg\min_{\xi \in \mathbb{R}} \frac{1}{T} \sum_{t=1}^{T} \left|\tau - \mathbf{1}(\eta_t < \xi)\right| (\eta_t - \xi)^2. \tag{18}$$

This estimator has been shown to be strongly consistent (Holzmann and Klar, 2016; Krätschmer and Zähle, 2017) for any fixed level τ. Write the innovations $\eta_t = y_t/\sigma_t(\boldsymbol{\theta}_0)$ and let $\psi_{t,\tau}(\boldsymbol{\theta}, \xi) = (\sigma_t(\boldsymbol{\theta})^{-1} y_t - \xi)\left|\mathbf{1}(\sigma_t(\boldsymbol{\theta})^{-1} y_t < \xi) - \tau\right|$. The empirical expectile $\xi = \widehat{\mathrm{XP}}_\tau^\eta = \widehat{\mathrm{XP}}_{\tau,\boldsymbol{\theta}_0}$ is the unique zero of

$$Q_n(\xi, \boldsymbol{\theta}_0) = \frac{1}{n} \sum_{t=1}^{n} \psi_{t,\tau}(\boldsymbol{\theta}_0, \xi) = 0. \tag{19}$$

Unfortunately, the population parameters are not available and will have to be estimated from the data. Let $\{\tilde{y}_t\}$ be arbitrary values and define

$$\tilde{\sigma}_{t,s}(\boldsymbol{\theta}) = \sigma(y_{t-1}, \dots, y_{t-s+1}, \tilde{y}_{t-s}, \tilde{y}_{t-s-1}, \dots; \boldsymbol{\theta}). \tag{20}$$

We use $\tilde{\sigma}_{t,t}$ to approximate $\sigma_t(\boldsymbol{\theta})$. Denote by $h(\cdot)$ the instrumental density used in the QML estimation:

$$\tilde{G}_n(\boldsymbol{\theta}) = \frac{1}{n} \sum_{t=1}^{n} g(y_t, \tilde{\sigma}_{t,t}(\boldsymbol{\theta})), \quad g(y, \sigma) = \log \frac{1}{\sigma} h\left(\frac{y}{\sigma}\right). \tag{21}$$

The QML estimator solves $\hat{\boldsymbol{\theta}}_n = \arg\max_{\boldsymbol{\theta} \in \Theta} \tilde{G}_n(\boldsymbol{\theta})$ for some compact subspace $\Theta \subseteq \mathbb{R}^m$. Francq and Zakoïan (2015) study this estimation procedure and show that the QML estimator is strongly consistent with $\boldsymbol{\theta}_0$ and derive a central limit theorem for the estimator.

Let $\widehat{\mathrm{XP}}_{\tau,\boldsymbol{\theta}}$ be the zero of $Q_n(\widehat{\mathrm{XP}}_{\tau,\boldsymbol{\theta}}, \boldsymbol{\theta})$ for $\boldsymbol{\theta} \in \Theta$. We define our residual expectile estimator as $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}$ for each $\hat{\boldsymbol{\theta}}_n$. Next theorem shows that this estimator is consistent.

Theorem 1 (Consistency)**.** *Suppose Assumptions 1, 2, 3, 4 with* $s = 2$*, 5, 6 with* $s = 3$*, 9 with* $a = b = 3, c = 0$*, 10, and 11 (only first derivative), hold. Then, the residual expectile estimator* $\widehat{XP}_{\tau,\hat{\boldsymbol{\theta}}_n} \to XP_\tau^\eta$ *in probability as* $n \to \infty$.

We now show asymptotic normality of the two-step estimator of the conditional expectile. In particular, we examine exactly how replacing the true innovations with standardized residuals affects the asymptotic distribution of the sample expectile under assumptions similar to those in

Francq and Zakoïan (2015), Heinemann and Telg (2018), and Beutner et al. (2024). To do so, we first establish additional notation. Let $g_1(y;\sigma) = \frac{\partial}{\partial\sigma}\ln h(y/\sigma)/\sigma$, where h is the instrumental density in the QML estimation. Although typically Gaussian, the latter must in general satisfy conditions A3, A4 and A9 in Francq and Zakoïan (2015). In addition, let $\partial_{\boldsymbol{\theta}}\sigma_T(\boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}}\sigma_T(\boldsymbol{\theta})$, $g_2(y;\sigma) = \frac{\partial}{\partial\sigma} g_1(y;\sigma)$, and $\dot{\rho}_\tau(y) = \frac{1}{2}\frac{\partial}{\partial y}\rho_\tau(y)$.

Theorem 2 (Asymptotic Distribution)**.** *Let Assumptions1 to 11 in Appendix A hold for any fixed* $0 < \underline{\tau} < \tau < \overline{\tau} < 1$. *It follows that*

$$\sqrt{T}\left(\widehat{XP}_{\tau,\hat{\boldsymbol{\theta}}_n} - XP_\tau^\eta\right) \xrightarrow{d} \dot{\Psi}_{XP}(XP_\tau^\eta)^{-1}\mathbb{Z}_{XP} + XP_\tau^\eta\, J(\boldsymbol{\theta}_0)\mathbb{Z}_{\boldsymbol{\theta}},$$

where $\mathbb{Z}_{\boldsymbol{\theta}}$ *and* $\mathbb{Z}_{XP}$ *are jointly zero-mean Gaussian random variables,*

$$\dot{\Psi}_{XP}(XP_\tau^\eta) = \tau\big[1 - F_\eta(XP_\tau^\eta)\big] + (1-\tau)F_\eta(XP_\tau^\eta)$$

is a positive scalar and $J(\boldsymbol{\theta}_0) = \mathbb{E}\left[\frac{\partial_{\boldsymbol{\theta}}\sigma_n(\boldsymbol{\theta}_0)}{\sigma_n(\boldsymbol{\theta}_0)}\right]$ *is nonzero.*

The covariance matrix of $\mathbb{Z}_{\boldsymbol{\theta}}$ *is* $\Sigma_{\boldsymbol{\theta}} = 4\tau_h^2 I(\boldsymbol{\theta}_0)^{-1}$, *with*

$$I(\boldsymbol{\theta}_0) = \mathbb{E}\left(\frac{\partial_{\boldsymbol{\theta}}\sigma_n(\boldsymbol{\theta}_0)}{\sigma_n^2(\boldsymbol{\theta}_0)}\frac{\partial_{\boldsymbol{\theta}'}\sigma_n(\boldsymbol{\theta}_0)}{\sigma_n^2(\boldsymbol{\theta}_0)}\right) \quad \textit{and} \quad \tau_h^2 = \frac{\mathbb{E}[g_1^2(\eta_1,1)]}{\{\mathbb{E}[g_2(\eta_1,1)]\}^2},$$

whereas the variance of $\mathbb{Z}_{XP}$ *is* $\sigma^2_{XP_\tau} = \mathbb{E}\left[(\eta_1 - XP_\tau^\eta)^2(\mathbf{1}(\eta_1 < XP_\tau^\eta) - \tau)^2\right]$ *and their covariance is* $\Sigma_{XP,\boldsymbol{\theta}} = \mathbb{E}\left[\partial_{\boldsymbol{\theta}}\sigma_T(\boldsymbol{\theta}_0)\, g_1(\eta_1;1)\dot{\rho}_\tau(\eta_1 - XP_\tau^\eta)\right]$.

Our asymptotic theory essentially says that, under the correct specification of the conditional volatility process, the first-step estimation error affects only the asymptotic variance of sample expectile in the second step. More importantly, Theorem 2 offers a way to construct asymptotically-valid confidence intervals for the conditional expectile estimates $\widehat{\mathrm{XP}}_{\tau,t} \equiv \widehat{\mathrm{XP}_\tau}(y_{t+1}|\mathcal{F}_t)$. In the next section, we show how to alternatively construct asymptotically-valid confidence intervals using bootstrap methods.

4 Bootstrap

Bootstrapping is the most popular alternative to assess the finite-sample performance of conditional risk measures in the literature. In particular, there are both fixed- and recursive-design residual bootstrap procedures to obtain asymptotically-valid prediction intervals for value-at-risk (Christoffersen and Gonçalves, 2005; Beutner et al., 2024) and expected shortfall (Gao and

Song, 2008; Heinemann and Telg, 2018). The recursive design resamples standardized residuals and then recursively generates bootstrap samples of asset returns from the conditional volatility model using the parameter estimates. In contrast, the fixed design resamples standardized residuals, but then forms bootstrap samples using the original conditional volatility estimates. See Cavaliere et al. (2018) and references therein for in-depth discussions.

We resort to a fixed-design residual bootstrap procedure to construct asymptotically-valid confidence intervals for conditional expectile forecasts, in a similar vein to Heinemann and Telg (2018) and Beutner et al. (2024). The algorithm is as follows.

Algorithm 1 (Fixed-Design Residual Bootstrap)**.** *For each* $b \in \{1, \ldots, B\}$*,*

1. *Draw a random sample* $(\eta_1^{(b)}, \ldots, \eta_T^{(b)})$ *from the empirical distribution function of the standardized residuals* $\hat{\eta}_t = \eta_t(\hat{\boldsymbol{\theta}}_T)$*, and then compute* $y_{t+1}^{(b)} = \sigma_t(y_t, \ldots, y_1; \hat{\boldsymbol{\theta}}_T)\, \eta_{t+1}^{(b)}$ *for* $t = 2, \ldots, T$*.*

2. *Obtain the QML estimates* $\hat{\boldsymbol{\theta}}_T^{(b)}$ *of the volatility parameters using the bootstrap sample* $(y_1^{(b)}, \ldots, y_T^{(b)})$*, and then compute for* $t = 1, \ldots, T$ *the corresponding bootstrap standardized residuals* $\eta_{t+1}(\hat{\boldsymbol{\theta}}_T^{(b)}) = y_{t+1}^{(b)} / \sigma_t(y_t^{(b)}, \ldots, y_1^{(b)}; \hat{\boldsymbol{\theta}}_T^{(b)})$*.*

3. *Estimate the expectile* $\widehat{XP_\tau}^{(b)}$ *of the bootstrap standardized residuals* $\eta_t(\hat{\boldsymbol{\theta}}_T^{(b)})$*, and then combine with the conditional volatility estimates to obtain* $\widehat{XP}_{\tau,t}^{(b)}$*.*

We then denote by $\hat{\boldsymbol{\theta}}_T^B = \frac{1}{B}\sum_{b=1}^B \hat{\boldsymbol{\theta}}_T^{(b)}$ *and* $\widehat{XP}_{\tau,t}^B = \frac{1}{B}\sum_{b=1}^B \widehat{XP}_{\tau,t}^{(b)}$ *the bootstrap estimators of the volatility parameters and of the conditional expectile, respectively.*

We focus mainly on the fixed design for two reasons. First, it is much faster than the recursive design. Second, the asymptotic validity of the residual bootstrap with a fixed design does not depend as much on the volatility specification. Cavaliere et al. (2018) remark that the recursive design must consider the dependence structure that the bootstrap procedure engenders. As such, it requires a stronger set of conditions to ensure bootstrap consistency for conditional risk measures. Given the fixed-design residual bootstrap performs better than Beutner et al.'s (2024) recursive-design bootstrap in our Monte Carlo simulations, we omit the latter results to save space.[1]

[1] They are obviously available from the authors upon request.

The next result establishes the first-order equivalence of the QML and fixed-design bootstrap estimators of the volatility parameters.

Lemma 1 (Beutner et al. (2024)). *Let Assumptions 3 to **??**, **??**(i), **??**(iii), **??**, 7, 9 and **??** hold with $a = \pm 12$, $b = 12$ and $c = 6$. It then follows that*

$$\sqrt{T}(\hat{\boldsymbol{\theta}}_T^B - \hat{\boldsymbol{\theta}}_T) \xrightarrow{d^*} \mathcal{N}\left(0, \frac{\kappa - 1}{4}\, J^{-1}\right),$$

where $J = \mathbb{E}\left(D_t D_t'\right)$ and $D_t = D_t(\boldsymbol{\theta}_0) = \sigma_t(\boldsymbol{\theta}_0)^{-1}\partial\sigma_t(\boldsymbol{\theta}_0)/\partial\boldsymbol{\theta}_0$.

To gauge estimation/prediction risk, we employ a two-step procedure to obtain bootstrap confidence intervals at approximately $100(1-\alpha)\%$ level. The first step is to generate bootstrap samples of the conditional expectile estimates using Algorithm 1. In the second stage, we construct bootstrap-based prediction intervals using three alternative methods: namely, equal-percentile (EP), reverse tail (RT), and symmetric (SY). Let $G_{B,T}^{-1}(x)$ and $H_{B,T}^{-1}(x)$ respectively denote the generalized inverses of $G_{B,T}(x) = \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\left(\sqrt{T}\big(\widehat{\mathrm{XP}}_{\tau,t}^{(b)} - \widehat{\mathrm{XP}}_{\tau,t}\big) \leq x\right)$ and $H_{B,T}(x) = \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\left(\sqrt{T}\big|\widehat{\mathrm{XP}}_{\tau,t}^{(b)} - \widehat{\mathrm{XP}}_{\tau,t}\big| \leq x\right)$. We then define the EP and RT intervals respectively as

$$\left[\widehat{\mathrm{XP}}_{\tau,t} - \frac{1}{\sqrt{T}}\, G_{B,T}^{-1}(1-\alpha/2),\quad \widehat{\mathrm{XP}}_{\tau,t} - \frac{1}{\sqrt{T}}\, G_{B,T}^{-1}(\alpha/2)\right] \tag{22}$$

and

$$\left[\widehat{\mathrm{XP}}_{\tau,t} + \frac{1}{\sqrt{T}}\, G_{B,T}^{-1}(\alpha/2),\quad \widehat{\mathrm{XP}}_{\tau,t} + \frac{1}{\sqrt{T}}\, G_{B,T}^{-1}(1-\alpha/2)\right], \tag{23}$$

whereas the symmetric interval reads

$$\left[\widehat{\mathrm{XP}}_{\tau,t} - \frac{1}{\sqrt{T}} H_{B,T}^{-1}(1-\alpha),\quad \widehat{\mathrm{XP}}_{\tau,t} + \frac{1}{\sqrt{T}}\, H_{B,T}^{-1}(1-\alpha)\right]. \tag{24}$$

The EP and SY intervals essentially replace the true quantile function in the corresponding unfeasible confidence intervals with the empirical quantile function. In turn, the RT interval flips tails around, yielding lower and upper bounds that coincide with the $\alpha/2$- and $(1-\alpha/2)$-quantiles of the empirical distribution of $\widehat{\mathrm{XP}}_{\tau,t}^{(b)}$. See Falk and Kaufmann (1991) for more motivation on RT intervals, and Beutner et al. (2024) for more details in general.

5 Monte Carlo study

We next run a Monte Carlo study to assess not only the finite-sample distribution of the two-step expectile estimator, as well as the performance of the fixed-design bootstrap procedure. In particular, we compute the bias and root mean squared error of the two-step estimator, as well as the coverage of the bootstrap confidence intervals. Our setup attempts to reproduce the main stylized facts we observe in asset returns. To reflect the usual challenges of empirical analyses, we also address the impact of misspecification.

We start with well-specified conditional volatility models that satisfy Assumptions 3 to **??**, as in Christoffersen and Gonçalves (2005). We simulate daily returns from a GARCH(1,1) process with Student-t errors: namely, $\sqrt{\nu/(\nu-2)}\,\eta_t \sim t_\nu(0,1)$ to ensure unit variance. To have a better grasp of the tail influence in the results, we entertain a heavy-tailed Student's t_ν-distribution with $\nu = 8$ degrees of freedom, as well as a virtually Gaussian alternative with $\nu = 500$. To ensure positivity and strict stationarity, we fix the unconditional volatility at 20% per year by setting $\omega_0 = 20^2/252 \times (1-\alpha_0-\beta_0)$, with $\alpha_0+\beta_0 < 1$. In particular, we contemplate $\alpha_0 = 0.10$ and $\beta_0 = 0.80$ in the benchmark case, as well as a highly persistent scenario with $\beta_0 = 0.89$. For each data generating process (DGP), we draw $S = 10,000$ sample paths of size $T \in \{500, 1000, 2500, 5000\}$, after burning the first 1,000 realizations.[2]

We first examine the finite-sample performance of the two-step estimator of the unconditional expectile, fixing τ at 5%. In particular, we compare via box plots the distributions of the difference between the sample and true expectiles using either true innovations or standardized GARCH residuals. For each replication, we also obtain expectile estimates for one fixed-design bootstrap sample, and then look at their distributions relative to the original QML estimates over the $S = 10,000$ replications. Figure 3 displays box plots of these distributions for the different scenarios of tail thickness, persistence and sample sizes we consider.

As expected, the distributions become less spread out as sample size increases. It is also apparent that estimating the GARCH process affects the distribution of the sample expectiles. The distributions based on standardized residuals are slightly less disperse than the corresponding distributions using true innovations, apart from exhibiting a very small bias in smaller samples ($T = 500$). Bootstrapping performs very well, resulting in distributions that are very close to the sample distributions based on true innovations. This indicates that resampling techniques

[2] We employ the `ugarchpath` function of the `rugarch` package in `R` (Ghalanos, 2018), whereas we run the numerical optimization of the quasi-likelihood function using the `nmlib` function. To reduce running time, we resort to parallel computation by means of the `mclapply` function of the `parallel` package.

can indeed mitigate the impact of the GARCH estimation in finite samples. Finally, persistence in volatility does not appear to matter much, apart from slightly thicker tails in the sampling distributions.

We next turn our attention to the conditional expectiles. In particular, for each replication, we proceed as follows. We first estimate the conditional expectiles at time $T+1$ given the entire sample, using the conditional volatility $\sigma_t(\boldsymbol{\theta}_0)$. This is feasible only because the conditional variance in a GARCH-type process depends exclusively on past returns (and the true parameter vector $\boldsymbol{\theta}_0$). We then compute the one-step-ahead volatility forecast $\sigma_t(\hat{\boldsymbol{\theta}}_T)$ based on the QML estimates, in order to compute the difference between the forecast and true values of the conditional expectiles: namely, $\widehat{\mathrm{XP}}_{\tau,t} - \mathrm{XP}_{\tau,t} = \sigma_t(\boldsymbol{\theta}_0)\,\mathrm{XP}_\tau^\eta - \sigma_t(\hat{\boldsymbol{\theta}}_T)\,\widehat{\mathrm{XP}_\tau}^{\hat{\eta}}$. Figure 4 plots the distributions of the forecast errors for the conditional expectiles at $\tau = 5\%$. As before, we consider the GARCH(1,1) models, with different levels of persistence in volatility and t_ν-innovations with $\nu \in \{8, 500\}$ degrees of freedom.

There is virtually no bias in the conditional estimates. In turn, the amount of persistence now seems to affect in a considerable manner the distributions of the two-step QML estimator of the conditional expectiles in smaller samples. For the benchmark GARCH(1,1) processes with $\alpha_0+\beta_0 = 0.90$, the distributions of the two-step forecast errors are reasonably symmetric around zero for every sample size. However, as we increase persistence in volatility to $\alpha_0+\beta_0 = 0.99$, the distributions become right skewed and leptokurtic for the smaller sample sizes ($T = 500, 1000$). As before, bootstrapping residuals leads to more symmetric finite-sample distributions.

We complement the analysis by assessing the relative bias and root mean squared error (RMSE) of the unconditional and conditional expectiles for a wider array of sample sizes: namely, $T \in \{250, 500, 750, \ldots, 4750, 5000\}$. Figure 5 displays bias in solid lines and RMSE in dotted lines, whereas colors identify two-step estimates/forecasts for the original and bootstrap samples (red and blue, respectively). There is very little relative bias even for $T = 250$, vanishing to zero very rapidly as sample size increases. Likewise, RMSE also decreases with the sample size, with a sharp reduction up to sample sizes of about 1,500 observations. By construction, the RMSE of the two-step estimators of the unconditional risk measures are always smaller than those based either on true innovations or on bootstrap residuals. In line with Figure 3, bootstrapping residuals with a fixed design yields very similar bias and RMSE to sample expectiles based on true innovations. The pattern of higher RMSE for the two-step forecasts based on bootstrap samples persists for conditional expectiles.

