Supplementary material for "On efficient estimation for Value-at-Risk via location-scale time series models"

Chaoxu Lei and Qianqian Zhu

School of Statistics and Management

Shanghai University of Finance and Economics

Supplementary Material

Abstract: This supplementary material includes theoretical properties for the parametric CQR under mis-specification in Section S1. Section S2 compares the proposed two CQR estimators and compares them with the Gaussian QMLE (GQMLE) and exponential QMLE (EQMLE) for model (2.1). The details for calculating AREs of Examples 1-3 are summarized in Section S3. Moreover, the proposed CQRs for DAR and NAR-GARCH type models are illustrated in Section S4. Technical details for Theorems 1-4 and Corollary S1 are provided in Section S5. Section S6 proves that Theorems 1-4 still hold for both CQR estimators in ARMA-GARCH, ALDAR and ESTAR-GARCH models under some regular conditions. Section S7 establishes the selection consistency with proof for the BIC proposed in Section 3.2 of the manuscript. In addition, Sections S8-S9 present additional results for simulation studies and empirical analysis. To show Theorems 1-2 and Corollary S1, Lemmas 1-7 are introduced with proofs. Throughout the supplement, the notation C is a generic constant that may take different values from line to line, and $\rho \in (0,1)$ is a generic constant that may take different values in different locations. $\to_p$ and $\to_{\mathcal{L}}$ denote the convergences in probability and in distribution, respectively, and $o_p(1)$ denotes a sequence of random variables converging to zero in probability. Moreover, the norm of a matrix or column vector is defined as $\|A\| = \sqrt{tr(AA')} = \sqrt{\sum_{i,j} a_{ij}^2}$.

S1 The parametric CQR under mis-specification

Note that Theorems 3-4 are established under the situation that the quantile function $Q_\tau(\boldsymbol{\lambda})$ is correctly specified for the innovation η_t; see Assumption 4(ii). If $Q_\tau(\boldsymbol{\lambda})$ is mis-specified, then the conditional quantile function $g_{t,\tau}(\boldsymbol{\psi})$ in (2.8) is a working model, and the resulting parametric CQR estimator $\widehat{\boldsymbol{\psi}}_n$ in (2.9) will be asymptotically biased. To establish the asymptotic properties for $\widehat{\boldsymbol{\psi}}_n$ under mis-specification, a pseudo-true parameter vector needs to be defined for $\widehat{\boldsymbol{\psi}}_n$ based on the working model, and Assumptions 1, 2(ii) and 4(ii) should be revised accordingly. Specifically, the pseudo-true parameter vector is defined as $\boldsymbol{\psi}_0^* = (\boldsymbol{\vartheta}_0^{*\prime}, \boldsymbol{\lambda}_0^{*\prime})' = \arg\min_{\psi\in\Psi}\sum_{k=1}^K E[\rho_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}))]$. Moreover, by imposing α-mixing condition on the process $\{y_t\}$ and replacing $\boldsymbol{\psi}_0$ with $\boldsymbol{\psi}_0^*$ in Assumptions 2(ii) and 4(ii), Assumptions 1, 2 and 4 are replaced by the following conditions under mis-specification.

Assumption S1. $\{y_t : t = 1, 2, \ldots\}$ *is strictly stationary and α-mixing with the mixing coefficient $\alpha(n)$ satisfying $\sum_{n\geqslant 1}[\alpha(n)]^{1-2/\delta} < \infty$ for some $\delta > 2$.*

Assumption S2. *(i) The parameter space Ψ is compact; (ii) the pseudo-true parameter $\boldsymbol{\psi}_0^*$ is an interior point in Ψ.*

Assumption S3. *(i) $g_{t,\tau}(\boldsymbol{\psi})$ is continuous in $\boldsymbol{\psi} \in \Psi$; (ii) if $g_{t,\tau}(\boldsymbol{\psi}) = g_{t,\tau}(\boldsymbol{\psi}_0^*)$, then $\boldsymbol{\psi} = \boldsymbol{\psi}_0^*$.*

Denote $\boldsymbol{Z}_t^* = \sum_{k=1}^K \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))$ and $M^* = E(\boldsymbol{Z}_t^*\boldsymbol{Z}_t^{*\prime}) + \lim_{n\to\infty} n^{-1}\sum_{t\neq s}^n E(\boldsymbol{Z}_t^*\boldsymbol{Z}_s^{*\prime})$, where $\psi_\tau(x) = \tau - I(x < 0)$. Define the matrix $J^* = \sum_{k=1}^K E\left[\ddot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))\right]$ and let $N^* = N - J^*$. Corollary S1 below establishes the asymptotic properties of $\widehat{\boldsymbol{\psi}}_n$ under mis-specification, which implies that $\widehat{\boldsymbol{\psi}}_n$ is asymptotically biased as $\boldsymbol{\psi}_0^* \neq \boldsymbol{\psi}_0$ if $Q_\tau(\boldsymbol{\lambda})$ is mis-specified. The effect from mis-specification of $Q_\tau(\boldsymbol{\lambda})$ is examined through simulations in Section 4. Simulation results indicate that the parametric CQR estimator $\widetilde{\boldsymbol{\vartheta}}_n$ is insensitive to the mis-specification due to $Q_\tau(\boldsymbol{\lambda})$, while the conditional quantile estimation and forecasting can be slightly affected. As a result, in practice we can choose a distribution such as the Tukey-lambda distribution for η_t, which not only has explicit quantile function but also can approximate various distributions; see Section 3 for details.

Corollary S1 (The parametric CQR under mis-specification)**.** *Suppose* $\Pi^* = N^{*-1}M^*N^{*-1}$ *is positive definite. For* $\{y_t\}$ *generated by model* (2.1)*, if Assumptions 3 and 5-S3 hold, then we have (i)* $\widehat{\boldsymbol{\psi}}_n \to_p \boldsymbol{\psi}_0^*$*; (ii)* $\sqrt{n}(\widehat{\boldsymbol{\psi}}_n - \boldsymbol{\psi}_0^*) \to_{\mathcal{L}} N(0, \Pi^*)$ *as* $n \to \infty$*.*

S2 Asymptotic efficiency comparison

This subsection compares the proposed two CQR estimators for model (2.1), and then compares them with the Gaussian QMLE (GQMLE) and exponential QMLE (EQMLE).

We first compare the semi-parametric CQR estimator $\widehat{\boldsymbol{\vartheta}}_n$ and parametric CQR estimator $\widetilde{\boldsymbol{\vartheta}}_n$ for model (2.1). Let d (or ℓ) be the dimension of $\boldsymbol{\vartheta}$ (or $\boldsymbol{\lambda}$). Denote $R_1 = (I_d, \mathbf{0}_{d\times K})$ and $R_2 = (I_d, \mathbf{0}_{d\times \ell})$, where I_m is an $m \times m$ identity matrix and $\mathbf{0}_{m\times n}$ is an $m \times n$ zero matrix. Note that $\boldsymbol{\vartheta}_0 = R_1\boldsymbol{\phi}_0 = R_2\boldsymbol{\psi}_0$. Then by Theorems 2 and 4, it follows that $\sqrt{n}(\widehat{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0) \to_{\mathcal{L}} N(0, R_1\Xi R_1')$ and $\sqrt{n}(\widetilde{\boldsymbol{\vartheta}}_n - \boldsymbol{\vartheta}_0) \to_{\mathcal{L}} N(0, R_2\Pi R_2')$ as $n \to \infty$, where $\Xi = \Sigma^{-1}\Omega\Sigma^{-1}$ and $\Pi = N^{-1}MN^{-1}$. To compare the efficiency of $\widehat{\boldsymbol{\vartheta}}_n$ with that of $\widetilde{\boldsymbol{\vartheta}}_n$, we may compare the covariance matrices $R_1\Xi R_1'$ and $R_2\Pi R_2'$. Note that Σ and N can be rewritten as follows

$$\Sigma = \sum_{k=1}^K f(b_{k0}) \begin{pmatrix} h_t^{-1}(\boldsymbol{\vartheta}_0)\dot{\boldsymbol{q}}_{t,k}\dot{\boldsymbol{q}}_{t,k}' & \dot{\boldsymbol{q}}_{t,k}\boldsymbol{e}_k' \\ \boldsymbol{e}_k\dot{\boldsymbol{q}}_{t,k}' & h_t(\boldsymbol{\vartheta}_0)\boldsymbol{e}_k\boldsymbol{e}_k' \end{pmatrix} =: \begin{pmatrix} A_\Sigma & B_\Sigma \\ C_\Sigma & D_\Sigma \end{pmatrix},$$

$$N = \sum_{k=1}^K f(Q_{\tau_k}(\boldsymbol{\lambda}_0)) \begin{pmatrix} h_t^{-1}(\boldsymbol{\vartheta}_0)\dot{\boldsymbol{g}}_{t,k}\dot{\boldsymbol{g}}_{t,k}' & \dot{\boldsymbol{g}}_{t,k}\dot{\boldsymbol{Q}}_k' \\ \dot{\boldsymbol{Q}}_k\dot{\boldsymbol{g}}_{t,k}' & h_t(\boldsymbol{\vartheta}_0)\dot{\boldsymbol{Q}}_k\dot{\boldsymbol{Q}}_k' \end{pmatrix} =: \begin{pmatrix} A_N & B_N \\ C_N & D_N \end{pmatrix},$$

where $\dot{\boldsymbol{q}}_{t,k} = \dot{\mu}_t(\boldsymbol{\vartheta}_0) + b_{k0}\dot{h}_t(\boldsymbol{\vartheta}_0)$, $\dot{\boldsymbol{g}}_{t,k} = \dot{\mu}_t(\boldsymbol{\vartheta}_0) + Q_{\tau_k}(\boldsymbol{\lambda}_0)\dot{h}_t(\boldsymbol{\vartheta}_0)$, $\dot{\boldsymbol{Q}}_k = \dot{Q}_{\tau_k}(\boldsymbol{\lambda}_0)$, and $\boldsymbol{e}_k \in \mathbb{R}^K$ is the vector with its kth element being one and the others being zero. Under the situation of correct specification that $b_{\tau 0} = Q_\tau(\boldsymbol{\lambda}_0)$ and $q_{t,\tau}(\boldsymbol{\phi}_0) = g_{t,\tau}(\boldsymbol{\psi}_0)$ for all $\tau \in (0,1)$, we can verify that $\widehat{\boldsymbol{\vartheta}}_n$ and $\widetilde{\boldsymbol{\vartheta}}_n$ are asymptotically equivalent (i.e. $R_1 \Xi R_1' = R_2 \Pi R_2'$) if the following Condition (S2.1) or (S2.2) holds:

$$\ell = d \text{ and } \dot{Q}_{\tau_k}(\boldsymbol{\lambda}_0) = \boldsymbol{e}_k, \quad k = 1, \ldots, K, \tag{S2.1}$$

$$(D_\Sigma - C_\Sigma A_\Sigma^{-1} B_\Sigma)^{-1} = H(D_N - C_N A_N^{-1} B_N)^{-1} H', \tag{S2.2}$$

where $H = (\dot{Q}_{\tau_1}'(\boldsymbol{\lambda}_0), \ldots, \dot{Q}_{\tau_K}'(\boldsymbol{\lambda}_0))'$. However, this condition imposes very strong restrictions on the distribution of η_t, and thus $\widehat{\boldsymbol{\vartheta}}_n$ and $\widetilde{\boldsymbol{\vartheta}}_n$ are unlikely to be equivalent for general situations. Moreover, it is difficult to determine whether $R_1 \Xi R_1' - R_2 \Pi R_2'$ is (semi)-positive or (semi)-negative definite for general specifications of $Q_\tau(\boldsymbol{\lambda})$. Alternatively, we study the asymptotic relative efficiency (ARE) of $\widehat{\boldsymbol{\vartheta}}_n$ to $\widetilde{\boldsymbol{\vartheta}}_n$, which can be calculated as $\text{ARE}(\widehat{\boldsymbol{\vartheta}}_n, \widetilde{\boldsymbol{\vartheta}}_n) = (|R_2 \Pi R_2'|/|R_1 \Xi R_1'|)^{1/d}$ via simulation given the true parameter vectors and density function $f(\cdot)$ of η_t, where $|\cdot|$ is the determinant of a matrix; see Serfling (2009). Simulation results in Section 4 indicate that the parametric CQR estimator $\widetilde{\boldsymbol{\vartheta}}_n$ is asymptotically more efficient than the semi-parametric CQR estimator $\widehat{\boldsymbol{\vartheta}}_n$ when the data is more heavy-tailed.