5.1 Bootstrap-based prediction intervals

The performance of the fixed-design residual bootstrap is so far encouraging. However, it remains to assess whether bootstrapping indeed offers reliable confidence intervals for the conditional expectiles. To do so, we compute the coverage rates of one-day-ahead prediction intervals with a confidence level of 90% based on $B = 999$ bootstrap samples. Apart from average coverage rates, we also compute average lengths of the bootstrap-based prediction intervals and violation rates at the lower and upper bounds across $S = 10,000$ simulations. The data generating processes are the same as before: namely, returns follow a GARCH(1,1) process with $\alpha_0+\beta_0 \in \{0.90, 0.99\}$ and t_ν-innovations with $\nu \in \{8, 500\}$. For each replication, we generate samples of $T \in \{500, 1000\}$ after discarding the first 1,000 observations.

We obtain bootstrap-based prediction intervals using the equal-tail percentile (EP), reverse tail (RT) and symmetric (SY) methods, as we describe in (22) to (24). For the sake of comparison, we do so not only for the expectile, but also for the traditional quantile-based measures (i.e., VaR and ES). Table 1 reports the performance of the alternative bootstrap-based prediction intervals for the GARCH(1,1) process with (approximately) Gaussian innovations. Coverage rates are very close to nominal levels for every risk measure, interval construction method, and sample size, especially for the benchmark level of persistence in volatility. It is interesting to observe that bootstrap-based intervals quite consistently underestimate upper tails, though RT and SY intervals apparently overestimate lower tails. As a result, they have better coverage rates than the EP interval, in line with Beutner et al.'s (2024) and Heinemann and Telg's (2018) findings. In addition, the prediction intervals of the expectiles are much narrower than those of the quantile-based measures. Finally, sample size affects more the average length of the prediction intervals than their coverage rates.

Table 2 reports similar statistics for the GARCH(1,1) process with t_8-innovations. Results are very similar, except perhaps to the prediction intervals of the expected shortfall. This is not surprising given that ES is less robust to tail events, which are now more likely to occur due to the heavier tail of the t-distribution. As such, despite their larger lengths on average, their prediction intervals greatly underestimate the lower tails, compromising their coverage rates in a substantial manner. Finally, although the coverage rates of the expectiles are similar to those of the value-at-risk, the average lengths of their prediction intervals are much smaller.

We next redo the analysis for $\tau = 1\%$ to assess how the bootstrap-based prediction intervals behave for more extreme risk measures. Table 3 reports the performances of the bootstrap-based prediction intervals at the 90% confidence level for the GARCH(1,1) processes with approximately Gaussian innovations ($\nu = 500$). As expected, coverage rates greatly deteriorate as we move τ from 5% to 1%. This is especially the case for the expected shortfall given it depends heavily on tail expectations, though it also affects to some extent the VaR coverage rates. The best coverage rates are for the RT and SY bootstrap-based prediction intervals of the conditional expectiles, mainly because they overestimate the lower tail ensuring fewer-than-5% exceptions. Finally, as for the length of the prediction intervals, we observe exactly the same pattern as for $\tau = 5\%$, that is to say, larger intervals for the expectiles than for the quantile-based measures (VaR and ES).

Table 4 documents the corresponding figures for the GARCH(1,1) processes with Student-t innovations ($\nu = 8$). We find very similar patterns, apart perhaps from the utter disruption of the bootstrap-based prediction intervals of the expected shortfall. Due to the thicker tails, coverage rates drop to a range between 37% and 46% for samples of 500 observations. The situation is even worse for samples of 1,000 observations, with coverage rates between 34% and 40%, because it is more likely to observe at least one extreme tail event in larger samples. These discouraging results are in line with Gao and Song (2008) and Heinemann and Telg (2018), who argue that only large sample sizes would produce acceptable coverage rates for the expected shortfall.

To sum up, our simulations reveal that the fixed-design bootstrap distribution approximates well the finite-sample distribution of the conditional expectiles, especially in larger samples. Our backtesting exercise also indicates that the coverage ratios of the bootstrap-based prediction intervals are reasonably close to nominal levels, ensuring a proper assessment of the estimation risk in conditional expectiles. In the next subsection, we revisit the performance of the bootstrap-based prediction intervals under different degrees of misspecification of the volatility process.

5.2 What happens under misspecification?

Recall that Assumption **??** in Appendix A requires the correct specification of the conditional volatility model. Otherwise, under misspecification, the QML estimator converges to a pseudo-true parameter vector that might lead to markedly different volatility forecasts and standardized residuals (Beutner et al., 2024). In what follows, we study the impact of misspecification issues in the estimation of conditional expectiles and their bootstrap-based prediction intervals.

For that purpose, we consider the family of asymmetric power autoregressive conditional heteroskedastic (APARCH) models, put forth by Ding et al. (1993): $y_{t+1} = \sigma_t\, \eta_{t+1}$, with

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p} \alpha_i (|y_{t-i+1}| - \gamma_i\, y_{t-i+1})^\delta + \sum_{j=1}^{q} \beta_j\, \sigma_{t-j}^\delta, \tag{25}$$

where $\alpha_i \geq 0$, $-1 < \gamma_i < 1$, $\beta_j \geq 0$, $\omega \geq 0$, and $\delta > 0$ to ensure the positiveness of the volatility process.[3] The APARCH(p,q) family nests several specifications in the literature, including the ARCH (Engle, 1982), GARCH (Bollerslev, 1986), GJR-GARCH (Glosten et al., 1993), and TGARCH (Zakoian, 1994) models.

In particular, we set the intercept to $\omega = \big(1 - \sum_{i=1}^{p} \alpha_i\, \kappa_i - \sum_{j=1}^{q} \beta_j\big)(20/\sqrt{252})^\delta$, where $\kappa_i = \mathbb{E}(|\eta| - \gamma_i\, \eta)$ with $\eta \sim t_{500}$. As for other parameters, we consider the following conditional volatility specifications: ARCH(2) with $\delta = 2$, $p = 2$, $q = 0$, $\gamma_1 = \gamma_2 = 0$, $\alpha_1 \in \{0.80, 0.89\}$, and $\alpha_2 = 0.10$; GARCH(2,2) with $\delta = 2$, $p = q = 2$, $\alpha_1 = 0.07$, $\alpha_2 = 0.03$, $\beta_1 \in \{0.70, 0.79\}$, and $\beta_2 = 0.10$; GJR-GARCH$(1,1)$ with $\delta = 2$, $p = q = 1$, $\alpha_1 = 0.10/\kappa_1 \approx 0.09$, $\gamma_1 = 1/3$, and $\beta_1 \in \{0.80, 0.89\}$; and TGARCH(1,1) with $\delta = 1$, $p = q = 1$, $\alpha_1 = 0.10/\kappa_1 \approx 0.125$, $\gamma_1 = 0.60$, and $\beta_1 \in \{0.80, 0.89\}$. The GJR-GARCH and TGARCH specifications are such that the impact of negative returns is fourfold that of positive returns, as in Francq and Zakoïan (2015).

We simulate daily returns that follow the above volatility processes with Gaussian innovations (i.e., t_{500}). We draw $S = 10,000$ sample paths of $T = 1000$ observations after burning the first 1,000 realizations. For each volatility specification, we estimate a GARCH(1,1) model by QML and then compute the expectiles of the standardized residuals with $\tau = 5\%$. Figure 6 exhibits

[3] We avoid exponential/log GARCH models as in Nelson (1991) for two reasons. First, they require different conditions for the consistency and asymptotic normality of the QML estimator (Allen, Chan, McAleer and Peiris, 2008; Wintenberger, 2013). Second, they fail to satisfy Assumption 8 (Beutner et al., 2024).

the box plots of the estimation errors of the unconditional expectiles for the different APARCH specifications. As before, apart from the estimates based on the standardized residuals of the GARCH(1,1) model in the original and bootstrap samples (in red and blue, respectively), we also consider the misspecification-free estimates based on true innovations (in green).

The GARCH(1,1) process we estimate filters better the GARCH(2,2) than the ARCH(2) volatility process, especially if persistence is very high. This is surprising because it should arguably capture well enough the persistence in volatility of any ARCH process, and not necessarily the autocorrelation structure of higher-order GARCH processes. The box plots nonetheless reveal that the sample expectiles of the GARCH(1,1) standardized residuals are negatively biased for the ARCH(2) data generating process. In contrast, we observe little estimation bias when we approximate GARCH(2,2) processes by a GARCH(1,1) model. Perhaps even more surprising is that the absence of bias when we miss out the leverage effects that characterize GJR-GARCH processes. The distributions of the sample expectiles using standardized residuals are very similar to those using true innovations. Unfortunately, the same is not true for the TGARCH data generating process. Employing GARCH(1,1) standardized residuals yield a relatively large negative bias in the estimation of the unconditional expectile. Finally, bootstrapping residuals using a fixed design alleviates the estimation error in the standardized residuals. Their distributions are indeed much closer to those based on true innovations. However, we still find evidence of bias in the ARCH(2) and TGARCH(1,1) cases, even if smaller in magnitude.

Figure 7 display the corresponding box plots for the conditional expectiles. Misspecification now affects not only the estimation of the unconditional expectiles through the standardized residuals, but also the volatility forecasts. As before, the distributions of the conditional expectiles in the bootstrap samples feature less dispersion than the corresponding distributions using standardized residuals. They are actually closer to the distributions of the conditional expectiles using true innovations, corroborating the evidence that bootstrapping helps control the first-step estimation error.

We next examine how well the fixed-design bootstrap performs under misspecification by looking at the coverage rates of the one-day-ahead prediction intervals. As before, we simulate $S = 10,000$ sample paths of the ARCH(2), GARCH(2,2), GJR-GARCH(1,1) and TGARCH(1,1) processes, and then employ $B = 999$ bootstrap replications based on GARCH(1,1) standardized residuals to construct prediction intervals at the 90% confidence level for each simulation. Tables 5 to 7 document the performance of the 90% bootstrap-based prediction intervals for the conditional risk measures at the $\tau = 0.05$ level in the case of Gaussian innovations.

The results indicate that the type of misspecification matters substantially. Missing the order of the GARCH process is not very damaging. Bootstrapping standardized residuals from a GARCH(1,1) model indeed works very well even if the true generating process is a GARCH(2,2) model, regardless of the level of persistence. The same is not true for the ARCH(2) data generating process, though. Although misspecification does not seem to hurt much the empirical coverage of the bootstrap-based prediction intervals for the benchmark case of $\alpha_1 + \alpha_2 = 0.90$, the resulting coverage for the highly persistent ARCH(2) process with $\alpha_1 + \alpha_2 = 0.99$ is poor. Lastly, bootstrap-based prediction intervals utterly fail if we compute standardized residuals using a simple GARCH(1,1) specification when leverage effects are at play. This is especially true for the TGARCH specifications, which focus on the conditional volatility (rather than variance), apart from featuring leverage effects (i.e., $\gamma_1 \neq 0$).

To sum up, it is apparent that the correct specification of the conditional volatility process is crucial for bootstrap-based prediction intervals, even if employing a fixed design. Accordingly, it is key to follow the best modeling practices by conducting exhaustive specification tests. For instance, the presence of skewness in the standardized residuals could well reflect unaccounted leverage effects in the volatility process. Alternatively, one could attempt to correct bias using double bootstrap procedures, as in Cavaliere, Gonçalves, Nielsen and Zanelli (2023).

6 Tail risk in cryptomarkets

We collect data from Yahoo Finance on the daily prices of the cryptocurrencies with the highest market values at the end of 2023: namely, Bitcoin (BTC), Ethereum (ETH), Binance Coin (BNB), and Cardano (ADA). For the sake of comparison, we also gather the daily prices of the S&P 500 equity index (SPX) and Euro-Dollar exchange rate (EUR). The sample ranges from January 2016 to December 2023 for SPX, EUR and BTC, whereas it runs from November 2017 to December 2023 for ADA, BNB and ETH. Figure 8 documents their evolution over time. Apart from the large drawdown during the Covid-19 outbreak, the SPX index grows almost linearly in time, as compared to the long swings of the EUR exchange rate. In contrast, cryptocurrencies gain value very rapidly from the last quarter of 2020 to the last quarter of 2021, only to experience a turbulent period in the next 12 months. For instance, Bitcoin's attains its maximum value of \$67,566.83 on November 8, 2021 but then drops to \$15,787.28 by November 21, 2022.

To assess downside risk, we compute the historical drawdowns of each time series. We define drawdown at time t as the difference between the maximum cumulative return up to time t and the cumulative return at time t. The idea is to measure the percentage drop from the all-time high. Figure 9 compares the historical drawdowns of the S&P 500 index and Euro-Dollar exchange rate with those of the cryptocurrencies, as from 2018. The box plots reveal there is a striking discrepancy between the historical drawdowns of traditional and crypto assets. In particular, EUR and SPX exhibit very modest drawdowns in comparison with those in the cryptocurrency market.

To deal with the nonstationarity of the S&P 500 equity index and exchange rates we compute daily continuously compounded returns $y_T = 100 \ln(P_T/P_{t-1})$, where P_T is the asset price at time t. Table 10 reports their main summary statistics. The sample sizes for the cryptocurrency returns are relatively larger than for traditional assets because they trade every day of the week, with no bank holidays.

The most striking feature is definitely the range into which daily cryptocurrency returns dwell. Their minimum values reflect drops in value of about 50%, with highest returns varying from 22.5% to over 85% in just one day! Such a variation leads to very high levels of daily volatility, from 3.75% to 6.31%. In line with the risk-return tradeoff, typical daily returns are also very high, with average values between 9% and 23%. Median returns are smaller, though: 2% for Cardano, 8% for Ethereum, 10% for Binance Coin, and 15% for Bitcoin. In comparison, changes in the Euro-Dollar exchange rate range from -2.81% to 1.82%, having zero mean and a daily volatility of 0.47%. In turn, the daily S&P 500 index returns vary from -12.77% to 8.97%, with typical values of 4% in average (7% in median) and annual volatility of 18.7% per year. While the skewness we observe especially for ADA, BTC, ETH and SPX might indicate leverage effects, the excess kurtosis in every asset return is consistent not only with heavy tails, but also with conditional heteroskedasticity.

Preliminary analyses show that MA(2) processes suffice to deal with the low persistence in asset returns. Given that missing out leverage effects might completely disrupt estimation and prediction intervals, we adopt a GJR-GARCH(1,1) specification for the conditional variance, even though we actually expect to observe a significant asymmetric response to negative returns only for the S&P 500 index. In what follows, we estimate the MA(2)-GJR-GARCH(1,1) models using rolling windows of 1,000 time-series observations. We then compute the unconditional expectiles of the standardized residuals for each estimation window in order to come up with one-day-ahead forecasts of the conditional expectiles, and their bootstrap-based prediction intervals. Given the encouraging results in the Monte Carlo study, we rely on Beutner et al.'s (2024) fixed-design residual bootstrap procedure to build 90% confidence intervals using the reversed tail method.

For comparison purposes, we compute one-step-ahead predictions of the value-at-risk and expected shortfall, as well. Interpretation changes as we move from one risk measure to another, though. Although every risk measure should dictate a capital requirement to serve as a buffer against unexpected losses, we cannot directly compare their values for a given level τ. Accordingly, we assess performance using statistical tests. Kupiec (1995) assesses unconditional coverage of value-at-risk models by looking at the expected and realized number of failures. In addition, Christoffersen (1998) and Christoffersen and Pelletier (2004) check also whether value-at-risk exceptions are independent over time using Markov and duration tests, respectively. In particular, the former tests whether the probability of a violation tomorrow depends on whether there is a violation today, whereas the latter assesses whether the absence of memory in the VaR violations. As for conditional coverage tests, they examine both unconditional coverage and independence by combining Kupiec's (1995) and Christoffersen's (1998) tests.

Backtesting tail expectations is not so straightforward. McNeil and Frey (2000) propose a statistical test based on the difference between realized returns and expected-shortfall estimates conditional on VaR violations. Under the null hypothesis of correct specification of the conditional mean and variance, standardized residuals should form a sequence of random variables with zero mean. Accordingly, they employ bootstrap methods to test the null hypothesis of mean zero in a nonparametric manner. For backtesting purposes, we adapt such residual test for conditional expectile models, as well.

Table 11 reports the backtesting analysis for each risk measure at $\tau = 0.01$ for a 90% confidence level. Value-at-risk performs very well in that it predicts a number of exceptions very close to the realized exceptions. As such, it is not surprising that the coverage tests cannot reject the specification of the model. Likewise, the bootstrap-based residual tests suggest the congruence of the expected shortfall, expectiles, and extremiles.

Figure 10 portrays daily returns from January 2, 2020 to December 30, 2023 as well as the one-step-ahead forecasts of the conditional risk measures at level $\tau = 0.01$ with their 90% bootstrap-based prediction intervals. It is apparent from the scale of the vertical axes that BTC is much riskier than SPX, with larger prediction intervals. The uncertainty around the BTC expected-shortfall estimates is also striking. The average length of the 90% bootstrap-based prediction intervals is of 6.46% for the expected shortfall, whereas they are just below 4.5% for expectiles and 3.22% for the value-at-risk forecasts. A similar pattern arises for SPX, in that the average length of the 90% bootstrap-based prediction intervals of the expected shortfall is relatively larger: 1.25% against 1.08% for value-at-risk and 0.974% for expectile.

Figure 11 reveals box plots for the difference between one-day-ahead forecasts of the conditional risk measures at $\tau = 0.01$ and realized returns. From a risk management standpoint, this difference indicates whether the predicted capital requirement is excessive or insufficient to absorb realized losses. For a fixed τ, we can rank the capital requirements implied by the different conditional risk measures as $\widehat{\mathrm{XP}}_\tau < \widehat{\mathrm{VaR}}_\tau < \widehat{\mathrm{ES}}_\tau$. The ES-based capital requirements are indeed very conservative, with first quartiles in the same order of magnitude than the third quartiles of the other risk measures. Conversely, expectiles assign the lowest capital requirements for every asset, but at the price of a large number of exceptions. This explains why we observe so many exceptions for the expectile forecasts, while very few for the expected shortfall in Table 11.

We so far consider a fixed τ across the different conditional risk measures, assuming that there is a mapping between their point predictions and capital requirements. Bellini and Di Bernardino (2017) argue however that we should choose τ in each risk measure for a given capital requirement. For instance, there is a capital equivalence between $\text{VaR}_{0.01}$ and $\text{ES}_{0.025}$ for regulatory purposes (Bank for International Settlements, 2013; Bank for International Settlements, 2014). Bellini and Di Bernardino (2017) and Nolde and Ziegel (2017) show that the expectile at $\tau = 0.145\%$ yields a similar capital requirement under normality. There is no need to restrict attention to the Gaussian case, though. Daouia et al. (2018) and Schmidt, Katzfuss and Gneiting (2021) investigate the connection between quantiles and expectiles in general, even if they do not discuss explicitly equivalence in terms of regulatory capital requirements. In fact, the one-to-one mapping in (4) ensures we can readily compute the expectile that corresponds to any value-at-risk measure by defining $\tau(\alpha)$ such that $\text{XT}_{\tau(\alpha)} = \text{VaR}_\alpha$ (Chen, 2018).

Some remarks are in order. First, once we establish capital requirements based on value-at-risk estimates, we can backtest them using traditional procedures (Kupiec, 1995; Christoffersen, 1998; Christoffersen and Pelletier, 2004). Second, for a fixed VaR_α, we can evaluate portfolio performance by means of the omega ratio in view that the conditional gain-loss ratio rests on the inverse of the expectile function. Third, we can establish a fixed gain-loss ratio using conditional expectiles, and then backtest the model using Gneiting's (2011) elicitable theory, in order to adjust VaR and ES capital requirements in a dynamical manner. Fourth, given that (13) cast expectiles as a function of value-at-risk and expected shortfall, we can jointly compare the forecast performances of both VaR_α and ES_α using the elicitable theory put forth by Fissler and Ziegel (2016). Altogether, value-at-risk and expectiles complement each other by offering a very complete tool for risk management purposes.

To begin with, we examine the out-of-sample performance of the conditional value-at-risk at the level $\alpha = 0.01$ by looking at the unconditional and conditional coverage tests (Kupiec, 1995; Christoffersen, 1998), as well as Christoffersen and Pelletier's (2004) duration test. Moreover, we also implement the scoring functions (7) and (8) to test our two-step method based on the empirical distribution of the QML standardized residuals against the most natural parametric alternatives based on Gaussian and Student-t errors. In the latter, we estimate the conditional mean and variance parameters by maximum likelihood, as well as the degrees of freedom in the Student-t case, and then employ the corresponding parametric quantiles of the standardized residuals. In particular, maximum likelihood yields Student-t distributions with 3 and 6 degrees of freedom for BTC and SPX, respectively.