Next we compare the semi-parametric CQR estimator $\widehat{\boldsymbol{\vartheta}}_n$ with the GQMLE and EQMLE for model (2.1). Define the GQMLE and EQMLE of model (2.1) as $\widehat{\boldsymbol{\vartheta}}_n^G = \arg\min_{\boldsymbol{\vartheta} \in \Theta} L_n^G(\boldsymbol{\vartheta})$ and $\widehat{\boldsymbol{\vartheta}}_n^E = \arg\min_{\boldsymbol{\vartheta} \in \Theta} L_n^E(\boldsymbol{\vartheta})$, respectively, where $L_n^G(\boldsymbol{\vartheta}) = n^{-1} \sum_{t=1}^n l_t^G(\boldsymbol{\vartheta})$ with $l_t^G(\boldsymbol{\vartheta}) = \ln h_t(\boldsymbol{\vartheta}) + 0.5[y_t - \mu_t(\boldsymbol{\vartheta})]^2/h_t^2(\boldsymbol{\vartheta})$, and $L_n^E(\boldsymbol{\vartheta}) = n^{-1} \sum_{t=1}^n l_t^E(\boldsymbol{\vartheta})$ with $l_t^E(\boldsymbol{\vartheta}) = \ln h_t(\boldsymbol{\vartheta}) + |y_t - \mu_t(\boldsymbol{\vartheta})|/h_t(\boldsymbol{\vartheta})$. Under the conditions that η_t has a zero mean and unit variance with $E(\eta_t^4) < \infty$, together with Assumptions 1-2, we have $\sqrt{n}(\widehat{\boldsymbol{\vartheta}}_n^G - \boldsymbol{\vartheta}_0) \to_{\mathcal{L}} N(0, S_G)$ as $n \to \infty$, where $S_G = U_G^{-1} V_G U_G^{-1}$ with $U_G = E[\partial^2 l_t^G(\boldsymbol{\vartheta}_0)/(\partial \boldsymbol{\vartheta} \partial \boldsymbol{\vartheta}')]$ and $V_G = E[\partial l_t^G(\boldsymbol{\vartheta}_0)/\partial \boldsymbol{\vartheta} \partial l_t^G(\boldsymbol{\vartheta}_0)/\partial \boldsymbol{\vartheta}']$. Meanwhile, under the conditions that η_t has a zero median with $E(|\eta_t|) = 1$ and $E(\eta_t^2) < \infty$, together with Assumptions 1-3, we have $\sqrt{n}(\widehat{\boldsymbol{\vartheta}}_n^E - \boldsymbol{\vartheta}_0) \to_{\mathcal{L}} N(0, S_E)$ as $n \to \infty$, where

$S_E = U_E^{-1} V_E U_E^{-1}/4$ with $U_E = f(0) E[h_t^{-2}(\boldsymbol{\vartheta}_0)\dot{\mu}_t(\boldsymbol{\vartheta}_0)\dot{\mu}_t'(\boldsymbol{\vartheta}_0)] + 0.5E[h_t^{-2}(\boldsymbol{\vartheta}_0)\dot{h}_t(\boldsymbol{\vartheta}_0)\dot{h}_t'(\boldsymbol{\vartheta}_0)]$ and $V_E = E[h_t^{-2}(\boldsymbol{\vartheta}_0)\dot{\mu}_t(\boldsymbol{\vartheta}_0)\dot{\mu}_t'(\boldsymbol{\vartheta}_0)] + E(\eta_t)E\{h_t^{-2}(\boldsymbol{\vartheta}_0)[\dot{\mu}_t(\boldsymbol{\vartheta}_0)\dot{h}_t'(\boldsymbol{\vartheta}_0) + \dot{h}_t(\boldsymbol{\vartheta}_0)\dot{\mu}_t'(\boldsymbol{\vartheta}_0)]\} + [E(\eta_t^2) - 1]E[h_t^{-2}(\boldsymbol{\vartheta}_0)\dot{h}_t(\boldsymbol{\vartheta}_0)\dot{h}_t'(\boldsymbol{\vartheta}_0)]$. To compare the asymptotic efficiency of the semi-parametric CQR estimator $\widehat{\boldsymbol{\vartheta}}_n$ with that of the GQMLE and EQMLE, it suffices to compare the asymptotic covariance $R_1 \Xi R_1'$ with S_G and S_E, respectively. Note that all the above asymptotic covariances depend on η_t. If η_t follows the standard normal (or Laplace) distribution, then GQMLE (or EQMLE) reduces to the MLE with the asymptotic covariance $S_G = U_G^{-1}$ (or $S_E = \{E[h_t^{-2}(\boldsymbol{\vartheta}_0)(\dot{\mu}_t(\boldsymbol{\vartheta}_0)\dot{\mu}_t'(\boldsymbol{\vartheta}_0) + \dot{h}_t(\boldsymbol{\vartheta}_0)\dot{h}_t'(\boldsymbol{\vartheta}_0))]\}^{-1}$) attaining the Cramér-Rao lower bound. Thus the GQMLE $\check{\boldsymbol{\vartheta}}_n$ (or EQMLE $\breve{\boldsymbol{\vartheta}}_n$) is the most efficient among these three estimators when η_t follows the standard normal (or Laplace) distribution. For general distributions of η_t, it is difficult to determine which estimator is asymptotically more efficient. Alternatively, as for the comparison between both CQRs, we also calculate the ARE of $\widehat{\boldsymbol{\vartheta}}_n$ to $\check{\boldsymbol{\vartheta}}_n$ or $\breve{\boldsymbol{\vartheta}}_n$ via simulation. Given the true parameter vectors and density function $f(\cdot)$ of η_t, the ARE is calculated as $\text{ARE}(\widehat{\boldsymbol{\vartheta}}_n, \check{\boldsymbol{\vartheta}}_n) = (|S_G|/|R_1 \Xi R_1'|)^{1/d}$ and $\text{ARE}(\widehat{\boldsymbol{\vartheta}}_n, \breve{\boldsymbol{\vartheta}}_n) = (|S_E|/|R_1 \Xi R_1'|)^{1/d}$, respectively. Simulation results in Section 4 indicate that the semi-parametric CQR is more efficient than the GQMLE and EQMLE when the data is more heavy-tailed.

S3 The ARE of $\widehat{\boldsymbol{\vartheta}}_n$ to $\widetilde{\boldsymbol{\vartheta}}_n$, $\check{\boldsymbol{\vartheta}}_n$ or $\breve{\boldsymbol{\vartheta}}_n$ for Examples 1-3

For the ARMA-GARCH, ALDAR and ESTAR-GARCH models in Examples 1-3, note that $\omega = 1$ is imposed by Assumptions $4'$, $4''$ and $4'''$ for identification of the semi-parametric CQR, whereas such condition is not required for the parametric CQR, GQMLE and EQMLE. Therefore, these models should be reparameterized such that $\omega = 1$, and then the ARE is calculated using the true parameter vectors of the reparameterized models. For illustration, we show how to calculate the ARE of $\widehat{\boldsymbol{\vartheta}}_n$ to $\widetilde{\boldsymbol{\vartheta}}_n$, and the ARE of $\widehat{\boldsymbol{\vartheta}}_n$ to $\check{\boldsymbol{\vartheta}}_n$ or $\breve{\boldsymbol{\vartheta}}_n$ can be similarly obtained.

The ARE of $\widehat{\boldsymbol{\vartheta}}_n$ to $\widetilde{\boldsymbol{\vartheta}}_n$ is $\text{ARE}(\widehat{\boldsymbol{\vartheta}}_n, \widetilde{\boldsymbol{\vartheta}}_n) = (|\dot{R}_2(\boldsymbol{\psi}_0)\Pi\dot{R}_2'(\boldsymbol{\psi}_0)|/|R_1\Xi R_1'|)^{1/d}$ after model reparameterization, where $\dot{R}_2(\boldsymbol{\psi}_0)$ is the first derivative of $R_2(\boldsymbol{\psi}_0)$. Specifically, for the ARMA-GARCH model (2.2), let $\boldsymbol{\vartheta}_0 = (\alpha_{10}, \ldots, \alpha_{p0}, \beta_{10}, \ldots, \beta_{q0}, \boldsymbol{\upsilon}_0')' \in \mathbb{R}^{p+q} \times [0,\infty)^{P+Q}$ with $\boldsymbol{\upsilon}_0 = (\gamma_{10}, \ldots, \gamma_{Q0}, \nu_{10}, \ldots, \nu_{P0})'$ and $d = p+q+P+Q$. Then $\dot{R}_2(\boldsymbol{\psi}_0)$ has the form of

$$\dot{R}_2(\boldsymbol{\psi}_0) = \begin{pmatrix} I_{p+q+P} & \mathbf{0}_{(p+q+P)\times 1} & \mathbf{0}_{(p+q+P)\times Q} & \mathbf{0}_{(p+q+P)\times \ell} \\ \mathbf{0}_{Q\times(p+q+P)} & -\omega_0^{-2}\boldsymbol{\upsilon}_0 & \omega_0^{-1} I_Q & \mathbf{0}_{Q\times\ell} \end{pmatrix}.$$

For the ALDAR model (2.3), let $\boldsymbol{\vartheta}_0 = (\varphi_{10}, \ldots, \varphi_{p0}, \boldsymbol{\alpha}_0')'$ with $\boldsymbol{\alpha}_0 = (\alpha_{10}^+, \ldots, \alpha_{q0}^+, \alpha_{10}^-, \ldots, \alpha_{q0}^-)'$ and $d = p + 2q$. Then $\dot{R}_2(\boldsymbol{\psi}_0)$ has the form of

$$\dot{R}_2(\boldsymbol{\psi}_0) = \begin{pmatrix} I_p & \mathbf{0}_{p\times 1} & \mathbf{0}_{p\times 2q} & \mathbf{0}_{p\times\ell} \\ \mathbf{0}_{2q\times p} & -\omega_0^{-2}\boldsymbol{\alpha}_0 & \omega_0^{-1} I_{2q} & \mathbf{0}_{2q\times\ell} \end{pmatrix}.$$

For the ESTAR-GARCH model (2.4), let $\boldsymbol{\vartheta}_0 = (\alpha_{00}, \ldots, \alpha_{0p}, \alpha_{10}, \ldots, \alpha_{1p}, \gamma_0, c_0, a_0, b_0)'$ and $d = 2p + 6$. Then $\dot{R}_2(\boldsymbol{\psi}_0)$ has the form of

$$\dot{R}_2(\boldsymbol{\psi}_0) = \begin{pmatrix} I_{2p+4} & \mathbf{0}_{(2p+4)\times 1} & \mathbf{0}_{(2p+4)\times 1} & \mathbf{0}_{(2p+4)\times 1} & \mathbf{0}_{(2p+4)\times\ell} \\ \mathbf{0}_{1\times(2p+4)} & -\omega_0^{-2} a_0 & 0 & \omega_0^{-1} & \mathbf{0}_{1\times\ell} \\ \mathbf{0}_{1\times(2p+4)} & 0 & 1 & 0 & \mathbf{0}_{1\times\ell} \end{pmatrix}.$$

S4 Additional illustrations for CQRs

S4.1 CQRs in DAR type models

This subsection investigates the proposed CQRs in the framework of DAR type models. Here we focus on the ALDAR model in (2.3) for illustration, and the proposed CQRs can be similarly applied to other DAR type models in Example 2.