Table 12 reports the backtest results of the VaR estimates at level $\alpha = 0.01$ for a confidence level of 90%. In general, we observe that the two-step estimator that employs the empirical quantiles of the QML standardized residuals outclasses both parametric alternatives, yielding more accurate numbers of violations, higher p-values in the coverage and duration tests. Besides, their average scores corroborate the evidence that their one-step-ahead predictions outperform both parametric alternatives, especially for SPX.

Figure 12 displays the one-step-ahead forecasts of ES_α, VaR_α and $\mathrm{XP}_{\tau(\alpha)}$. Note that we must change $\tau(\alpha)$ over time to ensure that the one-step-ahead forecasts of the expectile at $\tau(\alpha)$ coincide with those of the value-at-risk at $\alpha = 0.01$, as in (4). This occurs because the empirical quantile of the loss distribution varies over time due to the rolling estimation window. As for the distance between expected-shortfall and value-at-risk forecasts, it is much greater for BTC than for SPX, due to the heavier tails of the former returns.

Figure 12 also depicts the omega ratios implied by the one-step-ahead forecasts of the value-at-risk at $\alpha = 1\%$: namely, $\Omega_Y(\mathrm{VaR}_\alpha) = \Omega_Y(\mathrm{XP}_{\tau(\alpha)}) = 1/\tau(\alpha) - 1$. There are significant departures from normality for both asset returns given that the omega ratio evaluated at the first percentile of a Gaussian random variable is approximately 687.5 (Bellini and Di Bernardino, 2017). It follows from (15) that the relatively higher values of the BTC omega ratios reflect the larger differences between ES and VaR relative to what we observe for SPX returns. The latter indeed exhibit omega ratios that remain remarkably stable around 250 during the entire sample period.

Alternatively, one could instead evaluate the gain-loss ratio at the τ-th expectile, and then compute $\alpha(\tau)$ such that $\mathrm{VaR}_{\alpha(\tau)}$ and XP_τ yield exactly the same capital requirement. Figure 13 shows not only how $\alpha(\tau)$ evolves over time, but also the one-step-ahead forecasts of $\mathrm{ES}_{\alpha(\tau)}$, $\mathrm{VaR}_{\alpha(\tau)}$ and XP_τ, with $\tau = 0.01$. In particular, we employ (14) to compute $\mathrm{ES}_{\alpha(\tau)}$ in a straightforward manner. Finally, we do not plot gain-loss ratios because they are now constant over time: namely, $\Omega(\mathrm{XP}_\tau) = (1 - \tau)/\tau = 99$.

The capital requirements implied by $\mathrm{ES}_{\alpha(\tau)}$ and $\mathrm{VaR}_{\alpha(\tau)}$ decrease over time for SPX, mainly because there is an overall positive trend in $\alpha(\tau)$. In contrast, $\alpha(\tau)$ drops sharply for BTC during the first month of the COVID-19 pandemic, and then steadily decreases until April 2021 before reverting into a positive trend. These patterns are interesting because they reveal how risk managers who assess risk via expectiles would adjust their value-at-risk or expected-shortfall margins on a daily basis to keep their gain-loss ratios constant over time.

To backtest the one-step-ahead predictions of the conditional expectiles at $\tau = 0.01$, we compute the average scores in (5) and (6). Table 13 reveals that the conditional expectile forecasts based on the empirical quantiles of the QML standardized residuals best the parametric alternatives in almost every instance. The only exception refers to the 0-homogeneous scoring function in (8) for BTC. The expectile forecasts under the assumption of Gaussian errors yield, on average, slightly lower scores in this case. We actually do not observe much difference between the average scores of the two-step forecasts and those based on Gaussian errors, in general. The same does not apply to the conditional expectile forecasts under the assumption of Student-t errors, whose average scores markedly differ by taking significantly higher values.

As a sanity check, one can also compute one-step-ahead forecasts of the conditional expected shortfall using (14). The latter should obviously coincide with the conditional expected shortfall forecasts based on the empirical quantile of the QML standardized residuals. Figure 14 shows that the resulting expectile- and omega-based forecasts indeed are virtually indistinguishable. Table 14 complements the analysis by establishing that the average scores (9) and (10) of the different conditional expected-shortfall forecasts are very close to each other.

To sum up, the different lens at which we can examine margin requirements implied by expectile and value-at-risk forecasts help ameliorate risk assessment by enhancing interpretability. In particular, we can either fix the quantile α or the expectile level τ. If the former, we can then check how gain-loss ratios vary over time as a function of $\tau(\alpha)$. If the latter, we can compute how expected-shortfall and value-at-risk capital requirements change over time as a function of $\alpha(\tau)$.

7 Conclusion

In this paper, we extend Francq and Zakoïan's (2015) two-step approach for the estimation of conditional expectiles. In particular, we first estimate the parameters of the conditional mean and variance by Gaussian quasi-maximum likelihood (QML) and then employ standardized residuals to compute the unconditional expectile of the innovations' distribution. We quantify the impact of the first-step estimation error in the asymptotic distribution of the two-step estimator, as well as explore a fixed-design residual bootstrap method as in Beutner et al. (2024). Simulations show that bootstrap-based prediction intervals perform very well under the correct specification of the conditional mean and variance. However, they also reveal poor performance in terms of bias and coverage under misspecification. This is particularly true if we ignoring leverage effects in the volatility specification.

Empirically, we assess how the conditional expectile fares in comparison with traditional quantile-based risk measures in the analysis of daily cryptocurrency returns. We find that conditional expectiles yield reasonable capital requirements, balancing out the permissiveness of value-at-risk and the extreme conservatism of expected shortfall. Besides, we also exploit the connection between quantiles and expectiles to run risk assessment through the lens of gain-loss ratios.

As for future research, one can extend our asymptotic theory to deal with large portfolios, inasmuch as Francq and Zakoïan (2018) and Francq and Zakoian (2020) do using a conditional value-at-risk approach. Expectiles are particularly suitable to deal with several assets because they are coherent measures even under nonelliptical distributions Artzner et al. (1999).

References

Abdous, B., Remillard, B., 1995, Relating quantiles and expectiles under weighted-symmetry, Annals of the Institute of Statistical Mathematics 47(2), 371–384.

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

Acerbi, C., Nordio, C., Sirtori, C., 2001, Expected shortfall as a tool for financial risk management, arXiv preprint cond-mat/0102304.

Acerbi, C., Szekely, B., 2014, Back-testing expected shortfall, Risk 27(11), 76–81.

Allen, D., Chan, F., McAleer, M., Peiris, S., 2008, Finite sample properties of the QMLE for the Log-ACD model: Application to Australian stocks, Journal of Econometrics 147, 163–185.

Andrews, D. W. K., 1994, Empirical process methods in econometrics, in: R. F. Engle D. L. McFadden (eds), Handbook of Econometrics, Volume IV, Elsevier, chapter 37, pp. 2248–2294.

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

Bank for International Settlements 2013, Consultative document: Fundamental review of the trading book: A revised market risk framework, Technical report, Bank for International Settlements.

Bank for International Settlements 2014, Consultative document: Fundamental review of the trading book: Outstanding issues, Technical report, Bank for International Settlements.

Bellini, F., Bignozzi, V., 2015, On elicitable risk measures, Quantitative Finance 15(5), 725–733.

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

Bellini, F., Klar, B., Müller, A., 2018, Expectiles, omega ratios and stochastic ordering, Methodology and Computing in Applied Probability 20, 855–873.

Beutner, E., Heinemann, A., Smeekes, S., 2024, A residual bootstrap for conditional value-at-risk, Journal of Econometrics 238(2), 105554.

Bianchi, D., Babiak, M., 2022, On the performance of cryptocurrency funds, Journal of Banking and Finance 138, 106467.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31(3), 307–327.

Bollerslev, T., Engle, R. F., Nelson, D. B., 1994, ARCH models, HandBOOK of Econometrics 4, 2959–3038.

Borri, N., 2019, Conditional tail-risk in cryptocurrency markets, Journal of Empirical Finance 50, 1–19.

Breckling, J., Chambers, R., 1988, M-quantiles, Biometrika 75(4), 761–771.

Cavaliere, G., Gonçalves, S., Nielsen, M. Ø., Zanelli, E., 2023, Bootstrap inference in the presence of bias, Journal of the American Statistical Association. forthcoming.

Cavaliere, G., Pedersen, R. S., Rahbek, A., 2018, The fixed volatility bootstrap for a class of ARCH(q) models, Journal of Time Series Analysis 39(6), 920–941.

Chen, J. M., 2018, On exactitude in financial regulation: Value-at-risk, expected shortfall, and expectiles, Risks 6(2), 61.

Chen, Z., 1996, Conditional Lp-quantiles and their application to the testing of symmetry in non-parametric regression, Statistics and Probability Letters 29(2), 107–115.

Cheng, C., Parzen, E., 1997, Unified estimators of smooth quantile and quantile density functions, Journal of Statistical Planning and Inference 59(2), 291–307.

Christoffersen, P., 1998, Evaluating interval forecasts, International Economic Review 39(4), 841–862.

Christoffersen, P., Gonçalves, S., 2005, Estimation risk in financial risk management, Journal of Risk 7(3), 1–28.

Christoffersen, P., Pelletier, D., 2004, Backtesting value-at-risk: A duration-based approach, Journal of Financial Econometrics 2(1), 84–108.

Danielsson, J., Embrechts, P., Goodhart, C., Keating, C., Muennich, F., Renault, O., Shin, H. S., et al. 2001, An academic response to basel ii.

Daouia, A., Gijbels, I., Stupfler, G., 2019a, Extremiles: A new perspective on asymmetric least squares, Journal of the American Statistical Association 114(527), 1366–1381.

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

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

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

Daouia, A., Stupfler, G., Usseglio-Carleve, A., 2023, An expectile computation cookbook, Working Paper. Toulouse School of Economics.

Ding, Z., Granger, C. W. J., Engle, R. F., 1993, A long memory property of stock market returns and a new model, Journal of Empirical Finance 1(1), 83–106.

Efron, B., 1991, Regression percentiles using asymmetric squared error loss, Statistica Sinica pp. 93–125.

Engle, R. F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, Econometrica pp. 987–1007.

Falk, M., Kaufmann, E., 1991, Coverage probabilities of bootstrap confidence intervals for quantiles, Annals of Statistics 19(1), 485–495.

Fama, E. F., 1965, The behavior of stock-market prices, Journal of Business 38(1), 34–105.

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

Föllmer, H., Schied, A., 2002, Convex measures of risk and trading constraints, Finance and stochastics 6, 429–447.

Francq, C., Zakoïan, J.-M., 2015, Risk-parameter estimation in volatility models, Journal of Econometrics 184(1), 158–173.

Francq, C., Zakoïan, J.-M., 2018, Estimation risk for the VaR of portfolios driven by semi-parametric multivariate models, Journal of Econometrics 205(2), 381–401.

Francq, C., Zakoian, J.-M., 2020, Virtual historical simulation for estimating the conditional VaR of large portfolios, Journal of Econometrics 217(2), 356–380.

Gao, F., Song, F., 2008, Estimation risk in GARCH VaR and ES estimates, Econometric Theory 24(5), 1404–1424.

Ghalanos, A., 2018, Package `rugarch`, R Team Cooperation.

Girard, S., Stupfler, G., Usseglio-Carleve, A., 2021, Extreme conditional expectile estimation in heavy-tailed heteroscedastic regression models, Annals of Statistics 49(6), 3358–3382.

Gkillas, K., Katsiampa, P., 2018, An application of extreme value theory to cryptocurrencies, Economics Letters 164, 109–111.

Glosten, L. R., Jagannathan, R., Runkle, D. E., 1993, On the relation between the expected value and the volatility of the nominal excess return on stocks, Journal of Finance 48(5), 1779–1801.

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

Hansen, B., 1996, Stochastic equicontinuity for unbounded dependent heterogeneous arrays, Econometric Theory 12, 347–359.

Heinemann, A., Telg, S., 2018, A residual bootstrap for conditional expected shortfall, arXiv preprint arXiv:1811.11557.

Holzmann, H., Klar, B., 2016, Expectile asymptotics, Electronic Journal of Statistics 10(2), 2355–2371.

Hull, J., White, A., 2014, The shortfalls of expected shortfall, Risk p. 32.

Jones, M. C., 1994, Expectiles and M-quantiles are quantiles, Statistics and Probability Letters 20(2), 149–153.

Keating, C., Shadwick, W. F., 2002, A universal performance measure, Journal of Performance Measurement 6(3), 59–84.

Krätschmer, V., Zähle, H., 2017, Statistical inference for expectile-based risk measures, Scandinavian Journal of Statistics 44(2), 425–454.

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

Kupiec, P., 1995, Techniques for verifying the accuracy of risk management models, Journal of Derivatives 3(2), 73–84.

Makarov, I., Schoar, A., 2019, Price discovery in cryptocurrency markets, AEA Papers and Proceedings 109, 97–99.

Makarov, I., Schoar, A., 2020, Trading and arbitrage in cryptocurrency markets, Journal of Financial Economics 135(2), 293–319.

Makarov, I., Schoar, A., 2022, Cryptocurrencies and decentralized finance (DeFi), Brookings Papers on Economic Activity 2022(1), 141–215.

Martin, R., 2014, Expectiles behave as expected, Risk 27, 79–83.

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

McNeil, A. J., Frey, R., Embrechts, P., 2015, Quantitative Risk Management: Concepts, Techniques and Tools, Princeton University Press.

Nelson, D. B., 1991, Conditional heteroskedasticity in asset returns: A new approach, Econometrica 59(2), 347–370.

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

Nguyen, L. H., Chevapatrakul, T., Yao, K., 2020, Investigating tail-risk dependence in the cryptocurrency markets: A lasso quantile regression approach, Journal of Empirical Finance 58, 333–355.

Nolde, N., Ziegel, J. F., 2017, Elicitability and backtesting: Perspectives for banking regulation, Annals of Applied Statistics 11(4), 1833 – 1874.

Pascual, L., Romo, J., Ruiz, E., 2006, Bootstrap prediction for returns and volatilities in GARCH models, Computational Statistics and Data Analysis 50(9), 2293–2312.

Philipps, C., 2022, Interpreting expectiles, Available at SSRN 3881402.

Remillard, B., 2013, Statistical methods for financial engineering, CRC press.

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

Saerens, M., 2000, Building cost functions minimizing to some summary statistics, IEEE Transactions on neural networks 11(6), 1263–1271.

Scaillet, O., Treccani, A., Trevisan, C., 2018, High-frequency jump analysis of the Bitcoin market, Journal of Financial Econometrics 18(2), 209–232.

Schmidt, P., Katzfuss, M., Gneiting, T., 2021, Interpretation of point forecasts with unknown directive, Journal of Applied Econometrics 36(6), 728–743.

Sherman, M., Carlstein, E., 2004, Confidence intervals based on estimators with unknown rates of convergence, Computational Statistics and Data Analysis 46(1), 123–139.

Spierdijk, L., 2016, Confidence intervals for ARMA-GARCH value-at-risk: The case of heavy tails and skewness, Computational Statistics and Data Analysis 100, 545–559.

Taylor, J. W., 2008, Estimating value at risk and expected shortfall using expectiles, Journal of Financial Econometrics 6(2), 231–252.

Taylor, J. W., 2022, Forecasting value at risk and expected shortfall using a model with a dynamic omega ratio, Journal of Banking and Finance 140, 106519.

Thomson, W., 1979, Eliciting production possibilities from a well-informed manager, Journal of Economic Theory 20(3), 360–380.

Waltrup, L. S., Sobotka, F., Kneib, T., Kauermann, G., 2015, Expectile and quantile regression: David and goliath?, Statistical Modelling 15(5), 433–456.

Wintenberger, O., 2013, Continuous invertibility and stable QML estimation of the EGARCH(1,1) model, Scandinavian Journal of Statistics 40(4), 846–867.

Yamai, Y., Yoshiba, T., 2002, Comparative analyses of expected shortfall and value-at-risk: Their estimation error, decomposition, and optimization, Monetary and Economic Studies 20(1), 87–121.

Yao, Q., Tong, H., 1996, Asymmetric least squares regression estimation: A nonparametric approach, Journal of Nonparametric Statistics 6(2-3), 273–292.

Zakoian, J.-M., 1994, Threshold heteroskedastic models, Journal of Economic Dynamics and Control 18(5), 931–955.

Zhang, W., Li, Y., Xiong, X., Wang, P., 2021, Downside risk and the cross-section of cryptocurrency returns, Journal of Banking and Finance 133, 106246.

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

Figure 1: Minimum capital requirements for Gaussian and Student-t loss distributions

We contrast the inverse empirical distribution function of the daily losses (in black and grey) simulated from Gaussian ($\nu = 500$) and t-distributions ($\nu = 8$) with expected shortfall (in blue), expectile (in red), and value-at-risk (in green) for $\alpha = \tau = 0.99$ for small ($T = 250$), medium ($T = 500$) and large ($T = 1000$) sample sizes.

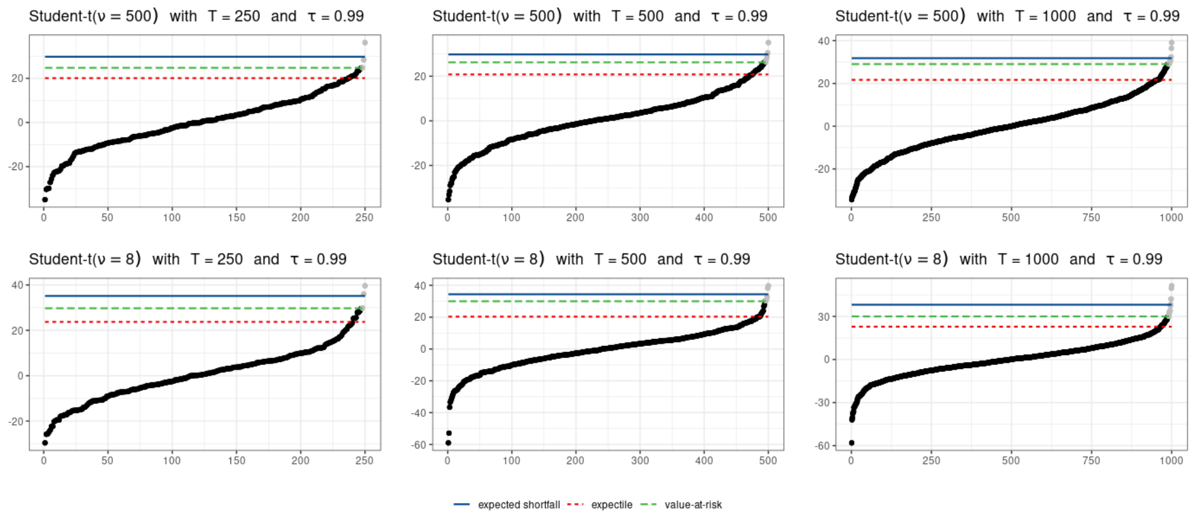

Figure 2: Quantiles and expectiles of the Gaussian and Student-t distributions

We illustrate the quantile (in green), expected shortfall (in blue) and expectile (in red) functions for the Gaussian and Student-t distributions ($\nu = 500, 8$, respectively).

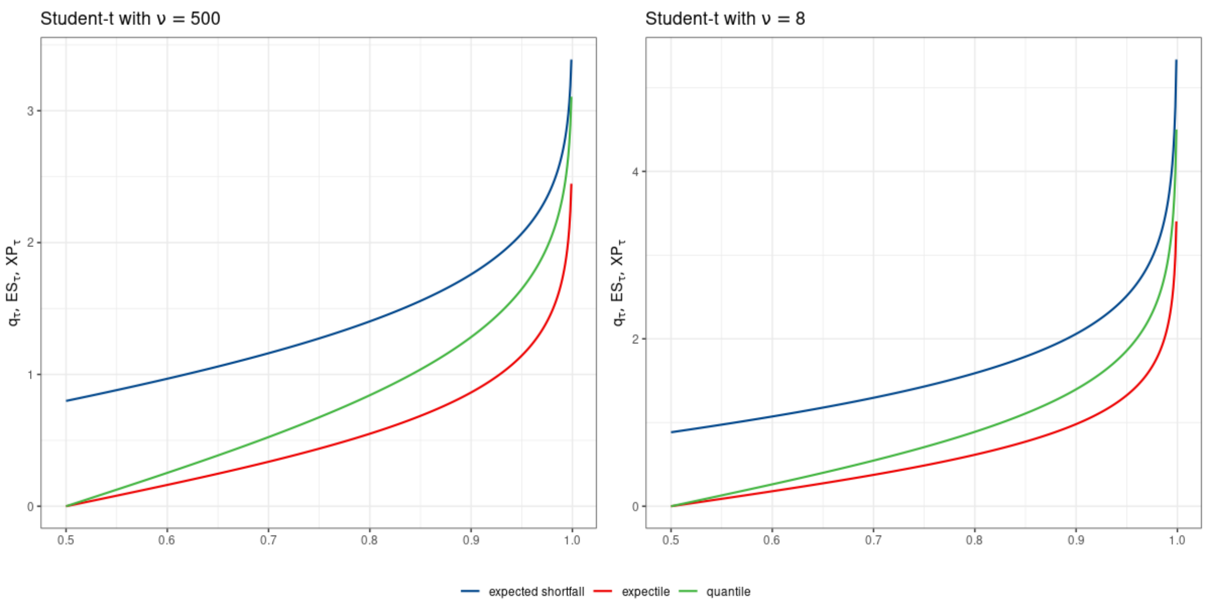

Figure 3: Distribution of the unconditional 5%-expectile estimates of GARCH innovations
We display box plots of the difference between the sample and true expectiles for the different data generating processes and sample sizes. In particular, we report the distributions of the sample expectiles using either true innovations (in green) or standardized residuals (in red), whereas the corresponding distribution for the fixed-design residual bootstrap appears in blue.

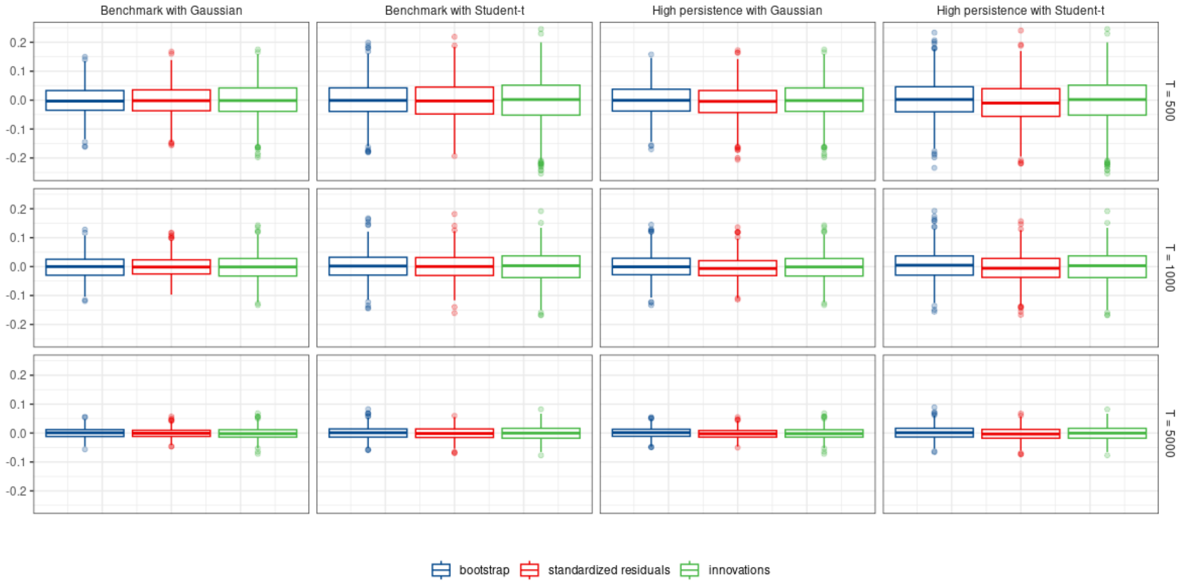

Figure 4: Distribution of the conditional 5%-expectile forecast errors
We exhibit box plots of the two-step forecast errors of the conditional 5%-expectile for the different data generating processes and sample sizes. We also display the corresponding distributions for the difference between the two-step forecasts in the fixed-design residual bootstrap and original samples.

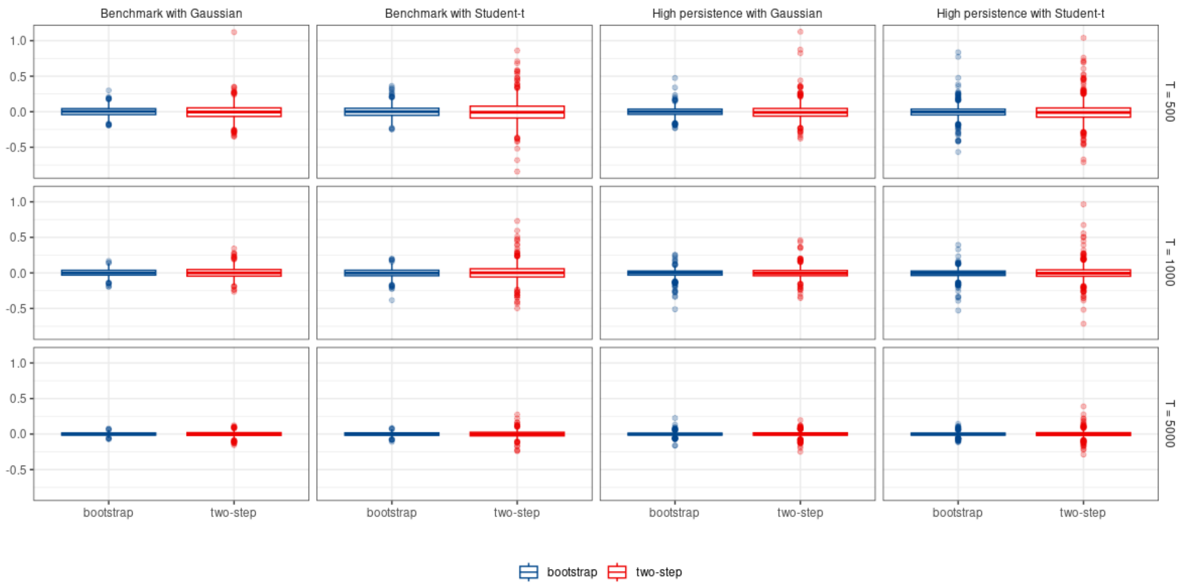

Figure 5: Relative bias and RMSE of the ALS-based risk measures at the $\tau = 5\%$ level

We plot bias (solid lines) and root mean squared error (dotted lines) of the unconditional and conditional expectile and extremile risk measures varying from 250 to 5000, in increments of 250 observations. For the unconditional risk measures, we plot the results for the sample estimators using the true innovations in green, as well as those for the two-step estimators using the original and bootstrap samples respectively in red and blue. For the conditional risk measures, we instead plot the bias and RMSE of the two-step forecasts of the conditional expectiles and extremiles.

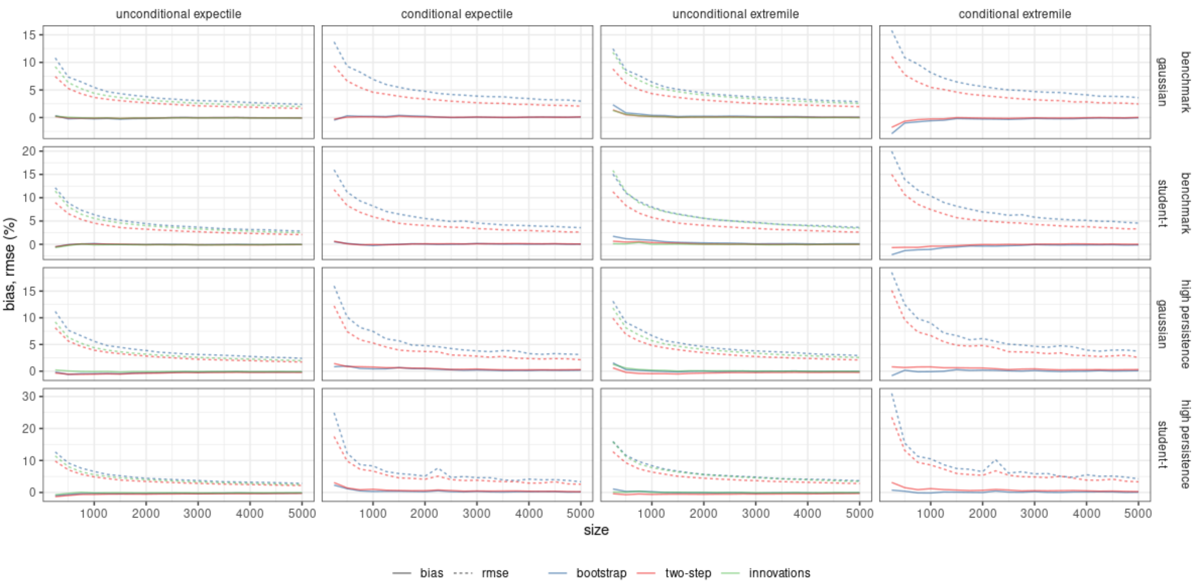

Figure 6: Distribution of the unconditional 5%-expectiles of APARCH innovations

We display box plots of the difference between the sample and true expectiles for the different data generating processes with Gaussian errors and $T = 1000$. In particular, we compute sample extremiles using either true innovations (in green) or GARCH(1,1) standardized residuals (in red), whereas the corresponding distribution for the fixed-design residual bootstrap appears in blue.

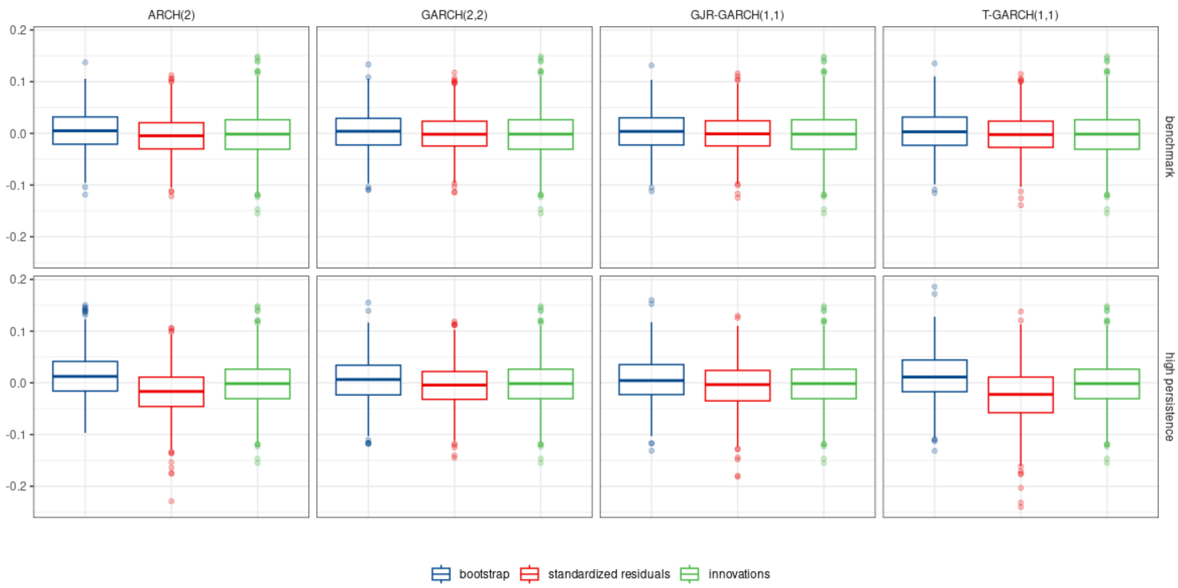

Figure 7: Distribution of the conditional 5%-expectile forecast errors under misspecification
We exhibit box plots of the two-step forecast errors of the conditional 5%-expectile based on a GARCH(1,1) volatility model for the different APARCH data generating processes and a sample size of $T = 1000$. We also display the corresponding distributions for the difference between the two-step forecasts in the fixed-design residual bootstrap and original samples.

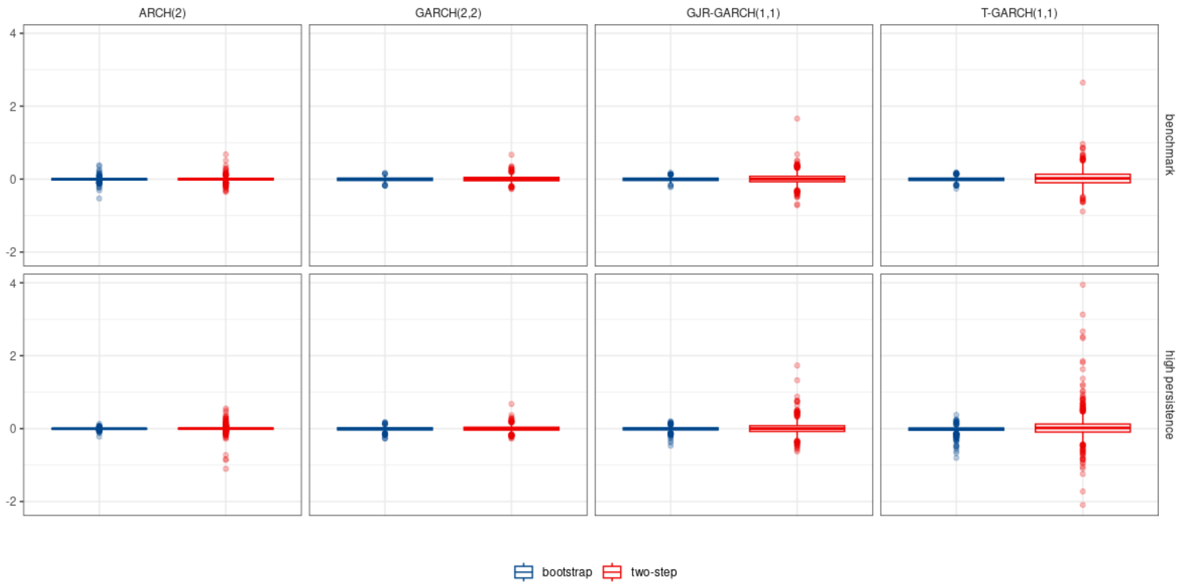

Figure 8: Daily time series of the S&P 500 equity index and exchange rates
We plot the historical daily prices of Cardano (ADA) Binance Coin (BNB), Bitcoin (BTC), Ethereum (ETH), Euro-Dollar exchange rate (EUR), and S&P 500 equity index (SPX) from January 2016 to December 2023.

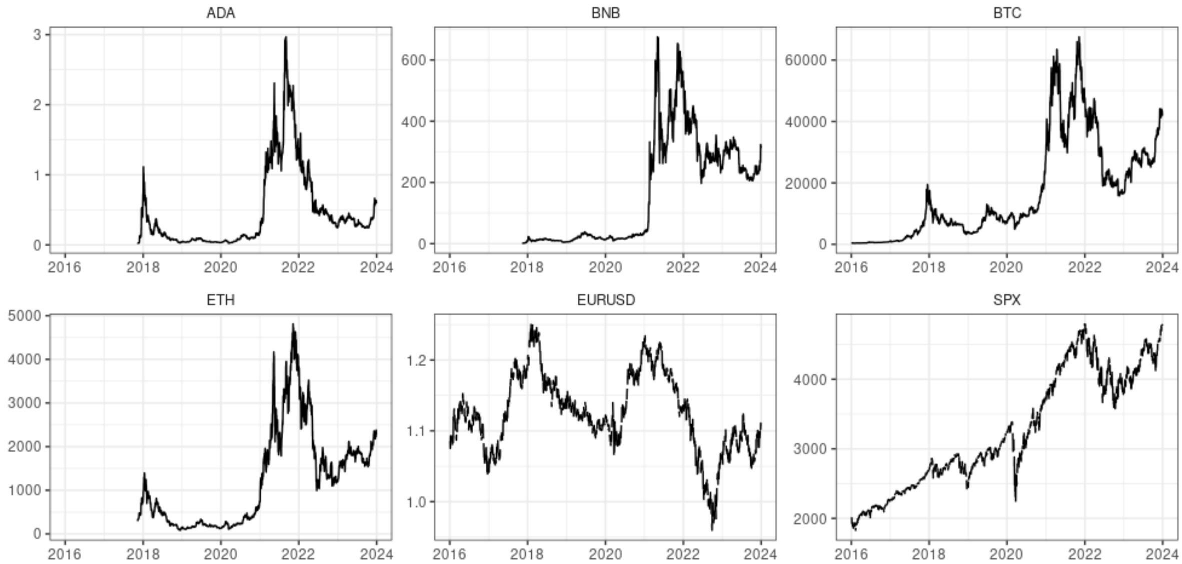

Figure 9: Box plots of the historical drawdowns of traditional and crypto assets

We plot the distributions of the historical drawdowns of the S&P 500 index (SPX) and Euro-Dollar exchange rate (EUR), as well as of the most liquid cryptocurrency exchange rates: namely, Cardano (ADA), Binance Coin (BNB), Bitcoin (BTC), and Ethereum (ETH). As we have data only as from 2017 for most cryptocurrencies, we restrict attention to the period between 2018 and 2023.

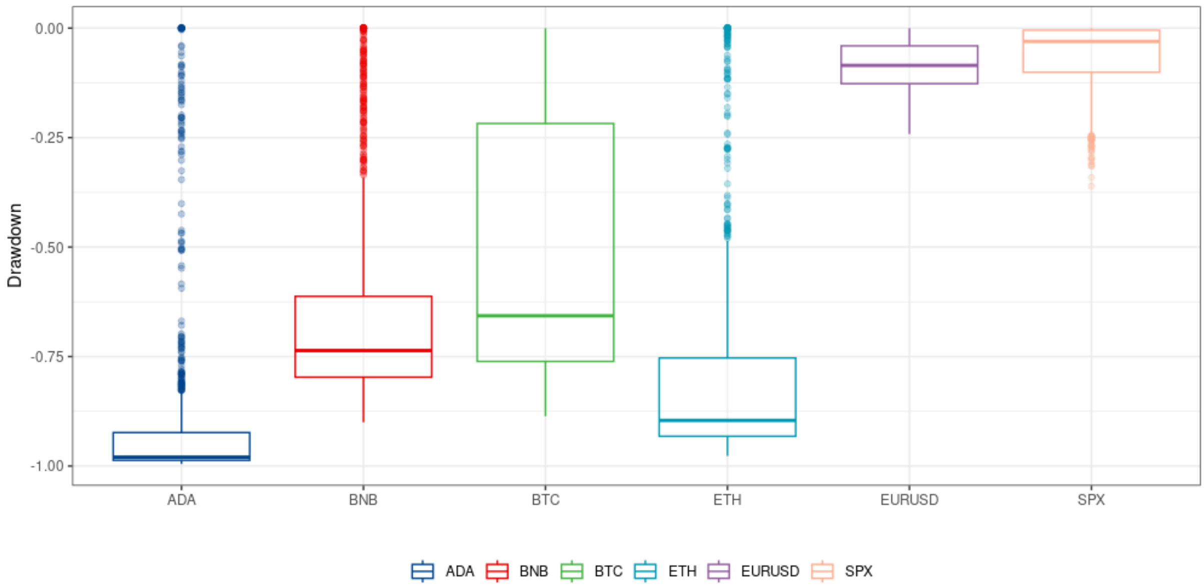

Figure 10: One-day-ahead forecasts of the conditional risk measures at $\tau = 1\%$

We plot daily returns (in black) and the one-step-ahead forecasts of the expected shortfall, expectile, and value-at-risk (in color). The corresponding 90% bootstrap-based prediction intervals are in grey.

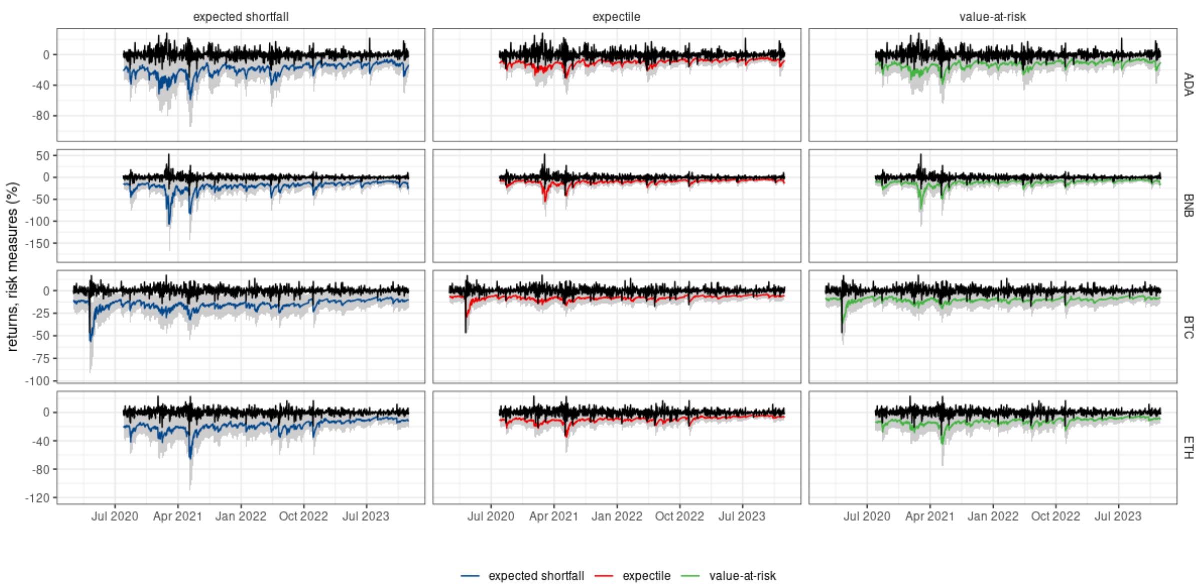

Figure 11: Distribution of the difference between capital requirement predictions and realized returns.