For the ALDAR model in (2.3), note that the location and scale functions $\mu_t(\boldsymbol{\vartheta}^{\text{II}})$ and $h_t(\boldsymbol{\vartheta}^{\text{II}})$ only depend on the finite past observations, and thus initial values are not required for estimation. Then the conditional quantile functions for the semi-parametric and parametric CQRs can be rewritten as $\widetilde{q}^{\text{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\text{II}}) = q^{\text{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\text{II}}) = \mu_t(\boldsymbol{\vartheta}^{\text{II}}) + b_k h_t(\boldsymbol{\vartheta}^{\text{II}})$ and $\widetilde{g}^{\text{II}}_{t,\tau}(\boldsymbol{\psi}^{\text{II}}) = g^{\text{II}}_{t,\tau}(\boldsymbol{\psi}^{\text{II}}) = \mu_t(\boldsymbol{\vartheta}^{\text{II}}) + Q_\tau(\lambda) h_t(\boldsymbol{\vartheta}^{\text{II}})$, respectively, where $\boldsymbol{\phi}^{\text{II}} = (\boldsymbol{\vartheta}^{\text{II}\prime}, b_1, \ldots, b_K)'$ and $\boldsymbol{\psi}^{\text{II}} = (\boldsymbol{\vartheta}^{\text{II}\prime}, \lambda)'$. As a result, the semi-parametric and parametric CQR estimators $\widehat{\boldsymbol{\phi}}^{\text{II}}_n$ and $\widehat{\boldsymbol{\psi}}^{\text{II}}_n$ can be calculated by (2.6) and (2.9), with $t = 1$ replaced by $t = \max(p, q) + 1$, and $\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi})$ and $\widetilde{g}_{t,\tau_k}(\boldsymbol{\psi})$ replaced by $\widetilde{q}^{\text{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\text{II}})$ and $\widetilde{g}^{\text{II}}_{t,\tau_k}(\boldsymbol{\psi}^{\text{II}})$, respectively.

It can be proved that the asymptotic properties of $\widehat{\boldsymbol{\phi}}^{\text{II}}_n$ and $\widehat{\boldsymbol{\psi}}^{\text{II}}_n$ in Theorems 1-4 still hold for the ALDAR model (2.3), with some regular conditions adjusted accordingly. Specifically, a sufficient condition for Assumption 1 is provided in Theorem 2.1 in Tan and Zhu (2022). Meanwhile, Assumptions 4″-5″ below provide the sufficient conditions for identification and moment conditions in Assumptions 4-5, respectively; see the detailed proofs in Section S6.2 of the Supplementary Material.

Assumption 4″ (Identification)**.** *(i)* $\omega = 1$ *and* $K \geqslant 2$ *for semi-parametric CQR; (ii)* $K \geqslant 4$ *and* $\lambda < 1$ *for parametric CQR.*

Assumption 5″ (Moments)**.** *(i)* $\underline{\varphi} \leqslant \varphi_i \leqslant \overline{\varphi}$ *for* $1 \leqslant i \leqslant p$*,* $0 < \underline{\omega} \leqslant \omega \leqslant \overline{\omega}$*,* $0 < \underline{\alpha} \leqslant \alpha_i^+, \alpha_i^- \leqslant \overline{\alpha}$ *for* $1 \leqslant i \leqslant q$*,* $\underline{b} \leqslant b_k \leqslant \overline{b}$ *for* $1 \leqslant k \leqslant K$ *and* $\underline{\lambda} \leqslant \lambda \leqslant \overline{\lambda}$*, where* $\underline{\varphi}$*,* $\overline{\varphi}$*,* $\underline{\omega}$*,* $\overline{\omega}$*,* $\underline{\alpha}$*,* $\overline{\alpha}$*,* $\underline{b}$*,*$\overline{b}$*,* $\underline{\lambda}$ *and* $\overline{\lambda}$ *are some constants; (ii)* $E(|y_t|^3) < \infty$*.*

S4.2 CQRs in NAR-GARCH models

This subsection studies the proposed CQRs in the framework of NAR-GARCH models. We focus on the 3-regime exponential STAR-GARCH model in (2.4) for illustration, and the proposed CQRs can be similarly applied to other NAR-GARCH models in Example 3.

For the ESTAR-GARCH model in (2.4), the scale function $h_t(\boldsymbol{\vartheta}^{\text{III}})$ depends on observations in the infinite past, and thus initial values are needed to calculate the feasible conditional quantile functions. We set $y_s = \epsilon_s = 0$ and $h_s = 1$ for $s \leqslant 0$ as initial values, and denote the feasible conditional quantile functions as $\widetilde{q}^{\mathrm{III}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{III}})$ and $\widetilde{g}^{\mathrm{III}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{III}})$ for the semi-parametric and parametric CQRs, where $\boldsymbol{\phi}^{\mathrm{III}} = (\boldsymbol{\vartheta}^{\mathrm{III}\prime}, b_1, \ldots, b_K)'$ and $\boldsymbol{\psi}^{\mathrm{III}} = (\boldsymbol{\vartheta}^{\mathrm{III}\prime}, \lambda)'$. It will be proved that the effect of initial values on the estimation is asymptotically negligible.

For model (2.4), the semi-parametric and parametric CQR estimators $\widehat{\boldsymbol{\phi}}^{\mathrm{III}}_n$ and $\widehat{\boldsymbol{\psi}}^{\mathrm{III}}_n$ are defined by (2.6) and (2.9), with $\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi})$ and $\widetilde{g}_{t,\tau_k}(\boldsymbol{\psi})$ replaced by $\widetilde{q}^{\mathrm{III}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{III}})$ and $\widetilde{g}^{\mathrm{III}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{III}})$, respectively. Denote their true parameter vectors by $\boldsymbol{\phi}^{\mathrm{III}}_0 = (\boldsymbol{\vartheta}^{\mathrm{III}\prime}_0, b_{10}, \ldots, b_{K0})'$ and $\boldsymbol{\psi}^{\mathrm{III}}_0 = (\boldsymbol{\vartheta}^{\mathrm{III}\prime}_0, \lambda_0)'$, where $\boldsymbol{\vartheta}^{\mathrm{III}}_0 = (\alpha^*_{00}, \alpha^*_{01}, \ldots, \alpha^*_{0p}, \alpha^*_{10}, \alpha^*_{11}, \ldots, \alpha^*_{1p}, \gamma_0, c_0, \omega_0, a_0, b_0)'$. We will prove that the asymptotic properties of $\widehat{\boldsymbol{\phi}}^{\mathrm{III}}_n$ and $\widehat{\boldsymbol{\psi}}^{\mathrm{III}}_n$ in Theorems 1-4 still hold for the ESTAR-GARCH model in (2.4), with some regular conditions adjusted accordingly. Specifically, a sufficient condition for Assumption 1 is provided in Theorem 3 of Chan et al. (2015) for ESTAR(1)-GARCH$(1,1)$ model. Assumptions 4‴-5‴ below give the sufficient conditions for identification and moment conditions in Assumptions 4-5, respectively; see the detailed proofs in Section S6.3 of the Supplementary Material.

Assumption 4‴ (Identification)**.** *(i)* $\omega = 1$ *for semi-parametric CQR; (ii)* $K \geqslant 4$ *and* $\lambda < 1$ *for parametric CQR.*

Assumption 5‴ (Moments)**.** *(i)* $\underline{\alpha} \leqslant \alpha_{ij} \leqslant \overline{\alpha}$ *for* $i = 0, 1$ *and* $0 \leqslant j \leqslant p$, $0 < \underline{\omega} \leqslant \omega \leqslant \overline{\omega}$, $0 < \underline{a} \leqslant a \leqslant \overline{a}$, $0 < \underline{\beta} \leqslant b \leqslant \overline{\beta}$, $\underline{b} \leqslant b_k \leqslant \overline{b}$ *for* $1 \leqslant k \leqslant K$ *and* $\underline{\lambda} \leqslant \lambda \leqslant \overline{\lambda}$, *where* $\underline{\alpha}$, $\overline{\alpha}$, $\underline{\omega}$, $\overline{\omega}$, $\underline{a}$, $\overline{a}$, $\underline{\beta}$, $\overline{\beta}$, $\underline{b}$, $\overline{b}$, $\underline{\lambda}$ *and* $\overline{\lambda}$ *are some constants; (ii)* $E(y_t^2) < \infty$.

Remark S1 (Advantages of CQR over QR)**.** *This paper considers the CQR instead of QR for conditional quantile estimation, owing to two advantages of CQR over QR: (a) CQR has efficiency gain than QR by combining data information at multiple quantile levels, whereas the* *efficiency of QR at a tail quantile level can be arbitrarily low; see also Zou and Yuan (2008). (b) CQR with a reasonable choice of K can avoid the identification issue in estimation for model* (2.1). *Particularly, Assumptions $4'$, $4''$ and $4'''$ indicate that $K \geqslant 4$ is needed for identification of the ARMA-GARCH, ALDAR and ESTAR-GARCH models estimated by the parametric CQR, while $K \geqslant 2$ is required for the ALDAR model estimated by the semi-parametric CQR. These findings suggest that the QR (i.e. CQR with $K = 1$) can be faced with identification issue in estimation for location-scale time series models. As a result, the CQR can not only improve estimation efficiency of QR, but also solve the identification problem of QR for many specifications of model* (2.1).

S5 Proofs of Theorems 1-4 and Corollary S1

In this section, we show the proofs of Theorems 1-4 and Corollary S1. Since the proofs of Theorems 3-4 are similar to the proofs of Theorems 1-2, we only verify Theorems 1-2 and Corollary S1 in this section.

S5.1 Proof of Theorem 1

To verify Theorem 1, we need the following lemma.

Lemma 1. *Define $L_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^{n} l_t(\boldsymbol{\phi})$, where $l_t(\boldsymbol{\phi}) = \sum_{k=1}^{K} \rho_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}))$. Let $\widetilde{L}_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^{n} \widetilde{l}_t(\boldsymbol{\phi})$, where $\widetilde{l}_t(\boldsymbol{\phi}) = \sum_{k=1}^{K} \rho_{\tau_k}(y_t - \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}))$. If Assumptions 1-4, 5 (i)* *and 6 (i) hold, then*

$$(1)\ E\left[\sup_{\boldsymbol{\phi}\in\Phi}|l_t(\boldsymbol{\phi})|\right]<\infty;$$

$$(2)\ E\ [l_t(\boldsymbol{\phi})]\ \textit{has a unique minimum at } \boldsymbol{\phi}_0;$$

$$(3)\ \sup_{\boldsymbol{\phi}\in\Phi}|L_n(\boldsymbol{\phi})-\widetilde{L}_n(\boldsymbol{\phi})|=o_p(1).$$

Proof. Recall that $q_{t,\tau_k}(\boldsymbol{\phi})=\mu_t(\boldsymbol{\vartheta})+b_kh_t(\boldsymbol{\vartheta})$ and $\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi})$ is the approximation of $q_{t,\tau_k}(\boldsymbol{\phi})$, where $\boldsymbol{\phi}=(\boldsymbol{\vartheta}',\boldsymbol{b}')'$ with $\boldsymbol{b}=(b_1,b_2,\ldots,b_K)'$. For Lemma 1 (1), since $\{y_t\}$ is ergodic and stationary by Assumption 1, $E\left[\sup_{\boldsymbol{\phi}\in\Phi}|q_{t,\tau_k}(\boldsymbol{\phi})|\right]<\infty$ by Assumption 5 (i), and by the fact that $|\rho_\tau(y)|\leqslant|y|$, it holds that

$$\begin{aligned}
E\sup_{\boldsymbol{\phi}\in\Phi}|l_t(\boldsymbol{\phi})| &\leqslant \sum_{k=1}^{K}E\left(\sup_{\boldsymbol{\phi}\in\Phi}|\rho_{\tau_k}(y_t-q_{t,\tau_k}(\boldsymbol{\phi}))|\right)\\
&\leqslant \sum_{k=1}^{K}E\left(\sup_{\boldsymbol{\phi}\in\Phi}|y_t-q_{t,\tau_k}(\boldsymbol{\phi})|\right)\\
&\leqslant E(|y_t|)+E\left(\sup_{\boldsymbol{\phi}\in\Phi}|q_{t,\tau_k}(\boldsymbol{\phi})|\right)\\
&<\infty.
\end{aligned}$$