We plot the difference between capital requirement forecasts based on each conditional risk measure with $\tau = 1\%$ and realized returns.

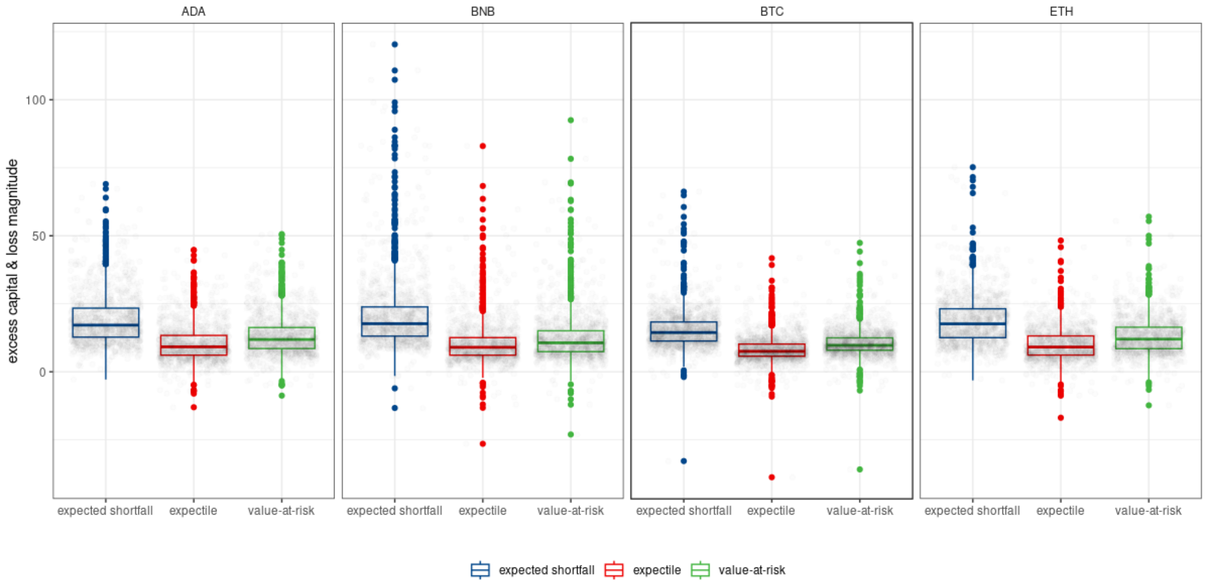

Figure 12: One-step-ahead forecasts of conditional risk measures at $\alpha = 1\%$

The first row plots the conditional expected shortfall and value-at-risk forecasts with $\alpha = 1\%$, as well as the conditional expectiles at $\tau(\alpha)$, for both BTC and SPX. We fix $\tau(\alpha)$ in order to match the capital requirements implied by the value-at-risk and expectile measures. The second row of plots display the corresponding omega ratios.

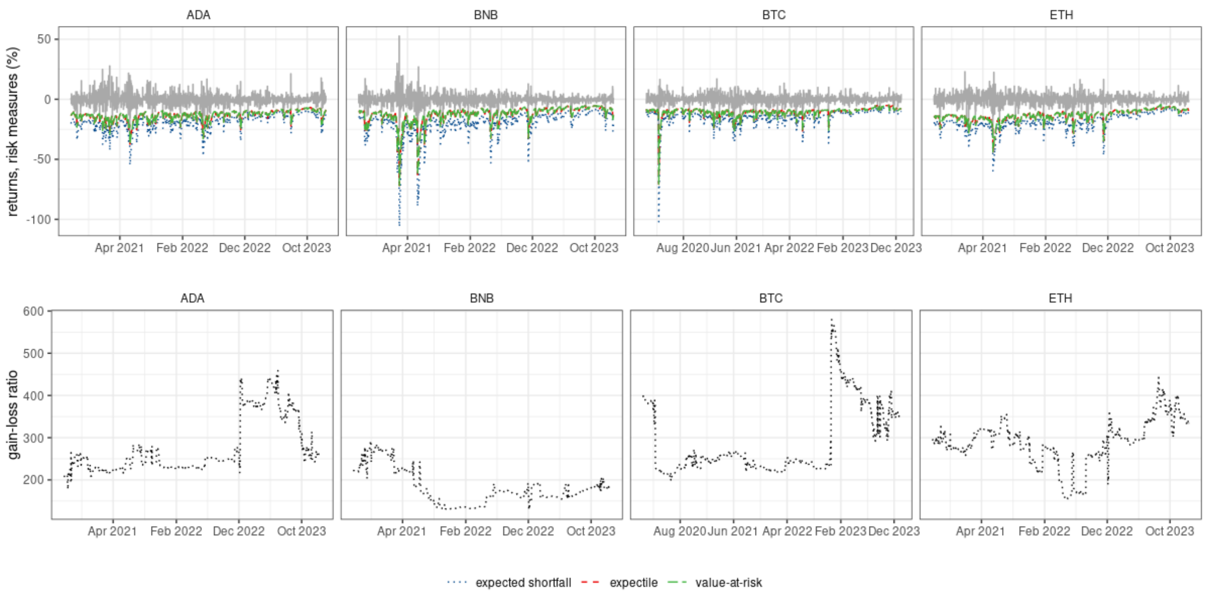

Figure 13: One-step-ahead forecasts of conditional risk measures at $\tau = 1\%$

The first row plots the conditional expectile forecasts at $\tau = 1\%$, as well as the conditional expected shortfall and value-at-risk forecasts at $\alpha(\tau)$, for both BTC and SPX. The second row of plots display how $\alpha(\tau)$ evolves over time so as to match the capital requirements implied by the value-at-risk and expectile measures.

Figure 14: Comparison of conditional expected shortfall forecasts

We plot the one-step-ahead forecasts of the conditional expected shortfall, as well as their expectile- and omega-implied forecasts, for BT and SPX.

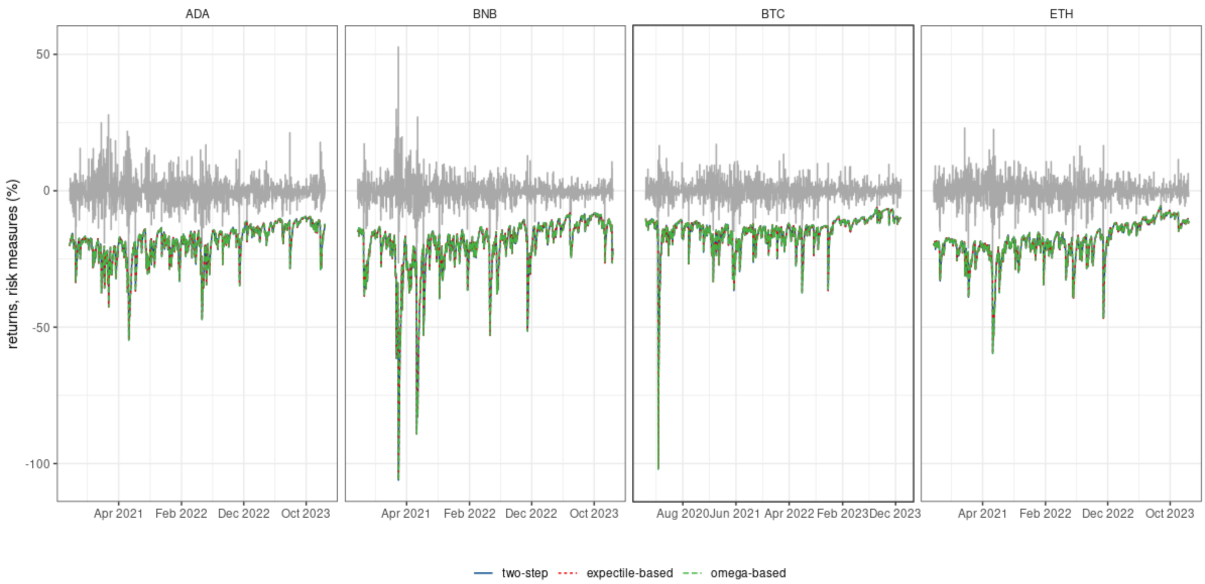

Table 1: Bootstrap-based 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, for GARCH(1,1) processes with Gaussian innovations ($\nu = 500$)

sample	risk	prediction	benchmark				high persistence			
size	measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
500	VaR	EP	5.91	7.39	86.70	0.49	6.13	9.01	84.86	0.41
		RT	2.51	7.39	90.10	0.49	2.57	9.01	88.42	0.43
		SY	2.95	7.67	89.38	0.50	3.12	9.06	87.82	0.44
	ES	EP	5.19	10.47	84.34	0.54	5.42	12.85	81.73	0.46
		RT	2.27	10.47	87.26	0.57	2.56	12.85	84.59	0.50
		SY	3.19	9.63	87.18	0.58	3.32	11.14	85.54	0.51
	XP	EP	4.63	7.48	87.89	0.33	5.27	8.95	85.78	0.28
		RT	3.48	7.48	89.04	0.33	3.40	8.95	87.65	0.29
		SY	3.74	7.30	88.96	0.33	3.93	8.59	87.48	0.29
1000	VaR	EP	5.15	5.94	88.91	0.35	5.17	7.44	87.39	0.30
		RT	3.05	5.94	91.01	0.35	2.83	7.44	89.73	0.31
		SY	3.23	6.34	90.43	0.36	3.24	7.37	89.39	0.31
	ES	EP	4.40	8.23	87.37	0.40	4.93	9.96	85.11	0.34
		RT	2.90	8.23	88.87	0.41	2.80	9.96	87.24	0.36
		SY	3.39	7.53	89.08	0.42	3.51	8.64	87.85	0.37
	XP	EP	4.29	6.53	89.18	0.24	4.48	8.06	87.46	0.20
		RT	3.68	6.53	89.79	0.24	3.35	8.06	88.59	0.21
		SY	3.77	6.28	89.95	0.24	3.74	7.31	88.95	0.21

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $\alpha_0 + \beta_0 = 0.90$ in the benchmark case and $\alpha_0 + \beta_0 = 0.99$ for high persistence.

Table 2: Bootstrap-based 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, for GARCH(1,1) processes with Student-t innovations ($\nu = 8$)

sample	risk	prediction	benchmark				high persistence			
size	measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
500	VaR	EP	5.81	6.59	87.60	0.60	6.01	8.11	85.88	0.49
		RT	3.22	6.59	90.19	0.58	3.08	8.11	88.81	0.49
		SY	3.37	7.80	88.83	0.59	3.45	8.90	87.65	0.51
	ES	EP	13.42	3.92	82.66	0.77	13.11	4.56	82.33	0.62
		RT	9.26	3.92	86.82	0.79	8.39	4.56	87.05	0.66
		SY	10.34	3.73	85.93	0.79	9.79	4.41	85.80	0.68
	XP	EP	4.85	8.07	87.08	0.42	4.94	9.59	85.47	0.34
		RT	3.63	8.07	88.30	0.41	3.40	9.59	87.01	0.35
		SY	3.79	8.59	87.62	0.42	3.81	9.47	86.72	0.35
1000	VaR	EP	5.24	6.18	88.58	0.43	5.69	7.02	87.29	0.34
		RT	3.42	6.18	90.40	0.42	3.30	7.02	89.68	0.35
		SY	3.39	6.80	89.81	0.43	3.57	7.51	88.92	0.35
	ES	EP	17.78	1.95	80.27	0.58	18.78	2.22	79.00	0.46
		RT	15.68	1.95	82.37	0.58	15.05	2.22	82.73	0.48
		SY	16.14	1.92	81.94	0.58	16.10	2.10	81.80	0.48
	XP	EP	4.26	7.27	88.47	0.30	4.47	8.01	87.52	0.24
		RT	3.91	7.27	88.82	0.30	3.85	8.01	88.14	0.25
		SY	3.80	7.28	88.92	0.30	3.80	7.94	88.26	0.25

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $\alpha_0 + \beta_0 = 0.90$ in the benchmark case and $\alpha_0 + \beta_0 = 0.99$ for high persistence.

Table 3: Bootstrap-based 90% prediction intervals of the conditional risk measures at the $\tau = 1\%$ level, for GARCH(1,1) processes with Gaussian innovations ($\nu = 500$)

sample	risk	prediction	benchmark				high persistence			
size	measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
500	VaR	EP	6.16	12.15	81.69	0.77	6.23	14.41	79.36	0.66
		RT	0.88	12.15	86.97	0.77	1.01	14.41	84.58	0.68
		SY	1.38	11.91	86.71	0.81	1.55	13.83	84.62	0.72
	ES	EP	5.94	19.55	74.51	0.76	6.00	22.19	71.81	0.65
		RT	0.77	19.55	79.68	0.86	0.83	22.19	76.98	0.76
		SY	1.76	15.80	82.44	0.89	1.78	18.01	80.21	0.79
	XP	EP	5.02	10.10	84.88	0.48	5.25	12.29	82.46	0.41
		RT	2.36	10.10	87.54	0.50	2.65	12.29	85.06	0.44
		SY	3.12	9.67	87.21	0.50	3.40	11.21	85.39	0.44
1000	VaR	EP	5.51	9.17	85.32	0.57	5.69	10.48	83.83	0.49
		RT	1.46	9.17	89.37	0.57	1.50	10.48	88.02	0.50
		SY	1.92	9.37	88.71	0.59	1.91	10.42	87.67	0.52
	ES	EP	4.80	13.82	81.38	0.61	4.76	15.15	80.09	0.52
		RT	1.39	13.82	84.79	0.65	1.46	15.15	83.39	0.58
		SY	2.38	11.47	86.15	0.66	2.36	12.40	85.24	0.59
	XP	EP	4.15	8.17	87.68	0.35	4.62	9.59	85.79	0.30
		RT	2.98	8.17	88.85	0.36	2.88	9.59	87.53	0.32
		SY	3.29	7.65	89.06	0.36	3.46	8.79	87.75	0.32

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 1\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $\alpha_0 + \beta_0 = 0.90$ in the benchmark case and $\alpha_0 + \beta_0 = 0.99$ for high persistence.

Table 4: Bootstrap-based 90% prediction intervals of the conditional risk measures at the $\tau = 1\%$ level, for GARCH(1,1) processes with Student-t innovations ($\nu = 8$)

sample size	risk measure	prediction interval	benchmark lower	benchmark upper	benchmark coverage	benchmark length	high persistence lower	high persistence upper	high persistence coverage	high persistence length
500	VaR	EP	5.70	12.00	82.30	1.19	5.79	13.20	81.01	0.99
		RT	1.16	12.00	86.84	1.12	1.20	13.20	85.60	0.95
		SY	1.54	12.80	85.66	1.17	1.51	14.03	84.46	1.00
	ES	EP	0.50	59.76	39.74	1.28	0.48	61.88	37.64	1.06
		RT	0.03	59.76	40.21	1.42	0.07	61.88	38.05	1.21
		SY	0.06	54.08	45.86	1.46	0.10	55.84	44.06	1.27
	XP	EP	4.17	12.19	83.64	0.73	4.13	13.74	82.13	0.60
		RT	2.26	12.19	85.55	0.73	2.28	13.74	83.98	0.62
		SY	2.81	11.99	85.20	0.74	2.65	13.07	84.28	0.63
1000	VaR	EP	5.16	9.54	85.30	0.87	5.30	10.35	84.35	0.71
		RT	1.63	9.54	88.83	0.83	1.76	10.35	87.89	0.69
		SY	1.94	10.27	87.79	0.86	1.95	10.78	87.27	0.72
	ES	EP	0.10	63.88	36.02	1.07	0.09	65.91	34.00	0.87
		RT	0.01	63.88	36.11	1.12	0.00	65.91	34.09	0.94
		SY	0.02	60.68	39.30	1.13	0.04	62.31	37.65	0.96
	XP	EP	3.18	9.74	87.08	0.55	3.64	10.94	85.42	0.44
		RT	2.82	9.74	87.44	0.54	2.67	10.94	86.39	0.45
		SY	2.75	9.70	87.55	0.54	2.88	10.57	86.55	0.45

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 1\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $\alpha_0 + \beta_0 = 0.90$ in the benchmark case and $\alpha_0 + \beta_0 = 0.99$ for high persistence.

Table 5: Bootstrap-based 90% prediction intervals of the conditional risk measures based on GARCH(1,1) standardized residuals at the $\tau = 5\%$ level, for ARCH(2) processes with Gaussian innovations ($\nu = 500$)

risk	prediction	benchmark				high persistence			
measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
VaR	EP	6.54	6.90	86.56	0.21	15.58	15.29	69.13	0.07
	RT	4.04	6.90	89.06	0.21	14.58	15.29	70.13	0.08
	SY	4.25	7.16	88.59	0.22	13.96	14.72	71.32	0.08
ES	EP	5.50	9.11	85.39	0.23	16.97	15.35	67.68	0.09
	RT	3.63	9.11	87.26	0.25	16.39	15.35	68.26	0.09
	SY	4.16	8.14	87.70	0.26	16.12	13.71	70.17	0.10
XP	EP	5.58	7.14	87.28	0.14	15.67	14.98	69.35	0.05
	RT	4.63	7.14	88.23	0.15	16.10	14.98	68.92	0.05
	SY	4.93	6.79	88.28	0.15	15.15	14.04	70.81	0.05

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $\alpha_1 = 0.80$ and $\alpha_2 = 0.10$ in the benchmark case and $\alpha_1 = 0.89$ and $\alpha_2 = 0.10$ for high persistence. We consider a sample size of $T = 1000$ observations.

Table 6: Bootstrap-based 90% prediction intervals of the conditional risk measures based on GARCH(1,1) standardized residuals at the $\tau = 5\%$ level, for GARCH(2,2) processes with Gaussian innovations ($\nu = 500$)

risk	prediction	benchmark				high persistence			
measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
VaR	EP	5.79	6.63	87.58	0.35	6.08	8.60	85.32	0.30
	RT	3.64	6.63	89.73	0.35	3.57	8.60	87.83	0.31
	SY	4.01	6.69	89.30	0.36	4.02	8.16	87.82	0.32
ES	EP	4.52	9.32	86.16	0.40	5.07	11.13	83.80	0.34
	RT	3.11	9.32	87.57	0.41	3.18	11.13	85.69	0.37
	SY	3.40	8.08	88.52	0.42	3.85	9.65	86.50	0.37
XP	EP	4.63	6.97	88.40	0.24	5.27	8.85	85.88	0.20
	RT	4.40	6.97	88.63	0.24	4.18	8.85	86.97	0.21
	SY	4.28	6.60	89.12	0.24	4.53	8.03	87.44	0.21

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (0.07, 0.03, 0.80, 0.10)$ in the benchmark case and $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (0.07, 0.03, 0.89, 0.10)$ for high persistence. We consider a sample size of $T = 1000$ observations..