We then consider Lemma 1 (2). For $x\neq0$, it holds that

$$\begin{aligned}
\rho_\tau(x-y)-\rho_\tau(x) &= -y\psi_\tau(x)+y\int_0^1\left[I(x\leqslant ys)-I(x\leqslant0)\right]ds\\
&= -y\psi_\tau(x)+(x-y)\left[I(0>x>y)-I(0<x<y)\right],
\end{aligned}\tag{S5.3}$$

where $\psi_\tau(x)=\tau-I(x<0)$; see Knight (1998). Let $\xi_{t,\tau_k}=y_t-q_{t,\tau_k}(\boldsymbol{\phi}_0)$ and $\upsilon_{t,\tau_k}(\boldsymbol{\phi})=q_{t,\tau_k}(\boldsymbol{\phi})-q_{t,\tau_k}(\boldsymbol{\phi}_0)$. By (S5.3), it follows that

$$\begin{aligned}
&l_t(\boldsymbol{\phi})-l_t(\boldsymbol{\phi}_0)\\
=&-\upsilon_{t,\tau_k}(\boldsymbol{\phi})\psi_{\tau_k}(\xi_{t,\tau_k})+[\xi_{t,\tau_k}-\upsilon_{t,\tau_k}(\boldsymbol{\phi})][I(0>\xi_{t,\tau_k}>\upsilon_{t,\tau_k}(\boldsymbol{\phi}))-I(0<\xi_{t,\tau_k}<\upsilon_{t,\tau_k}(\boldsymbol{\phi}))].
\end{aligned}$$

This together with $E[\psi_{\tau_k}(\xi_{t,\tau_k})|\mathcal{F}_{t-1}] = 0$, implies that

$$\begin{aligned} &E\left[l_t(\boldsymbol{\phi})\right] - E\left[l_t(\boldsymbol{\phi}_0)\right] \\ =&\sum_{k=1}^{K} E\left\{[\xi_{t,\tau_k} - \upsilon_{t,\tau_k}(\boldsymbol{\phi})][I(0 > \xi_{t,\tau_k} > \upsilon_{t,\tau_k}(\boldsymbol{\phi})) - I(0 < \xi_{t,\tau_k} < \upsilon_{t,\tau_k}(\boldsymbol{\phi}))]\right\} \geqslant 0. \end{aligned}$$

Since $f(\cdot)$ is continuous at the neighborhood of $q_{t,\tau_k}(\boldsymbol{\phi}_0)$ by Assumption 3 , then the above quality holds if and only if $\upsilon_{t,\tau_k}(\boldsymbol{\phi}) = 0$ with probability one for $t \in \mathbb{Z}$. This together with Assumption 4 (ii), implies that $\boldsymbol{\phi} = \boldsymbol{\phi}_0$ and Lemma 1 (2) is verified.

For Lemma 1 (3), since $\sum_{t=1}^{\infty} \sup_{\boldsymbol{\phi}\in\Phi} |q_{t,\tau_k}(\boldsymbol{\phi}) - \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi})| < \infty$ by Assumption 6 (i), it holds that

$$\begin{aligned} \sup_{\boldsymbol{\phi}\in\Phi} |L_n(\boldsymbol{\phi}) - \widetilde{L}_n(\boldsymbol{\phi})| \leqslant& \frac{1}{n}\sum_{t=1}^{n}\sum_{k=1}^{K}\sup_{\boldsymbol{\phi}\in\Phi} |\rho_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi})) - \rho_{\tau_k}(y_t - \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}))| \\ <& \frac{2}{n}\sum_{t=1}^{n}\sum_{k=1}^{K}\sup_{\boldsymbol{\phi}\in\Phi} |q_{t,\tau_k}(\boldsymbol{\phi}) - \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi})| = o_p(1). \end{aligned}$$

Hence, Lemma 1 (3) is verified. □

Proof of Theorem 1. Since $l_t(\boldsymbol{\phi}) - E[l_t(\boldsymbol{\phi})]$ is a measurable function of y_t in Euclidean space for each $\boldsymbol{\phi} \in \Phi$, which is also a continuous function of $\boldsymbol{\phi} \in \Phi$ for each y_t. Then by Theorem 3.1 of Ling and McAleer (2003), together with Lemma 1 (i) and the strict stationarity and ergodicity of $\{y_t\}$ by Assumption 1, we have

$$\sup_{\boldsymbol{\phi}\in\Phi} |L_n(\boldsymbol{\phi}) - E[l_t(\boldsymbol{\phi})]| = o_p(1).$$

This together with Lemma 1 (iii), implies that

$$\sup_{\boldsymbol{\phi}\in\Phi} |\widetilde{L}_n(\boldsymbol{\phi}) - E[l_t(\boldsymbol{\phi})]| = o_p(1). \tag{S5.4}$$

We next verify the uniform consistency. For any $c > 0$, with probability tending to 1 uniformly in $\epsilon \geqslant c$ and by $\widehat{\boldsymbol{\phi}}_n = \arg\min_{\boldsymbol{\phi}\in\Phi} \widetilde{L}_n(\boldsymbol{\phi})$, it holds that

$$\widetilde{L}_n(\widehat{\boldsymbol{\phi}}_n) \leqslant \widetilde{L}_n(\boldsymbol{\phi}_0) + \epsilon/3, \tag{S5.5}$$

and by (S5.4), it holds that

$$E[l_t(\widehat{\boldsymbol{\phi}}_n)] < \widetilde{L}_n(\widehat{\boldsymbol{\phi}}_n)) + \epsilon/3, \tag{S5.6}$$

$$\widetilde{L}_n(\boldsymbol{\phi}_0) < E[l_t(\boldsymbol{\phi}_0)] + \epsilon/3. \tag{S5.7}$$

Combining (S5.5), (S5.6) and (S5.7), with probability tending to 1, we have

$$E[l_t(\widehat{\boldsymbol{\phi}}_n)] < E[l_t(\boldsymbol{\phi}_0)] + \epsilon. \tag{S5.8}$$

Let $B_\delta(\boldsymbol{\phi}_0)$ be an open neighborhood of $\boldsymbol{\phi}_0$ with radius $\delta > 0$, then $B^c = \Phi/B_\delta(\boldsymbol{\phi}_0)$ is compact and $\inf_{\boldsymbol{\phi}\in B^c} E[l_t(\boldsymbol{\phi})]$ exists. Denote $\epsilon = \inf_{\boldsymbol{\phi}\in B^c} E[l_t(\boldsymbol{\phi})] - E[l_t(\boldsymbol{\phi}_0)]$. Since $\boldsymbol{\phi}_0 = \arg\min_{\boldsymbol{\phi}\in\Phi} E[l_t(\boldsymbol{\phi})]$ is unique by Lemma 1 (ii), we have $\epsilon > 0$. Select $c > 0$ for any $\epsilon > 0$ which satisfies $Pr(\epsilon > c) > 1-\epsilon$. Then combining with (S5.8), it follows that with probability greater than $1-\epsilon$,

$$E[l_t(\widehat{\boldsymbol{\phi}}_n)] < E[l_t(\boldsymbol{\phi}_0)] + \inf_{\boldsymbol{\phi}\in B^c} E[l_t(\boldsymbol{\phi})] - E[l_t(\boldsymbol{\phi}_0)] < \inf_{\boldsymbol{\phi}\in B^c} E[l_t(\boldsymbol{\phi})].$$

Therefore, we have $\widehat{\boldsymbol{\phi}}_n \in B_\delta(\boldsymbol{\phi}_0)$ and $\|\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0\| < \delta$ with probability greater than $1-\epsilon$. By the arbitrariness of ϵ, it holds that $\|\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0\| < \delta$ with probability tending to 1. The proof is accomplished. □

S5.2 Proof of Theorem 2

To show Theorem 2, we introduce Lemmas 2-4 below. Specifically, Lemma 2 verifies the stochastic differentiability condition defined by Pollard (1985), and the bracketing method in Pollard (1985) is used for its proof. Lemmas 3 and 4 are used to obtain the $\sqrt{n}$-consistency and asymptotic normality of $\widehat{\boldsymbol{\phi}}_n$, and the proof of Lemma 3 needs Lemma 2.

Lemma 2. *If Assumptions 1-5 hold, then for any sequence of random variables* $\boldsymbol{u}_n$ *such that*

$\boldsymbol{u}_n = o_p(1)$, *it holds that*

$$\pi_{1n}(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2),$$

where $\pi_{1n}(\boldsymbol{u}) = \boldsymbol{u}' \sum_{k=1}^{K} \sum_{t=1}^{n} \dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\{X_{t,\tau_k}(\boldsymbol{u}) - E\left[X_{t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right]\}$ *with*

$$X_{t,\tau_k}(\boldsymbol{u}) = \int_0^1 \left[I(y_t \leqslant q_{t,\tau_k}(\boldsymbol{\phi}_0) + \Delta_{t,\tau_k}(\boldsymbol{u})s) - I(y_t \leqslant q_{t,\tau_k}(\boldsymbol{\phi}_0))\right] ds$$

and $\Delta_{t,\tau_k}(\boldsymbol{u}) = q_{t,\tau_k}(\boldsymbol{\phi}_0 + \boldsymbol{u}) - q_{t,\tau_k}(\boldsymbol{\phi}_0)$.

Proof. Define $\aleph = \{\boldsymbol{u} : \boldsymbol{\phi}_0 + \boldsymbol{u} \in \Phi\}$, where $\boldsymbol{\phi}_0 = (\boldsymbol{\vartheta}_0', \boldsymbol{b}_0')'$ with $\boldsymbol{b}_0 = (b_{\tau_1}, \ldots, b_{\tau_K})'$ is the true parameter and Φ is the parameter space. For $\boldsymbol{u} \in \aleph$, note that

$$\pi_{1n}(\boldsymbol{u}) \leqslant \sqrt{n}\|\boldsymbol{u}\| \sum_{k=1}^{K} \sum_{j=1}^{d} \left| \frac{1}{\sqrt{n}} \sum_{t=1}^{n} m_{t,\tau_k,j}\{X_{t,\tau_k}(\boldsymbol{u}) - E\left[X_{t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right]\} \right|,$$

where d is the dimension of $\boldsymbol{\phi}$, $m_{t,\tau_k,j} = \partial q_{t,\tau_k}(\boldsymbol{\phi}_0)/\partial \phi_j$ with ϕ_j is jth element with $\boldsymbol{\phi}$. For $1 \leqslant j \leqslant d$ and $\tau \in [0,1]$, define $g_{t,\tau} = \max_j\{m_{t,\tau,j}, 0\}$ or $g_{t,\tau} = \max_j\{-m_{t,\tau,j}, 0\}$. Let $f_{t,\tau}(\boldsymbol{u}) = g_{t,\tau} X_{t,\tau}(\boldsymbol{u})$ and define

$$D_{n,\tau}(\boldsymbol{u}) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{f_{t,\tau}(\boldsymbol{u}) - E\left[f_{t,\tau}(\boldsymbol{u})|\mathcal{F}_{t-1}\right]\}.$$

To establish Lemma S.1, it suffices to show that, for any $\eta > 0$,

$$\sup_{\|\boldsymbol{u}\| \leqslant \eta} \frac{|D_{t,\tau}(\boldsymbol{u})|}{1 + \sqrt{n}\|\boldsymbol{u}\|} = o_p(1). \tag{S5.9}$$

We follow Lemma 4 of Pollard (1985) to verify (S5.9). Let $\mathfrak{F} = \{f_t(\boldsymbol{u}) : \|\boldsymbol{u}\| \leqslant \eta\}$ be a collection of functions indexed by $\boldsymbol{u}$. First, we verify that $\mathfrak{F}$ satisfies the bracketing condition defined on page 304 of Pollard (1985). Let $B_r(\boldsymbol{\zeta})$ be an open neighborhood of $\boldsymbol{\zeta}$ with radius $r > 0$, and define a constant C_1 to be selected later. For any fixed $\epsilon > 0$ and $0 < \delta \leqslant \eta$, there exists a sequence of small cubes $\{B_{\epsilon\delta/C_1}(\boldsymbol{u}_i)\}_{i=1}^{K_\epsilon}$ to cover $B_\delta(\mathbf{0})$, where K_ϵ is an integer less than $C_0\epsilon^{-d}$ and C_0 is depending on neither ϵ nor δ. Denote $V_i(\delta) = B_{\epsilon\delta/C_0}(\boldsymbol{u}_i) \bigcap B_\delta(\mathbf{0})$.