Table 7: Bootstrap-based 90% prediction intervals of the conditional risk measures based on GARCH(1,1) standardized residuals at the $\tau = 5\%$ level, for GJR-GARCH(1,1) processes with Gaussian innovations ($\nu = 500$)

risk	prediction	benchmark				high persistence			
measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
VaR	EP	16.69	17.41	65.90	0.36	21.38	24.19	54.43	0.29
	RT	15.64	17.41	66.95	0.35	20.15	24.19	55.66	0.30
	SY	15.41	17.21	67.38	0.36	20.13	23.60	56.27	0.31
ES	EP	18.87	19.09	62.04	0.41	24.54	24.19	51.27	0.34
	RT	16.96	19.09	63.95	0.42	22.07	24.19	53.74	0.36
	SY	17.64	18.17	64.19	0.42	22.95	22.97	54.08	0.36
XP	EP	17.23	17.27	65.50	0.24	22.33	24.05	53.62	0.20
	RT	16.62	17.27	66.11	0.24	21.18	24.05	54.77	0.20
	SY	16.82	17.11	66.07	0.24	21.45	23.19	55.36	0.21

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $(\alpha_1, \gamma_1, \beta_1) \approx (0.09, 0.33, 0.80)$ in the benchmark case and $(\alpha_1, \gamma_1, \beta_1) \approx (0.09, 0.33, 0.89)$ for high persistence. We consider a sample size of $T = 1000$ observations.

Table 8: Bootstrap-based 90% prediction intervals of the conditional risk measures based on GARCH(1,1) standardized residuals at the $\tau = 5\%$ level, for TGARCH(1,1) processes with Gaussian innovations ($\nu = 500$)

risk	prediction	benchmark				high persistence			
measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
VaR	EP	26.78	25.84	47.38	0.37	31.31	30.93	37.76	0.32
	RT	26.56	25.84	47.60	0.37	30.94	30.93	38.13	0.35
	SY	26.10	25.99	47.91	0.38	30.46	30.17	39.37	0.37
ES	EP	31.06	24.36	44.58	0.44	36.60	27.81	35.59	0.38
	RT	29.76	24.36	45.88	0.45	34.97	27.81	37.22	0.43
	SY	30.47	23.77	45.76	0.45	35.40	26.48	38.12	0.45
XP	EP	28.29	25.18	46.53	0.25	33.26	29.91	36.83	0.22
	RT	27.93	25.18	46.89	0.25	32.61	29.91	37.48	0.24
	SY	28.00	25.01	46.99	0.25	32.67	28.99	38.34	0.25

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $(\alpha_1, \gamma_1, \beta_1) \approx (0.125, 0.60, 0.80)$ in the benchmark case and $(\alpha_1, \gamma_1, \beta_1) \approx (0.125, 0.60, 0.89)$ for high persistence. We consider a sample size of $T = 1000$ observations.

Table 9: Double Bootstrap-based 90% prediction intervals of the conditional risk measures based on GARCH(1,1) standardized residuals at the $\tau = 5\%$ level, for TGARCH(1,1) processes with Gaussian innovations ($\nu = 500$)

risk	prediction	Single-Bootstrap				Double-Bootstrap			
measure	interval	lower	upper	coverage	length	lower	upper	coverage	length
VaR	EP								
	RT								
	SY								
ES	EP								
	RT								
	SY								
XP	EP								
	RT								
	SY								

We report average coverage rates (%), lower and upper exception rates (%) and lengths of the 90% prediction intervals of the conditional risk measures at the $\tau = 5\%$ level, over $S = 10,000$ replications. The bootstrap-based intervals rest on $B = 999$ bootstrap samples, with EP, RT and SY corresponding to the equal-tail percentile, reverse tail and symmetric intervals that we describe respectively in (22) to (24). We consider two degrees of persistence in volatility: namely, $(\alpha_1, \gamma_1, \beta_1) \approx (0.125, 0.60, 0.80)$ in the benchmark case and $(\alpha_1, \gamma_1, \beta_1) \approx (0.125, 0.60, 0.89)$ for high persistence. We consider a sample size of $T = 1000$ observations.

Table 10: Descriptive statistics of the daily continuously compounded returns

ticker	T	min	$q_{0.25}$	$q_{0.50}$	$q_{0.75}$	max	mean	std dev	skew	kurt
ADA	2,243	-50.36	-2.70	0.02	2.62	86.15	0.13	6.31	1.99	29.67
BNB	2,243	-54.31	-1.88	0.10	2.35	52.92	0.23	5.46	0.40	20.05
BTC	2,921	-46.47	-1.27	0.15	1.69	22.51	0.16	3.75	-0.72	14.83
ETH	2,243	-55.07	-1.90	0.08	2.38	23.47	0.09	4.80	-0.93	13.85
EUR	2,082	-2.81	-0.28	0.00	0.28	1.82	0.00	0.47	-0.05	4.91
SPX	2,011	-12.77	-0.38	0.07	0.58	8.97	0.04	1.18	-0.85	19.19

We report the minimum (min), average (mean) and maximum (max) log-returns for each asset, as well as their standard deviations (std dev), skewness (skew) and kurtosis (kurt), and first to third quantiles ($q_{0.25}$, $q_{0.50}$ and $q_{0.75}$, respectively). Assets include the S&P 500 equity index (SPX) and Euro-Dollar exchange rate (EUR), as well as the Bitcoin (BTC), Ethereum (ETH), Binance Coin (BNB), and Cardano (ADA) exchange rates. The sample ranges from January 2016 to December 2023 for SPX, EUR and BTC, and from November 2017 to December 2023 for ADA, BNB and ETH.

Table 11: Backtesting the conditional risk measures of BTC and SPX

asset	risk	exceptions		coverage test			residual
ticker	measure	expected	observed	unconditional	conditional	duration	test
ADA	VaR	12	11	0.68	0.83	0.47	
	ES		5				0.99
	XP		38				0.56
BNB	VaR	12	13	0.87	0.86	0.07	
	ES		5				0.60
	XP		38				0.54
BTC	VaR	19	21	0.30	0.73	0.23	
	ES		5				0.59
	XP		38				0.55
ETH	VaR	12	9	0.68	0.55	0.06	
	ES		5				0.90
	XP		38				0.56

We report the expected and observed number of exceptions, as well as the p-values of the coverage and residual tests. In particular, we compute the duration coverage test by (Christoffersen and Pelletier, 2004), the unconditional coverage test by (Kupiec, 1995), and the conditional coverage test by (Christoffersen, 1998) to assess the performance of the value-at-risk forecasts. The last column reports bootstrap-based p-values of the residual test by (McNeil and Frey, 2000) for the expected shortfall and expectile forecasts.

Table 12: Backtest performance of the conditional value-at-risk forecasts

		violations		coverage tests			scores	
ticker	method	expected	realized	unconditional	conditional	duration	(7)	(8)
BTC	two-step	19	21	0.68	0.73	0.23	14.57	2.60
	Gaussian	19	34	0.00	0.01	0.49	15.04	2.68
	Student-t	19	3	0.00	0.00	0.09	18.20	2.83
SPX	two-step	10	10	0.97	0.90	0.96	4.47	1.44
	Gaussian	10	21	0.00	0.00	0.26	15.04	1.55
	Student-t	10	8	0.49	0.73	0.66	18.21	1.63

We report the expected and realized number of violations, as well as the p-values of the duration, unconditional, conditional coverage tests to assess the one-step-ahead forecasts of $\mathrm{VaR}_{0.01}$ based on the empirical quantile of the QML standardized residuals relative to those under the assumptions of Gaussian or Student-t errors. We also compute the average scores in (7) and (8), whose scoring functions are strictly consistent for the value-at-risk measure (Nolde and Ziegel, 2017).

Table 13: Backtest performance of the conditional expectile forecasts

		average score	
ticker	method	(5)	(6)
BTC	two-step	172.23	1.39
	Gaussian	173.19	1.37
	Student-t	215.13	1.50
SPX	two-step	14.69	0.15
	Gaussian	15.01	0.16
	Student-t	23.00	0.28

We report the average scores of the one-step-ahead forecasts of the conditional expectile at $\tau = 1\%$. We employ the scoring functions in (5) and (6), which are both strictly consistent for expectiles (Nolde and Ziegel, 2017).

Table 14: Backtest performance of the conditional expected-shortfall forecasts

		average score	
ticker	method	(9)	(10)
BTC	two-step	3.8213	2.6927
	XP-based	3.8215	2.6910
	Ω-based	3.8213	2.6927
SPX	two-step	2.1038	1.4859
	XP-based	2.1039	1.4858
	Ω-based	2.1038	1.4859

We report the average scores of the one-step-ahead forecasts of the conditional expectile at $\tau = 1\%$. We employ the scoring functions in (9) and (10), which are both strictly consistent for expected shortfall (Nolde and Ziegel, 2017).

A Technical Appendix

In this section we will show the asymptotic properties of the estimated expectile with estimated parameters.

Conditions

In this section we list the assumptions used in the derivation of our results. These assumptions are in line with (Francq and Zakoïan, 2015; Heinemann and Telg, 2018; Francq and Zakoian, 2020) and (Beutner et al., 2024).

We consider the following model for the compound returns

$$y_{t+1} = \sigma_t(\boldsymbol{\theta}_0)\eta_{t+1}, \quad \text{with} \quad \sigma_t(\boldsymbol{\theta}_0) := \sigma(y_t, y_{t-1}, \ldots; \boldsymbol{\theta}), \quad t \in \mathbb{Z},$$

where $\sigma_t(\cdot)$, $\boldsymbol{\theta}_0$, and $\{\eta_t\}$ satisfy the regularity conditions below.

Assumption 1 (Innovation process)**.** *Innovations $\{\eta_t\}$ are independent and identically distributed according to an absolutely continuous distribution F_η with mean zero, variance one, and $\mathbb{E}[\eta_0^s] < \infty$. Its density function, f_η, is continuous and strictly positive around XP_τ^η.*

We restrict the distribution of the innovation process to have an absolutely continuous density that is continuous around the target expectile. This condition is parallel to that in Theorem 4 of (Francq and Zakoïan, 2015) for the quantile case and rules out distributions with jumps, for instance.

Assumption 2 (Stationarity & Ergodicity)**.** *$\{y_t\}$ is a strictly stationary and geometric strong mixing solution of* (16)*.*

This condition implies the process is ergodic and its dependence is sufficiently well behaved. This condition differs from that in (Francq and Zakoïan, 2015) (Condition A1). However, under a garch-type volatility process, the solution $\{y_t\}$ is geometrically beta mixing Carrasco and chen 2003, which implies our mixing assumption.

Assumption 3 (Compactness and interior)**.** *Θ is a compact subset of $\mathbb{R}^m$ and $\boldsymbol{\theta}_0$ is an interior point of Θ.*

This condition is standard in the literature.

Assumption 4 (Volatility process)**.** *The function $\sigma : \mathbb{R}^\infty \times \Theta \to (0, \infty)$ is known and for any real sequence $\{x_i\}$, the function $\theta \to \sigma(x_1, x_2, \ldots; \theta)$ is continuous. Almost surely $\inf_{\boldsymbol{\theta} \in \Theta} \sigma_t(\theta) \geq c_\sigma > 0$ for some constant c_σ and $\mathbb{E}[\sigma_t^s(\boldsymbol{\theta}_0)] < \infty$ for some $s > 0$. Moreover, for any $\theta \in \Theta$, we assume $\sigma_t(\boldsymbol{\theta}_0)/\sigma_t(\boldsymbol{\theta}) = 1$ almost surely if and only if $\boldsymbol{\theta} = \boldsymbol{\theta}_0$*

Assumptions (2) and 4 are satisfied by restricting the parameter space Θ for spacific forms of σ_t. For instance, in the GARCH(1,1), that is $\sigma_t^2 = \omega + \alpha y_{t-1}^2 + \beta\sigma_{t-1}^2$, case with parameters $\boldsymbol{\theta} = (\omega, \alpha, \beta)'$, we restrict $\omega \geq c_\sigma$, $\alpha > 0$, $\beta > 0$ and $\alpha + \beta < 1$.

Assumption 5 (Scaling Stability)**.** *There exists a function g such that for any $\boldsymbol{\theta} \in \Theta$, for any $\lambda > 0$, and any real sequence $\{x_i\}$*

$$\lambda\sigma(x_1, x_2, \ldots; \boldsymbol{\theta}) = \sigma(x_1, x_2, \ldots; \boldsymbol{\theta}_\lambda)$$

where $\theta_\lambda = g(\theta, \lambda)$ and g is differentiable in λ.

All commonly used volatility models satisfy this condition. Table 1 in (Francq and Zakoïan, 2015) lists garch-type models that satisfy this condition.

Let $\sigma_{t,r}(\boldsymbol{\theta}) = \sigma(y_t, ..., y_{t-r+1}, \tilde{y}_{t-r}, \tilde{y}_{t-r-1}, \ldots; \boldsymbol{\theta})$ for any sequence $\{\tilde{y}_k\}_{k \leq t-r}$ and $r \in \mathbb{N}$. Also, write $\partial_\theta \sigma_t(\boldsymbol{\theta}) = \frac{\partial \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$ and $\partial_{\theta\theta} \sigma_t(\boldsymbol{\theta}) = \frac{\partial^2 \sigma_t(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$.

Assumption 6 (Approximation)**.** *There exists a constant $\rho \in (0,1)$ and random variables C_r measurable with respect to $\mathcal{F}_{t-r}$ and $\mathbb{E}\left[|C_r|^s\right] < \infty$ for some $s > 0$ such that*

$$\sup_{\boldsymbol{\theta} \in \Theta} |\sigma_t(\boldsymbol{\theta}) - \sigma_{t,r}(\boldsymbol{\theta})| \leq C_r \rho^r$$

The function $\theta \to \sigma(x_1, x_2, \ldots; \theta)$ has continuous second-order derivatives satisfying

$$\sup_{\boldsymbol{\theta} \in \Theta} \left\| \partial_\theta \sigma_t(\boldsymbol{\theta}) - \partial_\theta \sigma_{t,r}(\boldsymbol{\theta}) \right\| + \left\| \partial_{\theta\theta} \sigma_t(\boldsymbol{\theta}) - \partial_{\theta\theta} \sigma_{t,r}(\boldsymbol{\theta}) \right\| \leq C_r \rho^r. \tag{26}$$

We require that the initial values of the volatility process are forgotten as we increase our observations. This is an approximation condition that is valid for all ARMA-type volatility processes with finite lags. For instance, if $\sigma_t = \sum_{i=1}^\infty a_i y_{t-i}^2$ admits a $AR(\infty)$ representation with coefficients $a_i \propto \rho^i$, then $|\sigma_t - \sigma_{t,r}| = \rho^{-r} C_r$ with $C_r = \sum_{i=0}^\infty \rho^r a_{i+r} y_{t-i+r}^2$, and $\mathbb{E}|C_r|^s < \infty$ if $\mathbb{E}|y_t|^{2s} < \infty$ and condition 2 holds. Typically, the coefficients a_i are functions of $\boldsymbol{\theta}$ and will still be geometrically decreasing as we differentiate. The same argument follows for the derivatives.

Assumption 7 (Non-degeneracy)**.** *There does not exist a non-zero $\boldsymbol{\lambda} \in \mathbb{R}^m$ such that,*

$$\lambda' \partial_\theta \sigma_t(\boldsymbol{\theta}_0) = 0, \quad a.s.$$

Assumption 8 (Monotonicity)**.** *For any real sequence $\{x_i\}$ and for any $\boldsymbol{\theta}_1, \boldsymbol{\theta}_2 \in \boldsymbol{\Theta}$ that satisfies $\boldsymbol{\theta}_1 \leq \boldsymbol{\theta}_2$ component-wise, we have $\sigma(x_1, x_2, \ldots; \boldsymbol{\theta}_1) \leq \sigma(x_1, x_2, \ldots; \boldsymbol{\theta}_2)$.*

Assumption 9 (Moments)**.** *There exists a ϵ-neighborhood V_ϵ^θ of $\boldsymbol{\theta}_0$ such that for some positive a, b, and c*

$$\mathbb{E}\left\{ \sup_{\boldsymbol{\theta} \in V_\epsilon^\theta} \left| \frac{\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta})} \right|^a + \sup_{\boldsymbol{\theta} \in V_\epsilon^\theta} \left\| \frac{\partial_\theta \sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})} \right\|^b + \sup_{\boldsymbol{\theta} \in V_\epsilon^\theta} \left\| \frac{\partial_{\theta\theta} \sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})} \right\|^c \right\} < \infty.$$

For particular choices of σ_t, this condition will be reflected by moment conditions on $\{y_t\}$.

We need to impose restrictions the instrumental density $g(y, \sigma) = \sigma^{-1} h(y/\sigma)$ used in the QML estimation.

Assumption 10. $\mathbb{E}g(\eta_0, \sigma) < \mathbb{E}g(\eta_0, 1)$*, for all $\sigma > 0$, $\sigma \neq 1$.*

As in (Francq and Zakoïan, 2015), this assumption, together with assumption **??**, ensures that

$$\lim_{n \to \infty} \tilde{G}_n(\boldsymbol{\theta}) = \mathbb{E}g(y_t, \sigma_t(\boldsymbol{\theta}) = \mathbb{E}\left\{ g(\eta_0, \sigma_t(\boldsymbol{\theta})/\sigma_t(\boldsymbol{\theta}_0)) - \log \sigma_t(\boldsymbol{\theta}_0) \right\},$$

is uniquely maximized at $\boldsymbol{\theta} = \boldsymbol{\theta}_0$. If $u \mapsto h(u)$ is differentiable, $\partial_\sigma \mathbb{E}g(\eta_0, \sigma) = 0$ if and only if $\sigma = 1$.

Assumption 11. *h is continuous on $\mathbb{R}$, twice differentiable except on a finite set A, and there exist finite constants $\delta > 0$ and c_h such that for all $u \in A^c$,*

$$|u \partial_u \log h(u)| + |u^2 \partial_{uu} \log h(u)| \leq c_h (1 + |u|^\delta)$$

with $\mathbb{E}|\eta_0|^\delta < \infty$.

This assumption is mild and is satisfied, for example, when h is the Gaussian kernel. More generally, if $h(u) \propto |u|^a \exp(c|u|^b)$, the inequality is satisfied with $\delta = b$.

Proofs o Section XX

Proposition 1 (Consistency of QMLE)**.** *Under Assumptions 1, ??, 3, 4, 5, 6 (of the volatility process), 10, and 11 (only first derivative), the QML estimador* $\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ *a.s. as* $n \to \infty$.

Proof. This result is parallel to Theorem 1 of (Francq and Zakoïan, 2015) and the proof is identical, replacing θ_0^* by $\boldsymbol{\theta}_0$. Their conditions are nested in our subset of assumptions for this proposition. $\square$

Lemma 2 (Approximation)**.** *Let* $\{x_t\}$ *be a stationary strong mixing process with mixing coefficients* $\{\alpha_m\}$ *and* $\mathcal{F}_s = \sigma\langle x_s, x_{s-1}, \ldots\rangle$. *Let* $f_{t,s} = f(y_t, \ldots, y_{t-s+1}, 0, 0, \ldots) \in \mathbb{R}$ *and* $f_t := f_{t,\infty}$ *be an* L_r*-integrable function, for some* $r > q > 0$*, satisfying for every* $s < t$, $|f_t - f_{t-s}| < C_{t,s}\rho^s$, *where* $\{C_{t,s}\}$ *is a sequence of* L_q*-integrable random variables and* $\rho \in (0,1)$. *Then* $\{f_t\}$ *is* L_q*-NED with respect to* $\mathcal{F}_t$ *and coefficients* $\gamma_m = c_1\rho^{m/2} + c_2\alpha_{m/2}^{1/q-1/r}$.

Proof. We have to show $\|f_t - \mathbb{E}_{t-m}f_t\|_q < \eta_m$, where $\|x\|_q = (\mathbb{E}|x|^q)^{1/q}$, $\mathbb{E}_s f_t = \mathbb{E}[f_t|\mathcal{F}_s]$ and $\mathcal{F}_s = \sigma\langle x_s, x_{s-1}, \ldots\rangle$. Observe that $\{f_{t,s}\}$ is strong mixing with coefficients $\tilde{\alpha}_m = \alpha_{(m-s)\vee 0}$. It follows from the triangle inequality that

$$\begin{aligned}
\|f_t - \mathbb{E}_{t-m}f_t\|_q &\le \|f_t - f_{t,m/2}\|_q + \|f_{t,m/2} - \mathbb{E}_{t-m}f_{t,m/2}\|_q + \|\mathbb{E}_{t-m}f_{t,m/2} - \mathbb{E}_{t-m}f_t\|_q \\
&\le 2\|C_{t,m/2}\|_q\rho^{m/2} + \|f_{t,m/2} - \mathbb{E}_{t-m}f_{t,m/2}\|_q \\
&\le 2\|C_{t,m/2}\|_q\rho^{m/2} + 2(\sqrt{2}+1)^{1/q}\alpha_{m/2}^{1/q-1/r}\|f_{t,m/2}\|_r,
\end{aligned}$$

where the last inequality follows from Theorem 13.XX in Davidson's book. The result follows with $c_1 = 2\|C_{t,m/2}\|_q$ and $c_2 = 2(\sqrt{2}+1)^{1/q}\|f_{t,m/2}\|_r$ $\square$

Proposition 2 (L_q-NED)**.** *Suppose Assumptions 2, 6 hold with* $s > q > 0$ *and 9 with* $a = b = c = s$. *Then, the processes* $\{\sup_{\boldsymbol{\theta}\in\Theta}|\sigma_t(\boldsymbol{\theta})|\}$, $\{\sup_{\boldsymbol{\theta}\in\Theta}\|\partial_\theta\sigma_t(\boldsymbol{\theta})\|\}$, *and* $\{\sup_{\boldsymbol{\theta}\in\Theta}\|\partial_{\theta\theta}\sigma_t(\boldsymbol{\theta})\|\}$ *are* L_p*-NED with respect to* $\mathcal{F}_t$ *with coefficients* $\gamma_m \propto e^{-cm}$, *for some* $c > 0$.

Proof. We show the result for $\{\sup_{\boldsymbol{\theta}\in\Theta}|\sigma_t(\boldsymbol{\theta})|\}$, and the remaining series follow after the exact same arguments.

By assumption 2, $\{y_t\}$ is a stationary geometric strong mixing process with coefficients $\alpha_m = \alpha^{-cm}$ for some $\alpha \in (0,1)$. Let $f_{t:r} = \sup_{\theta\in V_\epsilon^\theta}|\sigma_{t,r}(\boldsymbol{\theta})|$. By the inverse triangle inequality

$$\left|\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}|\sigma_t(\boldsymbol{\theta})| - \sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}|\sigma_{t,r}(\boldsymbol{\theta})|\right| \le \sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}|\sigma_t(\boldsymbol{\theta}) - \sigma_{t,r}(\boldsymbol{\theta})| \le C_r\rho^r,$$

where C_r is $\mathcal{F}_{t-r}$ measurable, L_r integrable random variable by assumption 6. From Lemma 2 it follows that $\{\sup_{\theta\in V_\epsilon^\theta}\sigma_t(\boldsymbol{\theta})\}$ is L_q-NED with respect to $\mathcal{F}_t$ with coefficients $\gamma_m = c_1\rho^{m/2} + c_2\alpha^{-c(1/q-1/r)m/2} \le c_3\tilde{\rho}^{-c_4m}$, where $\tilde{\rho} = \max(\rho, \alpha)$. $\square$

Theorem 3 (Consistency)**.** *Suppose Assumptions 1, 2, 3, 4 with* $s = 2$, *5, 6 with* $s = 3$, *9 with* $a = b = 3, c = 0$, *10, and 11 (only first derivative), hold. Then, the residual expectile estimator* $\widehat{XP}_{\tau,\hat{\boldsymbol{\theta}}_n} \to XP_\tau^\eta$ *in probability as* $n \to \infty$.

Proof. Under the assumptions of the theorem, the following is true:

1. $\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\| = o_p(1)$ from proposition 1;
2. $\inf_{\boldsymbol{\theta}\in V_\epsilon}\sigma_t(\boldsymbol{\theta}) > c_\sigma > 0$, where $V_\epsilon := \{\boldsymbol{\theta} \in \Theta : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < \epsilon\}$, for some $\epsilon > 0$, by assumption 4;

3. $\lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} \sup_{\boldsymbol{\theta}\in V_\epsilon} \left\|\frac{\partial}{\partial \boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta})\right\|(1+|y_t|) < \infty$, with probability converging to one, by assumption 2 and proposition 2 with $s = 3$.

In order to use Lemma 5.10 in Vaart, Asymptotic Statistics we have to show that (1) for each $\xi \in \mathbb{R}$, $Q_n(\hat{\boldsymbol{\theta}}_n;\xi) \to Q(\xi;\boldsymbol{\theta}_0) := \mathbb{E}(Q_n(\xi,\boldsymbol{\theta}_0))$ in probability, as $n \to \infty$; (2) $\xi \mapsto Q_n(\boldsymbol{\theta},\xi)$ is continuous; (3) $\xi = \widehat{\mathrm{XP}}_{\tau,\boldsymbol{\theta}}$ is the unique zero of $Q_n(\xi,\boldsymbol{\theta})$ for each $\boldsymbol{\theta}$; and (4) for every $\epsilon > 0$ $Q(\mathrm{XP}_\tau^{\boldsymbol{\theta}_0}-\epsilon;\boldsymbol{\theta}_0) < 0 < Q(\mathrm{XP}_\tau^{\boldsymbol{\theta}_0}+\epsilon;\boldsymbol{\theta}_0)$. Conditions (3) and (4) are already satisfied in (Holzmann and Klar, 2016; Krätschmer and Zähle, 2017), then we have to show (1) and (2).

For each $\epsilon > 0$, $\hat{\boldsymbol{\theta}} \in V_\epsilon$, $V_\epsilon := \{\boldsymbol{\theta} \in \Theta : \|\boldsymbol{\theta} - \boldsymbol{\theta}_0\| < \epsilon\}$, with probability converging to one. Hence, we constrain our parameter space to V_ϵ.

Let $\sigma_i = \sigma_t(\boldsymbol{\theta}_i)$ for $\boldsymbol{\theta}_i \in V_\epsilon$. A simple application of the mean value theorem shows that $\sigma_1/\sigma_2 = 1 + \sigma_2^{-1}\langle \partial_\theta \sigma_t(\tilde{\boldsymbol{\theta}}), \boldsymbol{\theta}_1 - \boldsymbol{\theta}_2\rangle$. By assumption, $\inf_{\boldsymbol{\theta}\in V_\epsilon} \sigma_t(\boldsymbol{\theta}) > c_\sigma$, so $|\sigma_1/\sigma_2 - 1| \leq (c_\sigma^{-1} \sup_{\boldsymbol{\theta}\in V_\epsilon} \|\partial\sigma_t(\boldsymbol{\theta})\|)\cdot\epsilon$. Now,

$$
\begin{aligned}
\psi_{t,\tau}(\boldsymbol{\theta}_1,\xi) - \psi_{t,\tau}(\boldsymbol{\theta}_2,\xi) &= (y_t/\sigma_1 - \xi)[|I(y_t/\sigma_1 - \xi \leq 0) - \tau| - |I(y_t/\sigma_1 - \xi \leq (\sigma_1/\sigma_2 - 1)\xi) - \tau|] \\
&\quad + \frac{y_t}{\sigma_1}(\sigma_1/\sigma_2 - 1)|I(y_t/\sigma_1 - \xi \leq (\sigma_1/\sigma_2 - 1)\xi) - \tau| \\
&\leq |\sigma_1/\sigma_2 - 1||[\xi + |y_t|/\sigma_1] \\
&\leq \left(c_\sigma^{-1} \sup_{\boldsymbol{\theta}\in V_\epsilon} \|\partial_\theta \sigma_t(\boldsymbol{\theta})\|\right)[\xi + c_\sigma^{-1}|y_t|]\cdot\epsilon.
\end{aligned}
$$

Finally, for any $\epsilon > 0$

$$
\lim_n \sup_{\boldsymbol{\theta}\in \mathrm{V}_\epsilon} |Q_n(\xi,\boldsymbol{\theta}) - Q_n(\xi,\boldsymbol{\theta}_0)| \leq \left\{\lim_n \frac{1}{n}\sum_{i=1}^{n}\left(c_\sigma^{-1} \sup_{\boldsymbol{\theta}\in V_\epsilon} \|\partial_\theta \sigma_t(\boldsymbol{\theta})\|\right)[\xi + c_\sigma^{-1}|y_t|]\right\}\cdot\epsilon.
$$

The limit in the bracket converges to its expectation by the ergodic theorem. Then, it follows that $|Q_n(\xi,\hat{\boldsymbol{\theta}}) - Q_n(\xi,\boldsymbol{\theta}_0)| = o_p(1)$. Now, $\psi_{t,\tau}(\xi;\boldsymbol{\theta}_0) = (\eta_t - \xi)|I(\eta_t \leq \xi) - \tau|$ is a function of i.i.d. random variables with finite mean, $Q_n(\xi,\boldsymbol{\theta}_0) \to \bar{Q}(\xi,\boldsymbol{\theta}_0)$ a.s., hence $Q_n(\xi,\hat{\boldsymbol{\theta}}) \to \bar{Q}(\xi,\boldsymbol{\theta}_0)$ in probability as $n \to \infty$.

Let $\delta > 0$ and $\boldsymbol{\theta}_1 \in \Theta$ be arbitrary. Then

$$
\begin{aligned}
\psi_{t,\tau}(\boldsymbol{\theta}_1,\xi+\delta) - \psi_{t,\tau}(\boldsymbol{\theta}_1,\xi+\delta) &\leq (y_t/\sigma_1 - \xi)[|I(y_t/\sigma_1 \leq \xi) - \tau| - |I(y_t/\sigma_1 \leq \xi+\delta) - \tau|] \\
&\quad + \delta|I(y_t/\sigma_1 \leq \xi+\delta) - \tau| \\
&\leq 2|\delta|,
\end{aligned}
$$

so that $Q_n(\xi,\boldsymbol{\theta}) = n^{-1}\sum_{i=1}\psi_{t,\tau}(\boldsymbol{\theta},\xi)$ is a uniformly continuous function of ξ, satisfying (2). □

Let $g_1(y;\sigma) = \frac{\partial}{\partial\sigma}\ln h(y/\sigma)/\sigma$, where h is the instrumental density in the QML estimation satisfying Assumptions 10 and 11. In addition, let $\partial_{\boldsymbol{\theta}}\sigma_T(\boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}}\sigma_T(\boldsymbol{\theta})$, $g_2(y;\sigma) = \frac{\partial}{\partial\sigma}g_1(y;\sigma)$.

Theorem 4 (Asymptotic distribution)**.** *Let Assumptions1 to 11 in Appendix A hold for any fixed* $0 < \underline{\tau} < \tau < \overline{\tau} < 1$. *It follows that*

$$
\sqrt{T}\left(\widehat{XP}_{\tau,\hat{\boldsymbol{\theta}}_n} - XP_\tau^\eta\right) \xrightarrow{d} \dot{\Psi}_{XP}(XP_\tau^\eta)^{-1}\mathbb{Z}_{XP} + XP_\tau^\eta\, J(\boldsymbol{\theta}_0)\mathbb{Z}_{\boldsymbol{\theta}},
$$

where $\mathbb{Z}_{\boldsymbol{\theta}}$ *and* $\mathbb{Z}_{XP}$ *are jointly zero-mean Gaussian random variables,*

$$
\dot{\Psi}_{XP}(XP_\tau^\eta) = \tau\big[1 - F_\eta(XP_\tau^\eta)\big] + (1-\tau)F_\eta(XP_\tau^\eta)
$$

is a positive scalar and $J(\boldsymbol{\theta}_0) = \mathbb{E}\left[\frac{\partial_{\boldsymbol{\theta}}\sigma_n(\boldsymbol{\theta}_0)}{\sigma_n(\boldsymbol{\theta}_0)}\right]$ *is nonzero.*

The covariance matrix of $\mathbb{Z}_{\boldsymbol{\theta}}$ *is* $\Sigma_{\boldsymbol{\theta}} = 4\tau_h^2 I(\boldsymbol{\theta}_0)^{-1}$, *with*

$$I(\boldsymbol{\theta}_0) = \mathbb{E}\left(\frac{\partial_{\boldsymbol{\theta}}\sigma_n(\boldsymbol{\theta}_0)}{\sigma_n^2(\boldsymbol{\theta}_0)}\frac{\partial_{\boldsymbol{\theta}'}\sigma_n(\boldsymbol{\theta}_0)}{\sigma_n^2(\boldsymbol{\theta}_0)}\right) \quad \text{and} \quad \tau_h^2 = \frac{\mathbb{E}[g_1^2(\eta_1, 1)]}{\{\mathbb{E}[g_2(\eta_1, 1)]\}^2},$$

whereas the variance of $\mathbb{Z}_{XP}$ *is* $\sigma^2_{XP_\tau} = \mathbb{E}\left[(\eta_1 - XP_\tau^\eta)^2(\mathbf{1}(\eta_1 < XP_\tau^\eta) - \tau)^2\right]$ *and their covariance is* $\Sigma_{XP,\boldsymbol{\theta}} = \mathbb{E}\left[\partial_{\boldsymbol{\theta}}\sigma_T(\boldsymbol{\theta}_0)\, g_1(\eta_1; 1)\dot{\rho}_\tau(\eta_1 - XP_\tau^\eta)\right]$.

Proof. The result follows after the following steps:

i. $\|\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0\| = o_p(1)$ and $\|\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} - \mathrm{XP}_\tau^\eta\| = o_p(1)$;

ii. stochastic equicontinuity of $\sqrt{n}(Q_n - Q)(\cdot; \boldsymbol{\theta}_0)$;

iii. $\sqrt{n}[(Q_n - Q)(\xi, \hat{\boldsymbol{\theta}}_n) - (Q_n - Q)(\xi, \boldsymbol{\theta}_0)] = o_p(\|\|\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)\|)$, for $|\xi - \mathrm{XP}_\tau^\eta| < \epsilon$;

iv. as $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} \to \mathrm{XP}_\tau^\eta$ in probability

$$Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) - Q(\mathrm{XP}_\tau^\eta, \boldsymbol{\theta}_0) = [Q_\xi(\mathrm{XP}_\tau^\eta, \boldsymbol{\theta}_0) + o_p(1)](\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} - \mathrm{XP}_\tau^\eta);$$

v. as $\hat{\boldsymbol{\theta}}_n \to \boldsymbol{\theta}_0$ and $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} \to \mathrm{XP}_\tau^\eta$, in probability

$$Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \hat{\boldsymbol{\theta}}_n) - Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) = [Q_\theta(\mathrm{XP}_\tau^\eta, \boldsymbol{\theta}_0) + o_p(1)](\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0);$$

vi. a central limit theorem for $(\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0), \sqrt{n}Q_n(\mathrm{XP}_\tau^\eta, \boldsymbol{\theta}_0))$.

Step (i): follows directly from Theorem 1 and proposition 1.
Step (ii):

In order to show stochastic equicontinuity in step (ii), we have to satisfy Pollard's entropy condition for the functions in display. We use results in (Andrews, 1994) and (Hansen, 1996) to simplify our task, i.e., the entropy condition will be replaced by a Lipschitz condition on $\psi_{t,\tau}$ and bound moment conditions on the Lipschitz coefficients and $\psi_{t,\tau}$.

Under Assumptions A, B and C in (Andrews, 1994), $\{\sqrt{n}(Q_n - Q)(\cdot; \boldsymbol{\theta}_0)\}$ is stochastically equicontinuous. Recall from the proof of Theorem 1 that $\psi_{t,\tau}(\boldsymbol{\theta}_0, \cdot)$ is Lipschitz, i.e., for $\delta > 0$ and any $\boldsymbol{\theta} \in \Theta$,

$$|\psi_{t,\tau}(\boldsymbol{\theta}, \xi + \delta) - \psi_{t,\tau}(\boldsymbol{\theta}, \xi)| \leq 2|\delta|.$$

It follows from Theorem 2 in (Andrews, 1994) that Pollard's condition is satisfied with envelope function $\sup_{\xi \in V_\epsilon^{\mathrm{XP}_\tau}} |\psi(\boldsymbol{\theta}_0, \xi)| \vee 2$, satisfying assumption A. Assumption B is satisfied because $\mathbb{E}(\sup_{\xi \in V_\epsilon^{\mathrm{XP}_\tau}} |\psi(\boldsymbol{\theta}_0, \xi)| \vee 2)^3 \leq \mathbb{E}(|\eta_t| + |\mathrm{XP}_\tau^\eta| + 3)^3 < \infty$ by assumption 1. Condition C holds because at $\boldsymbol{\theta}_0$, Q_n is a function of the independent innovations $\{\eta_t\}$.
Step (iii):

For any fixed ξ, and suppose without loss of generality that $\tau > 1/2$,

$$\begin{aligned}
\psi_{t,\tau}(\boldsymbol{\theta}, \xi) - \psi_{t,\tau}(\boldsymbol{\theta}_0, \xi) &= (1-2\tau)(y_t - \xi) I\left\{0 \leq y_t - \xi \leq (A_t(\boldsymbol{\theta}) - 1)\xi\right\} \\
&\quad + (2\tau - 1)(y_t - \xi) I\left\{(A_t(\boldsymbol{\theta}) - 1)\xi \leq y_t - \xi < 0\right\} \\
&+ (A_t(\boldsymbol{\theta}) - 1)\eta_t \left|I(\eta_t < \xi) - \tau\right| \\
&= (2\tau - 1)|y_t - \xi| I\left\{0 \wedge (A_t(\boldsymbol{\theta}) - 1)\xi \leq y_t - \xi \leq 0 \vee (A_t(\boldsymbol{\theta}) - 1)\xi\right\} \\
&+ (A_t(\boldsymbol{\theta}) - 1)\eta_t \left|I(\eta_t < \xi) - \tau\right| \\
&= |A_t(\boldsymbol{\theta}) - 1| \underbrace{(2\tau - 1)\left|\frac{y_t - \xi}{A_t(\boldsymbol{\theta}) - 1}\right| I\left\{0 \wedge \xi \leq \frac{y_t - \xi}{A_t(\boldsymbol{\theta}) - 1} \leq 0 \vee \xi\right\}}_{b_{1,t}} \\
&+ (A_t(\boldsymbol{\theta}) - 1) \underbrace{\eta_t \left|I(\eta_t < \xi) - \tau\right|}_{b_{2,t}}.
\end{aligned}$$

Also, follows from the mean value theorem that for some $\bar{\boldsymbol{\theta}}$ between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_0$, $A_t(\boldsymbol{\theta}) - 1 = \partial_\theta A_t(\bar{\boldsymbol{\theta}})'(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$. As $\partial_\theta A_t(\bar{\boldsymbol{\theta}})(b_{1,t} + b_{2,t})$ is a measurable, integrable function of $\{y_t\}$, it follows from the ergodic theorem that

$$\sup_{\xi \in V_\epsilon^{\mathrm{XP}_\tau}} \sqrt{n}[(Q_n - Q)(\xi, \hat{\boldsymbol{\theta}}_n) - (Q_n - Q)(\xi, \boldsymbol{\theta}_0)] = o_p(\|\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)\|),$$

for each $\xi \in V_\epsilon^{\mathrm{XP}_\tau}$.

Step (iv):

The function Q_n is discontinuous, but its expectation Q is not. Let $A_t(\boldsymbol{\theta}) = \sigma_t(\boldsymbol{\theta}_0)/\sigma_t(\boldsymbol{\theta})$, so that $\psi_{t,\tau}(\boldsymbol{\theta}, \xi) = (A_t(\boldsymbol{\theta})\eta_t - \xi)|I(A_t(\boldsymbol{\theta})\eta_t \leq \xi) - \tau|$. Let $\mathcal{F}_{t-1}$ be the σ-algebra generated by $(\eta_{t-1}, \eta_{t-2}, ...)$, then

$$\begin{aligned}
\mathbb{E}\left[\psi_{t,\tau}(\boldsymbol{\theta}, \xi)\middle|\mathcal{F}_{t-1}\right] &= A_t(\theta)\mathbb{E}\left[(\eta_t - A_t^{-1}(\boldsymbol{\theta})\xi)|I(\eta_t \leq A_t^{-1}(\boldsymbol{\theta})\xi) - \tau|\middle|\mathcal{F}_{t-1}\right] \\
&= A_t(\boldsymbol{\theta})\mathbb{E}\left[\psi_{t,\tau}(\theta_0, A_t^{-1}(\boldsymbol{\theta})\xi)\middle|\mathcal{F}_{t-1}\right] \\
&= A_t(\boldsymbol{\theta})\left\{\tau \int_{A_t^{-1}(\boldsymbol{\theta})\xi}^{\infty} (1 - F_\eta(x))dx - (1-\tau)\int_{-\infty}^{A_t^{-1}(\boldsymbol{\theta})\xi} F_\eta(x)dx\right\}. \qquad (27)
\end{aligned}$$

At $\boldsymbol{\theta} = \boldsymbol{\theta}_0$, $A_t(\boldsymbol{\theta}_0) = 1$ and $\mathbb{E}[\psi_{t,\tau}(\boldsymbol{\theta}_0, \xi)] = \tau\int_\xi^\infty (1 - F_\eta(x))dx - (1-\tau)\int_{-\infty}^\xi F_\eta(x)dx$, which is a continuously differentiable function of ξ. Applying the mean value theorem, there is some ξ between $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}$ and XP_τ^η satisfying

$$Q(\mathrm{XP}_\tau^\eta, \boldsymbol{\theta}_0) - Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) = Q_\xi(\xi, \boldsymbol{\theta}_0)(\mathrm{XP}_\tau^\eta - \widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}), \qquad (28)$$

where $Q_\xi(\xi, \boldsymbol{\theta}_0) = \tau(1 - F_\eta(\xi)) + (1-\tau)F_\eta(\xi)$. By continuity of $F_\eta(x)$ around XP_τ^η, $Q_\xi(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) = Q_\xi(\mathrm{XP}_\tau^\eta, \boldsymbol{\theta}_0) + o_p(1)$ as $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} \to \mathrm{XP}_\tau^\eta$ in probability, as requested.