Let $U_1(\delta) = V_1(\delta)$ and $U_i(\delta) = V_i(\delta) - \bigcup_{j=1}^{i-1} V_j(\delta)$ for $i \geqslant 2$, then $\{U_i(\delta)\}_{i=1}^{K_\epsilon}$ is a partition of $B_\delta(\mathbf{0})$. For each $\boldsymbol{u}_i \in U_i(\delta)$ with $1 \leqslant i \leqslant K_\epsilon$, define the following bracketing functions

$$f_{t,\tau}^{\pm}(\boldsymbol{u}) = g_{t,\tau} \int_0^1 \left[I\left(y_t \leqslant q_{t,\tau}(\boldsymbol{\phi}_0) + \Delta_{t,\tau}(\boldsymbol{u})s \pm \frac{\epsilon\delta}{C_1} \|\dot{q}_{t,\tau}(\boldsymbol{\phi}_0)\| \right) - I\left(y_t \leqslant q_{t,\tau}(\boldsymbol{\phi}_0)\right) \right] ds.$$

Since $I(\cdot)$ is non-decreasing and $g_{t,\tau} \geqslant 0$, for any $\boldsymbol{u} \in U_i(\delta)$, we have

$$f_{t,\tau}^{-}(\boldsymbol{u}_i) \leqslant f_{t,\tau}(\boldsymbol{u}) \leqslant f_{t,\tau}^{+}(\boldsymbol{u}_i). \tag{S5.10}$$

Furthermore, by Taylor expansion, it holds that

$$E\left[f_{t,\tau}^{+}(\boldsymbol{u}_i) - f_{t,\tau}^{-}(\boldsymbol{u}_i) | \mathcal{F}_{t-1} \right] \leqslant \frac{\epsilon\delta}{C_1} 2 \sup_x f(x) \frac{1}{h_t(\boldsymbol{\vartheta}_0)} \|\dot{q}_{t,\tau}(\boldsymbol{\phi}_0)\|^2. \tag{S5.11}$$

Denote $\varrho_{t,\tau} = 2\sup_x f(x) \|\dot{q}_{t,\tau}(\boldsymbol{\phi}_0)\|^2 / h_t(\boldsymbol{\vartheta}_0)$. By Assumption 3, we have $\sup_x f(x) < \infty$. Choose $C_1 = E(\varrho_{t,\tau})$. Then by iterative-expectation and Assumption 5 (ii), it follows that

$$E\left[f_{t,\tau}^{+}(\boldsymbol{u}_i) - f_{t,\tau}^{-}(\boldsymbol{u}_i) \right] = E\{ E\left[f_{t,\tau}^{+}(\boldsymbol{u}_i) - f_{t,\tau}^{-}(\boldsymbol{u}_i) | \mathcal{F}_{t-1} \right] \} \leqslant \epsilon\delta.$$

This together with (S5.10), implies that the family $\mathfrak{F}$ satisfies the bracketing condition.

Pick $\delta_k = 2^{-k}\eta$. Define $B(k) = B_{\delta_k}(\mathbf{0})$, and $A(k)$ to be the annulus $B(k) \backslash B(k+1)$. Fix $\epsilon > 0$, for each $1 \leqslant i \leqslant K_\epsilon$, by bracketing condition, there exists a partition $\{U_i(\delta_k)\}_{i=1}^{K_\epsilon}$ of $B(k)$. For $\boldsymbol{u} \in U_i(\delta_k)$, by (S5.11) with $\delta = \delta_k$, it holds that

$$\begin{aligned} D_{t,\tau}(\boldsymbol{u}) &\leqslant \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{ f_{t,\tau}^{+}(\boldsymbol{u}_i) - E\left[f_{t,\tau}^{+}(\boldsymbol{u}_i) | \mathcal{F}_{t-1} \right] \} + E\left[f_{t,\tau}^{+}(\boldsymbol{u}_i) - f_{t,\tau}^{-}(\boldsymbol{u}_i) | \mathcal{F}_{t-1} \right] \\ &\leqslant D_{t,\tau}^{+}(\boldsymbol{u}_i) + \sqrt{n}\epsilon\delta_k \frac{1}{nC_1} \sum_{t=1}^{n} \varrho_{t,\tau}, \end{aligned} \tag{S5.12}$$

where $D_{t,\tau}^{+}(\boldsymbol{u}_i) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{ f_{t,\tau}^{+}(\boldsymbol{u}_i) - E\left[f_{t,\tau}^{+}(\boldsymbol{u}_i) | \mathcal{F}_{t-1} \right] \}$. Define the event

$$E_n = \{ \omega : \frac{1}{nC_1} \sum_{t=1}^{n} \varrho_{t,\tau}(\omega) < 2 \}.$$

For $\boldsymbol{u} \in A(k)$, i.e. $\delta_{k+1} \leqslant \|\boldsymbol{u}\| \leqslant \delta_k$, we have $1 + \sqrt{n}\|\boldsymbol{u}\| > \sqrt{n}\delta_{k+1} = \sqrt{n}\delta_k/2$. Thus, by

(S5.12) and the Chebyshev's inequality, we have

$$Pr\left(\sup_{\boldsymbol{u}\in A(k)}\frac{D_{t,\tau}(\boldsymbol{u})}{1+\sqrt{n}\|\boldsymbol{u}\|}>6\epsilon,E_n\right)\leqslant Pr\left(\max_{1\leqslant i\leqslant K_\epsilon}\sup_{\boldsymbol{u}\in U_i(\delta_k)\cap A(k)}D_{t,\tau}(\boldsymbol{u})>3\sqrt{n}\epsilon\delta_k,E_n\right)$$
$$\leqslant K_\epsilon\max_{1\leqslant i\leqslant K_\epsilon}Pr\left(D^+_{t,\tau}(\boldsymbol{u}_i)>\sqrt{n}\epsilon\delta_k\right)$$
$$\leqslant K_\epsilon\max_{1\leqslant i\leqslant K_\epsilon}\frac{E\left[(D^+_{t,\tau}(\boldsymbol{u}_i))^2\right]}{n\epsilon^2\delta_k^2}. \quad \text{(S5.13)}$$

Then by iterative-expectation, Taylor expansion and the Cauchy-Schwarz inequality, together with $\|\boldsymbol{u}_i\|\leqslant\delta_k$ for $\boldsymbol{u}_i\in U_i(\delta_k)$, we have

$$E\{[f^+_{t,\tau}(\boldsymbol{u}_i)]^2\}=E\{E\{[f^+_{t,\tau}(\boldsymbol{u}_i)]^2|\mathcal{F}_{t-1}\}\}$$
$$\leqslant 2E\{g^2_{t,\tau}\left|\int_0^1\left[F\left(b_\tau+\frac{\Delta_{t,\tau}(\boldsymbol{u}_i)}{h_t(\boldsymbol{\vartheta}_0)}s+\frac{\epsilon\delta}{C_1}\frac{\|\dot{q}_{t,\tau}(\boldsymbol{\phi}_0)\|}{h_t(\boldsymbol{\vartheta}_0)}\right)-F(b_\tau)\right]ds\right|\}$$
$$\leqslant C\delta_k\sup_x f(x)E\left(\sup_{\boldsymbol{\phi}^\star\in\Phi}\frac{\|\dot{q}_{t,\tau}(\boldsymbol{\phi}^\star)\|^3}{h_t(\boldsymbol{\vartheta}_0)}\right)$$
$$:=\pi_n(\delta_k),$$

where $\boldsymbol{\phi}^\star$ is between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0+\boldsymbol{u}_i$. This, together with $E\left(h_t^{-1}(\boldsymbol{\vartheta}_0)\sup_\Phi\|\dot{q}_{t,\tau}(\boldsymbol{\phi})\|^3\right)<\infty$ by Assumption 5 (ii), $\sup_x f(x)<\infty$ by Assumption 3 and the fact that $f^+_{t,\tau}(\boldsymbol{u}_i)-E\left[f^+_{t,\tau}(\boldsymbol{u}_i)|\mathcal{F}_{t-1}\right]$ is a martingale difference sequence, implies that

$$E\{\left[D^+_{t,\tau}(\boldsymbol{u}_i)\right]^2\}=\frac{1}{n}\sum_{t=1}^n E\{\{f^+_{t,\tau}(\boldsymbol{u}_i)-E\left[f^+_{t,\tau}(\boldsymbol{u}_i)|\mathcal{F}_{t-1}\right]\}^2\}$$
$$\leqslant\frac{1}{n}\sum_{t=1}^n E\{\left[f^+_{t,\tau}(\boldsymbol{u}_i)\right]^2\}\leqslant\pi_n(\delta_k)<\infty \quad \text{(S5.14)}$$

Combining with (S5.13) and (S5.14), we have

$$Pr\left(\sup_{\boldsymbol{u}\in A(k)}\frac{D_{t,\tau}(\boldsymbol{u})}{1+\sqrt{n}\|\boldsymbol{u}\|}>6\epsilon,E_n\right)\leqslant\frac{K_\epsilon\pi_n(\delta_k)}{n\epsilon^2\delta_k^2},$$

Similar to the proof of the upper tail case, we can obtain the same bound for the lower tail case. Therefore, it holds that

$$Pr\left(\sup_{\boldsymbol{u}\in A(k)}\frac{|D_{t,\tau}(\boldsymbol{u})|}{1+\sqrt{n}\|\boldsymbol{u}\|}>6\epsilon,E_n\right)\leqslant\frac{2K_\epsilon\pi_n(\delta_k)}{n\epsilon^2\delta_k^2}. \quad \text{(S5.15)}$$

Since $\pi_n(\delta_k) \to 0$ as $k \to \infty$, we can choose k_ϵ such that $2K_\epsilon \pi_n(\delta_k)/(\epsilon^2\eta^2) < \epsilon$ for $k \geqslant k_\epsilon$. Let k_n be an integer satisfies $n^{-1/2} \leqslant 2^{-k_n} \leqslant 2n^{-1/2}$. Split $\{\boldsymbol{u} : \|\boldsymbol{u}\| \leqslant \eta\}$ into two sets $B := B(k_n+1)$ and $B^c := B(0) - B(k_n+1) = \bigcup_{k=0}^{k_n} A(k)$. Then by (S5.15), we have

$$\begin{aligned} Pr\left(\sup_{\boldsymbol{u}\in B^c}\frac{|D_{t,\tau}(\boldsymbol{u})|}{1+\sqrt{n}\|\boldsymbol{u}\|} > 6\epsilon\right) &\leqslant \sum_{k=0}^{k_n} Pr\left(\sup_{\boldsymbol{u}\in A(k)}\frac{|D_{t,\tau}(\boldsymbol{u})|}{1+\sqrt{n}\|\boldsymbol{u}\|} > 6\epsilon, E_n\right) + Pr\left(E_n^c\right) \\ &\leqslant \frac{1}{n}\sum_{k=0}^{k_\epsilon-1}\frac{2K_\epsilon\pi_n(\delta_k)}{\epsilon^2\eta^2}2^{2k} + \frac{\epsilon}{n}\sum_{k=k_\epsilon}^{k_n}2^{2k} + Pr(E_n^c) \\ &\leqslant O(\frac{1}{n}) + 4\epsilon + Pr(E_n^c). \end{aligned} \tag{S5.16}$$

Since $1+\sqrt{n}\|\boldsymbol{u}\| > 1$ for $\boldsymbol{u} \in B$ and $\sqrt{n}\delta_{k_n+1} < 1$, similar to the proof of (S5.13) and (S5.14), we have

$$Pr\left(\sup_{\boldsymbol{u}\in B}\frac{D_{t,\tau}(\boldsymbol{u})}{1+\sqrt{n}\|\boldsymbol{u}\|} > 3\epsilon, E_n\right) \leqslant Pr\left(\max_{1\leqslant i\leqslant K_\epsilon} D_{t,\tau}^+(\boldsymbol{u}_i) > \epsilon, E_n\right) \leqslant \frac{K_\epsilon\pi_n(\delta_{k_n+1})}{\epsilon^2}.$$

We can get the same bound for the lower tail. Therefore, we have

$$\begin{aligned} Pr\left(\sup_{\boldsymbol{u}\in B}\frac{|D_{t,\tau}(\boldsymbol{u})|}{1+\sqrt{n}\|\boldsymbol{u}\|} > 3\epsilon\right) &\leqslant Pr\left(\sup_{\boldsymbol{u}\in B}\frac{|D_{t,\tau}(\boldsymbol{u})|}{1+\sqrt{n}\|\boldsymbol{u}\|} > 3\epsilon, E_n\right) + Pr\left(E_n^c\right) \\ &\leqslant \frac{2K_\epsilon\pi_n(\delta_{k_n+1})}{\epsilon^2} + Pr(E_n^c). \end{aligned} \tag{S5.17}$$

Note that $\pi_n(\delta_{k_n+1}) \to 0$ as $n \to \infty$. Furthermore, by the ergodic theorem, $Pr(E_n) \to 1$ and thus $Pr(E_n^c) \to 0$ as $n \to \infty$. Finally, (S5.9) follows by (S5.16) and (S5.17). The proof of this lemma is accomplished. □

Lemma 3. *If Assumptions 1-5 hold, then for any sequence of random variables* $\boldsymbol{u}_n$ *such that* $\boldsymbol{u}_n = o_p(1)$, *it holds that*

$$n\left[L_n(\boldsymbol{\phi}_0+\boldsymbol{u}_n) - L_n(\boldsymbol{\phi}_0)\right] = -\sqrt{n}\boldsymbol{u}_n'\boldsymbol{T}_n + \sqrt{n}\boldsymbol{u}_n'\Sigma_1\sqrt{n}\boldsymbol{u}_n + o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2),$$

where $L_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^n\sum_{k=1}^K \rho_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}))$, $\boldsymbol{T}_n = n^{-1/2}\sum_{t=1}^n\sum_{k=1}^K \dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\psi_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0))$, *and* $\Sigma_1 = \Sigma/2 = \sum_{k=1}^K f(b_{\tau_k})E\left[h_t^{-1}(\boldsymbol{\vartheta}_0)\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\dot{q}_{t,\tau_k}'(\boldsymbol{\phi}_0)\right]/2.$

Proof. Denote $\boldsymbol{u} = \boldsymbol{\phi} - \boldsymbol{\phi}_0$, where $\boldsymbol{\phi} = (\boldsymbol{\vartheta}', \boldsymbol{b}')'$ and $\boldsymbol{\phi}_0 = (\boldsymbol{\vartheta}_0', \boldsymbol{b}_0')'$. Recall that $L_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^n \sum_{k=1}^K \rho_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}))$. Let $X_{t,\tau}(\boldsymbol{u}) = \int_0^1 [I(\xi_{t,\tau} \leqslant \Delta_{t,\tau}(\boldsymbol{u})s) - I(\xi_{t,\tau} \leqslant 0)]\, ds$ with $\Delta_{t,\tau}(\boldsymbol{u}) = q_{t,\tau}(\boldsymbol{\phi}_0 + \boldsymbol{u}) - q_{t,\tau}(\boldsymbol{\phi}_0)$ and $\xi_{t,\tau} = y_t - q_{t,\tau}(\boldsymbol{\phi}_0)$. By the Knight identity (S5.3), it holds that

$$n[L_n(\boldsymbol{\phi}_0 + \boldsymbol{u}) - L_n(\boldsymbol{\phi}_0)] = \sum_{t=1}^n \sum_{k=1}^K [\rho_{\tau_k}(\xi_{t,\tau_k} - \Delta_{t,\tau_k}(\boldsymbol{u})) - \rho_{\tau_k}(\xi_{t,\tau_k})]$$

$$= K_{1n}(\boldsymbol{u}) + K_{2n}(\boldsymbol{u}), \tag{S5.18}$$

where $\boldsymbol{u} \in \aleph \equiv \{\boldsymbol{u} \in \mathbb{R}^d : \boldsymbol{u} + \boldsymbol{\phi}_0 \in \Phi\}$,

$$K_{1n}(\boldsymbol{u}) = -\sum_{t=1}^n \sum_{k=1}^K \Delta_{t,\tau_k}(\boldsymbol{u})\psi_{\tau_k}(\xi_{t,\tau_k}), \quad \text{and} \quad K_{2n}(\boldsymbol{u}) = \sum_{t=1}^n \sum_{k=1}^K \Delta_{t,\tau_k}(\boldsymbol{u}) X_{t,\tau_k}(\boldsymbol{u}).$$

By Taylor expansion, we have $\Delta_{t,\tau}(\boldsymbol{u}) = \boldsymbol{u}'\dot{q}_{t,\tau}(\boldsymbol{\phi}_0) + \boldsymbol{u}'\ddot{q}_{t,\tau}(\boldsymbol{\phi}^\star)\boldsymbol{u}/2$, where $\boldsymbol{\phi}^\star$ is between $\boldsymbol{\phi}_0 + \boldsymbol{u}$ and $\boldsymbol{\phi}_0$. Then,

$$K_{1n}(\boldsymbol{u}) = -\boldsymbol{u}' \sum_{t=1}^n \sum_{k=1}^K \dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\psi_{\tau_k}(\xi_{t,\tau_k}) - \boldsymbol{u}' \sum_{k=1}^K \sum_{t=1}^n \ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\psi_{\tau_k}(\xi_{t,\tau_k})\boldsymbol{u}$$

$$= -\sqrt{n}\boldsymbol{u}'\boldsymbol{T}_n - \sqrt{n}\boldsymbol{u}' R_{1n}(\boldsymbol{\phi}^\star)\sqrt{n}\boldsymbol{u}, \tag{S5.19}$$

where

$$\boldsymbol{T}_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n \sum_{k=1}^K \dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\psi_{\tau_k}(\xi_{t,\tau_k}) \text{ and } R_{1n}(\boldsymbol{\phi}^\star) = \frac{1}{n} \sum_{k=1}^K \sum_{t=1}^n \ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\psi_{\tau_k}(\xi_{t,\tau_k}).$$

Since $E\left[\sup_{\boldsymbol{\phi}^\star \in \Phi} \|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|\right] < \infty$ by Assumption 5 (iii) and the fact that $|\psi_{\tau_k}(\xi_{t,\tau_k})| \leqslant 1$, we have

$$E\left[\sup_{\boldsymbol{\phi}^\star \in \Phi} \|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\psi_{\tau_k}(\xi_{t,\tau_k})\|\right] < \infty.$$

Then by iterative-expectation and the fact $E[\psi_{\tau_k}(\xi_{t,\tau_k})|\mathcal{F}_{t-1}] = 0$, it holds that

$$E[\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\psi_{\tau_k}(\xi_{t,\tau_k})] = 0.$$

Therefore, by Theorem 3.1 in Ling and McAleer (2003), we have

$$\sup_{\boldsymbol{\phi}^\star\in\Phi}\|R_{1n}(\boldsymbol{\phi}^\star)\| = o_p(1). \tag{S5.20}$$

This together with (S5.19), implies that

$$K_{1n}(\boldsymbol{u}_n) = -\sqrt{n}\boldsymbol{u}_n'\boldsymbol{T}_n + o_p(n\|\boldsymbol{u}_n\|^2). \tag{S5.21}$$

Denote $X_{t,\tau}(\boldsymbol{u}) = X_{1t,\tau}(\boldsymbol{u}) + X_{2t,\tau}(\boldsymbol{u})$, where

$$X_{1t,\tau}(\boldsymbol{u}) = \int_0^1 \left[I(\xi_{t,\tau} \leqslant \Delta_{t,\tau}(\boldsymbol{u})s) - I(\xi_{t,\tau} \leqslant \boldsymbol{u}'\dot{q}_{t,\tau}(\boldsymbol{u})s)\right] ds,$$

and

$$X_{2t,\tau}(\boldsymbol{u}) = \int_0^1 \left[I(\xi_{t,\tau} \leqslant \boldsymbol{u}'\dot{q}_{t,\tau}(\boldsymbol{u})s) - I(\xi_{t,\tau} \leqslant 0)\right] ds.$$

For $K_{2n}(\boldsymbol{u})$, by Taylor expansion, it holds that

$$K_{2n}(\boldsymbol{u}) = R_{2n}(\boldsymbol{u}) + R_{3n}(\boldsymbol{u}) + R_{4n}(\boldsymbol{u}) + R_{5n}(\boldsymbol{u}), \tag{S5.22}$$

where

$$\begin{aligned}
R_{2n}(\boldsymbol{u}) &= \boldsymbol{u}'\sum_{k=1}^{K}\sum_{t=1}^{n}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\{X_{t,\tau_k}(\boldsymbol{u}) - E\left[X_{t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right]\},\\
R_{3n}(\boldsymbol{u}) &= \boldsymbol{u}'\sum_{k=1}^{K}\sum_{t=1}^{n}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)E\left[X_{1t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right],\\
R_{4n}(\boldsymbol{u}) &= \boldsymbol{u}'\sum_{k=1}^{K}\sum_{t=1}^{n}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)E\left[X_{2t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right],\\
R_{5n}(\boldsymbol{u}) &= \frac{1}{2}\boldsymbol{u}'\sum_{k=1}^{K}\sum_{t=1}^{n}\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)X_{t,\tau_k}(\boldsymbol{u})\boldsymbol{u}.
\end{aligned}$$

For $R_{2n}(\boldsymbol{u})$, by Lemma S.1, it holds that

$$R_{2n}(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}\|^2). \tag{S5.23}$$

For $R_{3n}(\boldsymbol{u})$, note that

$$E\left[X_{1t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right] = \int_0^1 \left[F(b_{\tau_k} + \frac{\Delta_{t,\tau_k}(\boldsymbol{u})s}{h_t(\boldsymbol{\vartheta}_0)}) - F(b_{\tau_k} + \frac{\boldsymbol{u}'\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)s}{h_t(\boldsymbol{\vartheta}_0)})\right] ds. \tag{S5.24}$$

Then by iterative-expectation, Taylor expansion and the Cauchy-Schwarz inequality, together with $\sup_x f(x) < \infty$ by Assumption 3, $E\left[h_t^{-1}(\boldsymbol{\vartheta}_0)\sup_{\boldsymbol{\phi}\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^2\right] < \infty$ by Assumption 5 (ii), and $E\left[h_t^{-1}(\boldsymbol{\vartheta}_0)\sup_{\boldsymbol{\phi}\in\Phi}\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^2\right] < \infty$ by Assumption 5 (iii), for any $\eta > 0$, it holds that

$$\begin{aligned}
E\left(\sup_{\|\boldsymbol{u}\|\leqslant\eta}\frac{|R_{3n}(\boldsymbol{u})|}{n\|\boldsymbol{u}\|^2}\right) &\leqslant \frac{1}{4}\eta\sup_x f(x)\sum_{k=1}^K E\left[\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\|\frac{\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|}{h_t(\boldsymbol{\vartheta}_0)}\right]\\
&\leqslant C\eta\sum_{k=1}^K\left\{E\left[\sup_{\boldsymbol{\phi}\in\Phi}\frac{\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^2}{h_t(\boldsymbol{\vartheta}_0)}\right]\right\}^{1/2}\left\{E\left[\sup_{\boldsymbol{\phi}\in\Phi}\frac{\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^2}{h_t(\boldsymbol{\vartheta}_0)}\right]\right\}^{1/2}\\
&\to 0 \quad as \quad \eta\to 0.
\end{aligned}$$