Step (v):

Similarly, using the mean value theorem on $Q(\xi, \cdot)$, there is some $\tilde{\boldsymbol{\theta}}$ between $\hat{\boldsymbol{\theta}}_n$ and $\boldsymbol{\theta}_0$ such that

$$Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \hat{\boldsymbol{\theta}}_n) - Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) = Q_\theta(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \tilde{\boldsymbol{\theta}})'(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0),$$

where $Q_\theta(\xi\,\boldsymbol{\theta}) = \partial Q(\xi,\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ is

$$\begin{aligned}
\frac{\partial Q(\xi,\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} &= \mathbb{E}\left[\frac{\partial}{\partial\boldsymbol{\theta}}A_t(\boldsymbol{\theta})\mathbb{E}\left\{\psi_{t,\tau}(\boldsymbol{\theta}_0, A_t(\boldsymbol{\theta})^{-1}\xi)|\mathcal{F}_{t-1}\right\}\right] \\
&+ \mathbb{E}\left[A_t(\boldsymbol{\theta})\frac{\partial}{\partial\boldsymbol{\theta}}A_t(\boldsymbol{\theta})^{-1}\tau(1-F_\eta(A_t(\boldsymbol{\theta})^{-1}\xi)) + (1-\tau)F_\eta(A_t(\boldsymbol{\theta})^{-1}\xi)\right] \\
&= -\mathbb{E}\left[A_t(\boldsymbol{\theta})\frac{\partial_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\mathbb{E}\left\{\psi_{t,\tau}(\boldsymbol{\theta}_0, A_t(\boldsymbol{\theta})^{-1}\xi)|\mathcal{F}_{t-1}\right\}\right] \\
&+ \mathbb{E}\left[\frac{\partial_{\boldsymbol{\theta}}\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\xi\left\{\tau(1-F_\eta(A_t(\boldsymbol{\theta})^{-1}\xi)) + (1-\tau)F_\eta(A_t(\boldsymbol{\theta})^{-1}\xi)\right\}\right].
\end{aligned}$$

We will show that Q_θ is a bounded continuous function of $\boldsymbol{\theta}$ and ξ for $\boldsymbol{\theta}\in V_\epsilon^\theta$ and $\xi\in V_\epsilon^\xi$. It follows from (Holzmann and Klar, 2016, Proposition 1) that $A_t(\boldsymbol{\theta})^{-1}\mathrm{XP}_\tau^\eta$ is the unique zero of (27). Then, for $\xi\in V_\epsilon^\xi$, it follows from the mean value theorem that

$$\begin{aligned}
|\mathbb{E}[\psi_{t,\tau}(\boldsymbol{\theta},\xi)|\mathcal{F}_{t-1}]| &= |A_t(\theta)\left[\mathbb{E}\left\{\psi_{t,\tau}(\boldsymbol{\theta}_0, A_t(\boldsymbol{\theta})^{-1}\xi)|\mathcal{F}_{t-1}\right\} - \mathbb{E}\left\{\psi_{t,\tau}(\boldsymbol{\theta}_0, A_t(\boldsymbol{\theta})^{-1}\mathrm{XP}_\tau^\eta)|\mathcal{F}_{t-1}\right\}\right]| \\
&= \left|\left[\tau(1-F_\eta(A_t(\boldsymbol{\theta})^{-1}\tilde{\xi})) + (1-\tau)F_\eta(A_t(\boldsymbol{\theta})^{-1}\tilde{\xi})(\xi-\mathrm{XP}_\tau^\eta)\right]\right| \\
&\le |\xi-\mathrm{XP}_\tau^\eta| < \epsilon.
\end{aligned}$$

Then, $\|Q_\theta(\xi,\boldsymbol{\theta})\| \lesssim \mathbb{E}\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|^2 < \infty$. To show continuity, let $B_1(\xi,\boldsymbol{\theta}) = \mathbb{E}[\psi_{t,\tau}(\boldsymbol{\theta}_0, A_t(\boldsymbol{\theta})^{-1}\xi)|\mathcal{F}_{t-1}]$ and $B_2(\xi,\boldsymbol{\theta}) = \tau(1-F_\eta(A_t(\boldsymbol{\theta})^{-1}\xi)) + (1-\tau)F_\eta(A_t(\boldsymbol{\theta})^{-1}\xi)$. Let $\boldsymbol{\theta}_1,\boldsymbol{\theta}_2\in V_\epsilon^\theta$ and $\xi_1,\xi_2\in V_\epsilon^\xi$ be such that $|\xi_1-\xi_2| + \|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\| < \delta$, then

$$\begin{aligned}
\|Q_\theta(\xi_1,\boldsymbol{\theta}_1) - Q_\theta(\xi_2,\boldsymbol{\theta}_2)\| &\le \left\|\mathbb{E}\left[\left\{\frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_1)}{\sigma_t(\boldsymbol{\theta}_1)} - \frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_2)}{\sigma_t(\boldsymbol{\theta}_2)}\right\}(\xi_1 B_2(\xi_1,\boldsymbol{\theta}_1) - A_t(\boldsymbol{\theta}_1)B_1(\xi_1,\boldsymbol{\theta}_1)\right]\right\| \\
&+ \left\|\mathbb{E}\left[\frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_2)}{\sigma_t(\boldsymbol{\theta}_2)}\left\{\xi_1 B_2(\xi_1,\boldsymbol{\theta}_1) - \xi_2 B_2(\xi_2,\boldsymbol{\theta}_2)\right\}\right]\right\| \\
&+ \left\|\mathbb{E}\left[\frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_2)}{\sigma_t(\boldsymbol{\theta}_2)}A_t(\boldsymbol{\theta}_2)\left\{B_1(\xi_1,\boldsymbol{\theta}_1) - B_1(\xi_2,\boldsymbol{\theta}_2)\right\}\right]\right\| \\
&+ \left\|\mathbb{E}\left[\frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_2)}{\sigma_t(\boldsymbol{\theta}_2)}B_1(\xi_1,\boldsymbol{\theta}_1)\left\{A_t(\boldsymbol{\theta}_1) - A_t(\boldsymbol{\theta}_2)\right\}\right]\right\| \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

We show that $I_i \lesssim \delta$, $i = 1,..,4$. Starting with I_1, we have shown that $B_1(\xi,\boldsymbol{\theta})$ and $B_2(\xi,\boldsymbol{\theta})$ are uniformly bounded inside V_ϵ^ξ, then

$$I_1 \lesssim \mathbb{E}\left[\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_{\theta\theta}\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\| + \sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\|\partial_\theta\sigma_t(\boldsymbol{\theta})\|\right]\cdot\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\| \lesssim \delta. \tag{29}$$

In I_2, $|\xi_1 B_1(\xi_1\boldsymbol{\theta}_1) - \xi_2 B_1(\xi_2\boldsymbol{\theta}_2)| \le |\xi_1-\xi_2|$, then

$$I_2 \le \mathbb{E}\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|\cdot|\xi_1-\xi_2| \lesssim \delta. \tag{30}$$

Now, $B_1(\xi_1,\boldsymbol{\theta}_1) - B_1(\xi_2,\boldsymbol{\theta}_2) \leq A_t(\boldsymbol{\theta}_1)^{-1}\xi_1 - A_t(\boldsymbol{\theta}_2)^{-1}\xi_2$ so that

$$\begin{aligned}
I_3 &\leq \left\|\mathbb{E}\left[\frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_2)}{\sigma_t(\boldsymbol{\theta}_2)}A_t(\boldsymbol{\theta}_2)\left\{A_t(\boldsymbol{\theta}_1)^{-1}\xi_1 - A_t(\boldsymbol{\theta}_2)^{-1}\xi_2\right\}\right]\right\| \\
&\leq \mathbb{E}\left[\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|\left\{\left|\frac{\sigma_t(\boldsymbol{\theta}_1)}{\sigma_t(\boldsymbol{\theta}_2)} - 1\right||\xi_1| + |\xi_1-\xi_2|\right\}\right] \\
&\leq \left[\mathbb{E}\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|^2 \mathbb{E}\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\|\partial_\theta\sigma_t(\boldsymbol{\theta})\|^2\right]^{1/2}\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\|\frac{|\xi_1|}{c_\sigma} + \mathbb{E}\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\||\xi_1-\xi_2| \qquad (31)\\
&\lesssim \delta.
\end{aligned}$$

Finally by the mean value theorem $A_t(\boldsymbol{\theta}_1) - A_t(\boldsymbol{\theta}_2) = -A_t(\tilde{\boldsymbol{\theta}})\frac{\partial_\theta\sigma_t(\tilde{\boldsymbol{\theta}})}{\sigma_t(\theta)}'(\boldsymbol{\theta}_1-\boldsymbol{\theta}_2)$, for some $\tilde{\boldsymbol{\theta}}$ between $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$, then

$$\begin{aligned}
I_4 &\leq \mathbb{E}\left[\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|^2|B_1(\xi_1,\boldsymbol{\theta}_1)A_t(\tilde{\boldsymbol{\theta}})|\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\|\right] \\
&\leq \mathbb{E}\left[\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|^2\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}|\sigma_t(\boldsymbol{\theta})|\right]c_\sigma^{-1}\epsilon\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\| \\
&\leq \mathbb{E}\left[\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}\left\|\frac{\partial_\theta\sigma_t(\boldsymbol{\theta})}{\sigma_t(\boldsymbol{\theta})}\right\|^3\right]^{2/3}\left[\mathbb{E}\sup_{\boldsymbol{\theta}\in V_\epsilon^\theta}|\sigma_t(\boldsymbol{\theta})|^3\right]^{1/3}\|\boldsymbol{\theta}_1-\boldsymbol{\theta}_2\| \lesssim \delta \qquad (32)
\end{aligned}$$

Combining (29)–(32), we conclude that $|Q_\theta(\xi_1,\boldsymbol{\theta}_1) - Q_\theta(\xi_2,\boldsymbol{\theta}_2)| < \varepsilon$ for some ε. Then Q_θ is a continuous function of ξ and $\boldsymbol{\theta}$. Convergence follows from the continuous mapping theorem.

Step (vi):

Replace $\boldsymbol{\theta}_0^*$ by $\boldsymbol{\theta}_0$ in Theorem 2 of (Francq and Zakoïan, 2015), which uses a Taylor expansion of the QMLE criterion function around $\boldsymbol{\theta}_0$. After isolating the dominant term, we obtain

$$\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0) = H(\theta_0)^{-1}G_n(\boldsymbol{\theta}_0) + o_p(1),$$

with $H(\boldsymbol{\theta}_0) = \mathbb{E}[g_2(\eta_0,1)]^2 I(\boldsymbol{\theta}_0)/4$ a positive definite matrix and $G_n(\boldsymbol{\theta}) = n^{-1/2}\sum_{t=1}^n \frac{\partial}{\partial\boldsymbol{\theta}}g(y_t;\sigma_t(\boldsymbol{\theta}_0))$. Here $g(x;\sigma) = \log\sigma^{-1}h(x/\sigma)$, $g_1(x,\sigma) = \partial g(x;\sigma)/\partial\sigma = \partial_\sigma g(x;\sigma)$ and $g_2(x;\sigma) = \partial g_1(x;\sigma)/\partial\sigma$, and $I(\boldsymbol{\theta}) = [\{\partial_\theta\sigma_t(\boldsymbol{\theta})/\sigma_t(\boldsymbol{\theta})\}\{\partial_\theta\sigma_t(\boldsymbol{\theta})/\sigma_t(\boldsymbol{\theta})\}']$ is nonsingular for $\boldsymbol{\theta} \in V_\epsilon^\theta$. Moreover, $G_n(\boldsymbol{\theta}_0) \Rightarrow N(0;\mathbb{E}[g_1^2(\eta_0,1)]I(\boldsymbol{\theta}_0)/4)$.

For any $\alpha := (\alpha_1,\boldsymbol{\alpha}_2) \in \mathbb{R}^{m+1}$ write

$$\begin{aligned}
Z(\boldsymbol{\alpha}) &= \alpha_1[\sqrt{n}Q_n(\xi_0,\boldsymbol{\theta}_0)] + \boldsymbol{\alpha}_2'[G_n(\boldsymbol{\theta}_0)] \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n \alpha_1\psi_{t,\tau}(\boldsymbol{\theta}_0,\xi_0) + g_1(\eta_t;1)\frac{\boldsymbol{\alpha}_2'\partial_\theta\sigma_t(\theta_0)}{\sigma_t(\boldsymbol{\theta}_0)} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^n z_t(\boldsymbol{\alpha}).
\end{aligned}$$

The sequence $z_t(\boldsymbol{\alpha})$ is a strictly stationary, ergodic, martingale difference with respect to filtration $\mathcal{F}_t$, therefore the central limit theorem for martingales holds and $Z(\boldsymbol{\alpha}) \Rightarrow N(0,\mathbb{E}(z_1(\boldsymbol{\alpha})^2)$. It follows from our identification assumption 10 that $\mathbb{E}[\eta_1 h'(\eta_1)/h(\eta_1)] = -1$ and $\mathbb{E}[\psi_{t,\tau}(\boldsymbol{\theta}_0,\xi_0)|\mathcal{F}_{t-1}] = 0$. Then, for all $\boldsymbol{\alpha}$,

$$\mathbb{E}[z_t(\boldsymbol{\alpha})|F_{t-1}] = \alpha_1\mathbb{E}[\psi_{t,\tau}(\boldsymbol{\theta}_0,\xi_0)|\mathcal{F}_{t-1}] - \frac{\boldsymbol{\alpha}_2'\partial_\theta\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta}_0)}\mathbb{E}\left[\eta_t\frac{h'(\eta_t)}{h(\eta_t)} + 1\right] = 0.$$

The variance of $z_t(\boldsymbol{\alpha})$ is

$$\mathbb{E}(z_t(\boldsymbol{\alpha})^2) = \alpha_1^2 \sigma^2_{\mathrm{XP}_\tau} + \boldsymbol{\alpha}_2' I(\boldsymbol{\theta}0)\boldsymbol{\alpha}_2 \frac{\mathbb{E}[g_1(\eta_0,1)^2]}{4} + 2\alpha_1 \boldsymbol{\alpha}_2' \Sigma_{\mathrm{XP}^\eta_\tau,\boldsymbol{\theta}},$$

with $\Sigma_{\mathrm{XP}^\eta_\tau,\boldsymbol{\theta}} := \mathbb{E}[\sigma_t(\boldsymbol{\theta}_0)^{-1}\partial_\theta \sigma_t(\boldsymbol{\theta}_0) g_1(\eta_0,1)\psi_{1,\tau}(\boldsymbol{\theta}_0;\xi_0)]$ and $\sigma^2_{\mathrm{XP}_\tau} := \mathbb{E}[\psi_{t,\tau}(\boldsymbol{\theta}_0,\xi_0)^2|\mathcal{F}_{t-1}] = \mathbb{E}[(\eta_0-\xi_0)^2|I(\eta_0<\xi_0)-\tau|^2]$.

The central limit theorem follows from the Cramer-Wold device, that is, $(\sqrt{n}Q_n(\mathrm{XP}^\eta_\tau,\boldsymbol{\theta}_0), \sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}_0)') \Rightarrow (\mathbb{Z}_{\mathrm{XP}_\tau}, \mathbb{Z}_\theta')$. The random variables $\mathbb{Z}_{\mathrm{XP}_\tau}$ and $\mathbb{Z}_\theta$ are jointly Gaussian with zero mean and covariance matrix

$$\Sigma = \begin{pmatrix} \sigma^2_{\mathrm{XP}_\tau} & \sigma'_{\mathrm{XP}_\tau,\theta} \\ \sigma_{\mathrm{XP}_\tau,\theta} & \Sigma_\theta \end{pmatrix} = \begin{pmatrix} \mathbb{E}[(\eta_0-\xi_0)^2|I(\eta_0<\xi_0)-\tau|^2] & \cdot \\ \mathbb{E}[\sigma_t(\boldsymbol{\theta}_0)^{-1}\partial_\theta\sigma_t(\boldsymbol{\theta}_0)g_1(\eta_0,1)\psi_{1,\tau}(\boldsymbol{\theta}_0;\xi_0)] & 4\frac{\mathbb{E}[g_1(\eta_0,1)^2]}{[\mathbb{E}\{g_2(\eta_0,1)\}]^2} I(\boldsymbol{\theta}_0)^{-1} \end{pmatrix}.$$

Combining steps (i)–(vi):
The solution $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}$ and $\hat{\boldsymbol{\theta}}_n$ solves

$$\begin{aligned}
0 &= Q_n(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \hat{\boldsymbol{\theta}}_n) - Q(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0) \\
&= [\sqrt{n}(Q_n - Q)(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \hat{\boldsymbol{\theta}}_n) - \sqrt{n}(Q_n - Q)(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0)] + Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \hat{\boldsymbol{\theta}}_n) - Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) \\
&\quad - [\sqrt{n}(Q_n - Q)(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) - \sqrt{n}(Q_n - Q)(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0)] - Q(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) - Q(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0) \\
&\quad + \sqrt{n}Q_n(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta_0}) \\
&= Q(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0) + Q_\theta(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0)\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta_0}) + o_p(1) \\
&\quad - \{Q_\xi(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta_0}) + o_p(1)\}\sqrt{n}(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} - \mathrm{XP}^\eta_\tau).
\end{aligned}$$

Consistency of $\hat{\boldsymbol{\theta}}_n$ and $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}$ imply that

$$\sqrt{n}(Q_n - Q)(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \hat{\boldsymbol{\theta}}_n) - \sqrt{n}(Q_n - Q)(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) = o_p(1).$$

Under the same arguments used in (Andrews, 1994), equation (3.36), stochastic equicontinuity of $\{\sqrt{n}(Q_n - Q)(\cdot;\boldsymbol{\theta}_0) : \xi \in V_\epsilon^{\mathrm{XP}_\tau}\}$ and consistency of $\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}$,

$$\sqrt{n}(Q_n - Q)(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n}, \boldsymbol{\theta}_0) - \sqrt{n}(Q_n - Q)(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0) = o_p(1).$$

Then,

$$\begin{aligned}
\sqrt{n}(\widehat{\mathrm{XP}}_{\tau,\hat{\boldsymbol{\theta}}_n} - \mathrm{XP}^\eta_\tau) &= Q_\xi(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta_0})^{-1}\left\{Q_n(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0) + Q_\theta(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0)\sqrt{n}(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta_0})\right\} + o_p(1) \\
&\Rightarrow Q_\xi(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta_0})^{-1}\left\{\mathbb{Z}_{\mathrm{XP}_\tau} + Q_\theta(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0)\mathbb{Z}_\theta\right\}.
\end{aligned}$$

Finally,

$$Q_\xi(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta_0})^{-1} = \tau\big[1 - F_\eta(\mathrm{XP}^\eta_\tau)\big] + (1-\tau)F_\eta(\mathrm{XP}^\eta_\tau) =: \dot{\Psi}_{\mathrm{XP}_\tau}(\mathrm{XP}^\eta_\tau),$$

and

$$\begin{aligned}
Q_\xi(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta_0})^{-1}Q_\theta(\mathrm{XP}^\eta_\tau, \boldsymbol{\theta}_0) &= \dot{\Psi}_{\mathrm{XP}_\tau}(\mathrm{XP}^\eta_\tau)^{-1}\frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta}_0)}\mathrm{XP}^\eta_\tau\,\dot{\Psi}_{\mathrm{XP}_\tau}(\mathrm{XP}^\eta_\tau) \\
&= \frac{\partial_\theta\sigma_t(\boldsymbol{\theta}_0)}{\sigma_t(\boldsymbol{\theta}_0)}\mathrm{XP}^\eta_\tau.
\end{aligned}$$

□