Therefore, by Markov's theorem, for any $\epsilon, \delta > 0$, there exists $\eta_0 = \eta_0(\epsilon)$ such that

$$Pr\left(\sup_{\|\boldsymbol{u}\|\leqslant\eta_0}\frac{|R_{3n}(\boldsymbol{u})|}{n\|\boldsymbol{u}\|^2} > \delta\right) < \frac{\epsilon}{2} \tag{S5.25}$$

for all $n \geqslant 1$. Since $\boldsymbol{u}_n = o_p(1)$, it follows that

$$Pr\left(\|\boldsymbol{u}_n\| > \eta_0\right) < \frac{\epsilon}{2} \tag{S5.26}$$

as n is large enough. From (S5.25) and (S5.26), we have

$$\begin{aligned}
Pr\left(\frac{|R_{3n}(\boldsymbol{u}_n)|}{n\|\boldsymbol{u}_n\|^2} > \delta\right) &= Pr\left(\frac{|R_{3n}(\boldsymbol{u}_n)|}{n\|\boldsymbol{u}_n\|^2} > \delta, \|\boldsymbol{u}_n\| \leqslant \eta_0\right) + Pr\left(\frac{|R_{3n}(\boldsymbol{u}_n)|}{n\|\boldsymbol{u}_n\|^2} > \delta, \|\boldsymbol{u}_n\| > \eta_0\right)\\
&\leqslant Pr\left(\sup_{\|\boldsymbol{u}_n\|\leqslant\eta_0}\frac{|R_{3n}(\boldsymbol{u}_n)|}{n\|\boldsymbol{u}\|^2} > \delta\right) + Pr\left(\|\boldsymbol{u}_n\| > \eta_0\right) < \epsilon
\end{aligned}$$

as n is large enough. Therefore,

$$R_{3n}(\boldsymbol{u}_n) = o_p(n\|\boldsymbol{u}_n\|^2). \tag{S5.27}$$

For $R_{4n}(\boldsymbol{u})$, note that

$$E\left[X_{2t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right] = \int_0^1\left[F(b_{\tau_k} + \frac{\boldsymbol{u}'\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)s}{h_t(\boldsymbol{\vartheta}_0)}) - F(b_{\tau_k})\right]ds. \tag{S5.28}$$

Then by Taylor expansion and Assumption 3, it follows that

$$\begin{aligned} E\left[X_{2t,\tau_k}(\boldsymbol{u})|\mathcal{F}_{t-1}\right] = & \frac{\boldsymbol{u}'}{2} f(b_{\tau_k}) \frac{\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)}{h_t(\boldsymbol{\vartheta}_0)} \\ & + \boldsymbol{u}' \frac{\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)}{h_t(\boldsymbol{\vartheta}_0)} \int_0^1 \left[f(b_{\tau_k} + \frac{\boldsymbol{u}'\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)}{h_t(\boldsymbol{\vartheta}_0)} s^\star) - f(b_{\tau_k}) \right] sds, \end{aligned}$$

where $s^\star$ is between 0 and s. Therefore, it holds that

$$R_{4n}(\boldsymbol{u}) = \sqrt{n}\boldsymbol{u}'\Sigma_{1n}\sqrt{n}\boldsymbol{u} + \sqrt{n}\boldsymbol{u}'\pi_{3n}(\boldsymbol{u})\sqrt{n}\boldsymbol{u}, \tag{S5.29}$$

where $\Sigma_{1n} = (2n)^{-1}\sum_{k=1}^K f(b_{\tau_k})\sum_{t=1}^n h_t^{-1}(\boldsymbol{\vartheta}_0)\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\dot{q}'_{t,\tau_k}(\boldsymbol{\phi}_0)$ and

$$\pi_{3n}(\boldsymbol{u}) = \frac{1}{n}\sum_{k=1}^K\sum_{t=1}^n \frac{1}{h_t(\boldsymbol{\vartheta}_0)}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\dot{q}'_{t,\tau_k}(\boldsymbol{\phi}_0)\int_0^1 \left[f(b_{\tau_k} + \frac{\boldsymbol{u}'\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)}{h_t(\boldsymbol{\vartheta}_0)} s^\star) - f(b_{\tau_k}) \right] sds.$$

By Taylor expansion, together with $E\left[h_t^{-1}(\boldsymbol{\vartheta}_0)\sup_{\boldsymbol{\phi}\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^3\right] < \infty$ by Assumption 5 (ii) and $\sup_x\left|\dot{f}(x)\right| < \infty$ by Assumption 3, for any $\eta > 0$, it holds that

$$\begin{aligned} E\left(\sup_{\|\boldsymbol{u}\|\leqslant\eta}\|\pi_{3n}(\boldsymbol{u})\|\right) & \leqslant \frac{1}{2n}\sup_x\left|\dot{f}(x)\right|\sum_{k=1}^K\sum_{t=1}^n E\left(\sup_{\|\boldsymbol{u}\|\leqslant\eta}\left\|\frac{1}{h_t(\boldsymbol{\vartheta}_0)}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\dot{q}'_{t,\tau_k}(\boldsymbol{\phi}_0)\ \frac{\boldsymbol{u}'\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)}{h_t(\boldsymbol{\vartheta}_0)}\right\|\right) \\ & \leqslant C\eta\sup_x\left|\dot{f}(x)\right|\sum_{k=1}^K E\left[\sup_{\boldsymbol{\phi}\in\Phi}\frac{\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^3}{h_t^2(\boldsymbol{\vartheta}_0)}\right] \to 0 \quad as \quad \eta \to 0. \end{aligned}$$

Similar to (S5.25) and (S5.26), we can show that $\pi_{3n}(\boldsymbol{u}_n) = o_p(1)$. This together with (S5.29) and $\Sigma_1 = \Sigma_{1n} + o_p(1)$ by ergodic theorem, implies that

$$R_{4n}(\boldsymbol{u}_n) = \sqrt{n}\boldsymbol{u}_n'\Sigma_1\sqrt{n}\boldsymbol{u}_n + o_p(n\|\boldsymbol{u}_n\|^2), \tag{S5.30}$$

where $\Sigma_1 = \sum_{k=1}^K f(b_{\tau_k})E\left[h_t^{-1}(\boldsymbol{\vartheta}_0)\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\dot{q}'_{t,\tau_k}(\boldsymbol{\phi}_0)\right]/2$.

Finally, we consider $R_{5n}(\boldsymbol{u})$. Since $I(x \leqslant a) - I(x \leqslant b) = I(b \leqslant x \leqslant a) - I(b \geqslant x \geqslant a)$ and $\Delta_{t,\tau}(\boldsymbol{u}) = \boldsymbol{u}'\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^*)$ with $\boldsymbol{\phi}^*$ between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0 + \boldsymbol{u}$. Then by Taylor expansion, we have

$$\begin{aligned}\sup_{\|\boldsymbol{u}\|\leqslant\eta}|X_{t,\tau_k}(\boldsymbol{u})|\leqslant&\int_0^1\sup_{\|\boldsymbol{u}\|\leqslant\eta}\left|I\left(b_{\tau_k}\leqslant\eta_t\leqslant b_{\tau_k}+\frac{\Delta_{t,\tau_k}(\boldsymbol{u})}{h_t(\boldsymbol{\vartheta}_0)}s\right)\right|ds\\&+\int_0^1\sup_{\|\boldsymbol{u}\|\leqslant\eta}\left|I\left(b_{\tau_k}\geqslant\eta_t\geqslant b_{\tau_k}+\frac{\Delta_{t,\tau_k}(\boldsymbol{u})}{h_t(\boldsymbol{\vartheta}_0)}s\right)\right|ds\\\leqslant&I\left(b_{\tau_k}\leqslant\eta_t\leqslant b_{\tau_k}+\eta\sup_{\boldsymbol{\phi}^*\in\Phi}\frac{\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^*)\|}{h_t(\boldsymbol{\vartheta}_0)}\right)\\&+I\left(b_{\tau_k}\geqslant\eta_t\geqslant b_{\tau_k}-\eta\sup_{\boldsymbol{\phi}^*\in\Phi}\frac{\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^*)\|}{h_t(\boldsymbol{\vartheta}_0)}\right).\end{aligned}$$

Then by iterative-expectation and the Cauchy-Schwarz inequality, together with $\sup_x\left|\dot{f}(x)\right|<\infty$ by Assumption 3, $E\left[h_t^{-1}(\boldsymbol{\vartheta}_0)\sup_{\boldsymbol{\phi}\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^2)\right]<\infty$ by Assumption 5 (ii), and $E[h_t^{-1}(\boldsymbol{\vartheta}_0)$ $\sup_{\boldsymbol{\phi}\in\Phi}\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi})\|^2]<\infty$ by Assumption 5 (iii), for any $\eta>0$, it follows that

$$\begin{aligned}E\left(\sup_{\|\boldsymbol{u}\|\leqslant\eta}\frac{|R_{5n}(\boldsymbol{u})|}{n\|\boldsymbol{u}\|^2}\right)\leqslant&\frac{1}{2n}\sum_{k=1}^K\sum_{t=1}^nE\left[\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|E\left(\sup_{\|\boldsymbol{u}\|\leqslant\eta}|X_{t,\tau_k}(\boldsymbol{u})|\,|\mathcal{F}_{t-1}\right)\right]\\\leqslant&\eta\sup_xf(x)\sum_{k=1}^KE\left[\sup_{\boldsymbol{\phi}^\star\in\Phi}\frac{\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|}{h_t(\boldsymbol{\vartheta}_0)}\sup_{\boldsymbol{\phi}^*\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^*)\|\right]\\\leqslant&C\eta\sum_{k=1}^K\left\{E\left[\sup_{\boldsymbol{\phi}^*\in\Phi}\frac{\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^*)\|^2}{h_t(\boldsymbol{\vartheta}_0)}\right]\right\}^{1/2}\left\{E\left[\sup_{\boldsymbol{\phi}^\star\in\Phi}\frac{\|\ddot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|^2}{h_t(\boldsymbol{\vartheta}_0)}\right]\right\}^{1/2}\\\to&0\quad as\quad\eta\to0.\end{aligned}$$

Similar to (S5.25) and (S5.26), we can show that

$$R_{5n}(\boldsymbol{u}_n)=o_p(n\|\boldsymbol{u}_n\|^2).\tag{S5.31}$$

From (S5.23), (S5.27), (S5.30) and (S5.31), we have

$$K_{2n}(\boldsymbol{u}_n)=\sqrt{n}\boldsymbol{u}_n'\Sigma_1\sqrt{n}\boldsymbol{u}_n+o_p(\sqrt{n}\|\boldsymbol{u}_n\|+n\|\boldsymbol{u}_n\|^2).\tag{S5.32}$$

In view of (S5.18), (S5.21) and (S5.32), we accomplish the proof of this lemma. □

Lemma 4. *If Assumptions 1-6 hold, then for any sequence of random variables $\boldsymbol{u}_n$ such that*

$\boldsymbol{u}_n = o_p(1)$, *it holds that*

$$n\left[\widetilde{L}_n(\boldsymbol{\phi}_0+\boldsymbol{u}_n)-\widetilde{L}_n(\boldsymbol{\phi}_0)\right]-n\left[L_n(\boldsymbol{\phi}_0+\boldsymbol{u}_n)-L_n(\boldsymbol{\phi}_0)\right]=o_p(\sqrt{n}\|\boldsymbol{u}_n\|+n\|\boldsymbol{u}_n\|^2),$$

where $\widetilde{L}_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^n\sum_{k=1}^K \rho_{\tau_k}(y_t-\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}))$ *and* $L_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^n\sum_{k=1}^K \rho_{\tau_k}(y_t-q_{t,\tau_k}(\boldsymbol{\phi}))$.

Proof. Recall that $\xi_{t,\tau_k} = y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0)$, $\Delta_{t,\tau_k}(\boldsymbol{u}) = q_{t,\tau_k}(\boldsymbol{\phi}_0+\boldsymbol{u}) - q_{t,\tau_k}(\boldsymbol{\phi}_0)$ and $X_{t,\tau_k}(\boldsymbol{u}) = \int_0^1 \left[I(y_t \leqslant q_{t,\tau_k}(\boldsymbol{\phi}_0)+\Delta_{t,\tau_k}(\boldsymbol{u})s) - I(y_t \leqslant q_{t,\tau_k}(\boldsymbol{\phi}_0))\right]ds$. Define $\widetilde{\xi}_{t,\tau_k} = y_t - \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)$, $\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u}) = \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0+\boldsymbol{u}) - \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)$ and $\widetilde{X}_{t,\tau_k}(\boldsymbol{u}) = \int_0^1 \left[I(y_t \leqslant \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)+\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})s) - I(y_t \leqslant \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0))\right]ds$. Then by (S5.3), it holds

$$\begin{aligned}
&n\left[\widetilde{L}_n(\boldsymbol{\phi}_0+\boldsymbol{u}_n)-\widetilde{L}_n(\boldsymbol{\phi}_0)\right]-n\left[L_n(\boldsymbol{\phi}_0+\boldsymbol{u}_n)-L_n(\boldsymbol{\phi}_0)\right]\\
=&\sum_{t=1}^n\sum_{k=1}^K\left\{\left[-\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})\psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})+\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})\widetilde{X}_{t,\tau_k}(\boldsymbol{u})\right]-\left[-\Delta_{t,\tau_k}(\boldsymbol{u})\psi_{\tau_k}(\xi_{t,\tau_k})+\Delta_{t,\tau_k}(\boldsymbol{u})X_{t,\tau_k}(\boldsymbol{u})\right]\right\}\\
=&\sum_{k=1}^K\left[\widetilde{A}_{k,1n}(\boldsymbol{u})+\widetilde{A}_{k,2n}(\boldsymbol{u})+\widetilde{A}_{k,3n}(\boldsymbol{u})+\widetilde{A}_{k,4n}(\boldsymbol{u})\right], \qquad \text{(S5.33)}
\end{aligned}$$

where $\boldsymbol{u} \in \aleph \equiv \{\boldsymbol{u} : \boldsymbol{u}+\boldsymbol{\phi}_0 \in \Phi\}$,

$$\begin{aligned}
\widetilde{A}_{k,1n}(\boldsymbol{u}) &= \sum_{t=1}^n\left[\Delta_{t,\tau_k}(\boldsymbol{u})-\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})\right]\psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k}),\\
\widetilde{A}_{k,2n}(\boldsymbol{u}) &= \sum_{t=1}^n\left[\psi_{\tau_k}(\xi_{t,\tau_k})-\psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})\right]\Delta_{t,\tau_k}(\boldsymbol{u}),\\
\widetilde{A}_{k,3n}(\boldsymbol{u}) &= \sum_{t=1}^n\left[\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})-\Delta_{t,\tau_k}(\boldsymbol{u})\right]\widetilde{X}_{t,\tau_k}(\boldsymbol{u}), \quad \text{and}\\
\widetilde{A}_{k,4n}(\boldsymbol{u}) &= \sum_{t=1}^n\left[\widetilde{X}_{t,\tau_k}(\boldsymbol{u})-X_{t,\tau_k}(\boldsymbol{u})\right]\Delta_{t,\tau_k}(\boldsymbol{u}).
\end{aligned}$$

For $\widetilde{A}_{k,1n}(\boldsymbol{u})$, since $|\psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})| \leqslant 1$, $\sum_{t=1}^\infty \sup_{\boldsymbol{\phi}\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi})-\dot{\widetilde{q}}_{t,\tau_k}(\boldsymbol{\phi})\| < \infty$ by Assumption 6 (ii), then by Taylor expansion, it holds

$$\begin{aligned}
\sup_{\boldsymbol{u}\in\aleph}\frac{|\widetilde{A}_{k,1n}(\boldsymbol{u})|}{\sqrt{n}\|\boldsymbol{u}\|} &\leqslant \frac{1}{\sqrt{n}}\sum_{t=1}^n\sup_{\boldsymbol{u}\in\aleph}\frac{|\Delta_{t,\tau_k}(\boldsymbol{u})-\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})|}{\|\boldsymbol{u}\|}|\psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})|\\
&\leqslant \frac{1}{\sqrt{n}}\sum_{t=1}^n\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)-\dot{\widetilde{q}}_{t,\tau_k}(\boldsymbol{\phi}^\star)\| = o_p(1),
\end{aligned}$$

where $\boldsymbol{\phi}^\star$ is between $\boldsymbol{\phi}_0$ and $\boldsymbol{\phi}_0 + \boldsymbol{u}$. Therefore, for $\boldsymbol{u}_n = o_p(1)$, we have

$$\widetilde{A}_{k,1n}(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\|). \tag{S5.34}$$

For $\widetilde{A}_{k,2n}(\boldsymbol{u})$, by Taylor expansion and Cauchy-Schwarz inequality, together with the strict stationarity and ergodicity of $\{y_t\}$ by Assumption 1, $E\left[\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|\right] < \infty$ by Assumption 5 (ii), it holds that

$$\begin{aligned}\sup_{\boldsymbol{u}\in\aleph}\frac{|\widetilde{A}_{k,2n}(\boldsymbol{u})|}{\sqrt{n}\|\boldsymbol{u}\|} &\leqslant \frac{1}{\sqrt{n}}\sum_{t=1}^{n}\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\||\psi_{\tau_k}(\xi_{t,\tau_k}) - \psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})| \\ &\leqslant \frac{1}{\sqrt{n}}\max_{1\leqslant t\leqslant n}\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|\sum_{t=1}^{n}|\psi_{\tau_k}(\xi_{t,\tau_k}) - \psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})| \\ &\leqslant o_p(1)\sum_{t=1}^{\infty}|\psi_{\tau_k}(\xi_{t,\tau_k}) - \psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})|. \end{aligned} \tag{S5.35}$$

Since $I(x<a) - I(x<b) = I(0 < x-b < a-b) - I(0 > x-b > a-b)$ and $\psi_{\tau_k}(\xi_{t,\tau_k}) - \psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k}) = I(y_t < \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - I(y_t < q_{t,\tau_k}(\boldsymbol{\phi}_0))$, we have

$$\begin{aligned}E\left[|\psi_{\tau_k}(\xi_{t,\tau_k}) - \psi_{\tau_k}(\widetilde{\xi}_{t,\tau_k})||\mathcal{F}_{t-1}\right] \leqslant & E\left[I(0 < y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0) < |\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|)|\mathcal{F}_{t-1}\right] \\ & + E\left[I(0 > y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0) > -|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|)|\mathcal{F}_{t-1}\right] \\ \leqslant & F(b_{\tau_k} + \frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)}) \\ & - F(b_{\tau_k} - \frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)}) \\ \leqslant & 2\sup_x f(x)\frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)}.\end{aligned}$$

This together with $\sup_x f(x) < \infty$ by Assumption 3, $\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0)\sup_{\boldsymbol{\phi}\in\Phi}|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}) - q_{t,\tau_k}(\boldsymbol{\phi})| < \infty$ by Assumption 6 (i), and Corollary 2.3 in Hall and Heyde (2014), we have

$$\sup_{\boldsymbol{u}\in\aleph}\frac{|\widetilde{A}_{k,2n}(\boldsymbol{u})|}{\sqrt{n}\|\boldsymbol{u}\|} \leqslant o_p(1)\sum_{t=1}^{\infty}\frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)} \leqslant o_p(1).$$

Therefore, for $\boldsymbol{u}_n = o_p(1)$, we have

$$\widetilde{A}_{k,2n}(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\|). \tag{S5.36}$$

For $\widetilde{A}_{k,3n}(\boldsymbol{u})$, since $|\widetilde{X}_{t,\tau_k}(\boldsymbol{u})| \leqslant 2$, then similar to the proof of $\widetilde{A}_{k,1n}(\boldsymbol{u})$, for $\boldsymbol{u}_n = o_p(1)$, we have

$$\widetilde{A}_{k,3n}(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\|). \tag{S5.37}$$

We finally consider $\widetilde{A}_{k,4n}(\boldsymbol{u})$. Denote

$$\widetilde{c}_t = \int_0^1 \left[I(y_t \leqslant \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0) + \widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u})s) - I(y_t \leqslant q_{t,\tau_k}(\boldsymbol{\phi}_0) + \Delta_{t,\tau_k}(\boldsymbol{u})s)\right] ds,$$

and

$$\widetilde{d}_t = \int_0^1 \left[I(y_t \leqslant \widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)) - I(y_t \leqslant q_{t,\tau_k}(\boldsymbol{\phi}_0))\right] ds.$$

By using $I(x \leqslant a) - I(x \leqslant b) = I(0 \leqslant x - b \leqslant a - b) - I(0 \geqslant x - b \geqslant a - b)$ and Taylor expansion, it holds that

$$\begin{aligned}
& E\left(|\widetilde{c}_t||\mathcal{F}_{t-1}\right) \\
\leqslant & E\left[\int_0^1 I\left(|y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0) - \Delta_{t,\tau_k}(\boldsymbol{u})s)| \leqslant |\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0) - q_{t,\tau_k}(\boldsymbol{\phi}_0)| + |\widetilde{\Delta}_{t,\tau_k}(\boldsymbol{u}) - \Delta_{t,\tau_k}(\boldsymbol{u})|s\right) ds\right] \\
\leqslant & C \sup_x f(x) \left[\frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)} + \|\boldsymbol{u}\| \sup_{\boldsymbol{\phi}^\star \in \Phi} \frac{\|\dot{\widetilde{q}}_{t,\tau_k}(\boldsymbol{\phi}^\star) - \dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|}{h_t(\boldsymbol{\vartheta}_0)}\right],
\end{aligned}$$

and

$$\begin{aligned}
E\left(|\widetilde{d}_t||\mathcal{F}_{t-1}\right) \leqslant & E\left[\int_0^1 I\left(|y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0)| \leqslant |\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|\right) ds\right] \\
\leqslant & 2 \sup_x f(x) \frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0) - q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)}.
\end{aligned}$$

Then by iterative-expansion,Cauchy-Schwarz and Taylor expansion, together with $\widetilde{X}_{t,\tau_k}(\boldsymbol{u}) - X_{t,\tau_k}(\boldsymbol{u}) = \widetilde{c}_t - \widetilde{d}_t$, $\sup_x f(x) < \infty$ by Assumption 3, $E\left[\sup_{\boldsymbol{\phi}^\star \in \Phi} \|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|\right] < \infty$ by Assumption 5 (ii), $\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0) \sup_{\boldsymbol{\phi} \in \Phi} |\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}) - q_{t,\tau_k}(\boldsymbol{\phi})| < \infty$ by Assumption 6 (i), $\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0)$

$\sup_{\phi\in\Phi}\|\dot{\tilde{q}}_{t,\tau_k}(\boldsymbol{\phi}) - \dot{q}_{t,\tau_k}(\boldsymbol{\phi})\| < \infty$ by Assumption 6 (ii), and and Corollary 2.3 in Hall and Heyde (2014), we have

$$\begin{aligned}
\sup_{\boldsymbol{u}\in\aleph}\frac{|\widetilde{A}_{k,4n}(\boldsymbol{u})|}{\sqrt{n}\|\boldsymbol{u}\|+n\|\boldsymbol{u}\|^2} \leqslant & \sum_{t=1}^{n}\sup_{\boldsymbol{u}\in\aleph}\frac{\Delta_{t,\tau_k}(\boldsymbol{u})}{\|\boldsymbol{u}\|}\frac{|\widetilde{X}_{t,\tau_k}(\boldsymbol{u})-X_{t,\tau_k}(\boldsymbol{u})|}{\sqrt{n}+n\|\boldsymbol{u}\|} \\
\leqslant & \frac{C}{\sqrt{n}}\sum_{t=1}^{n}\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|\frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)-q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)} \\
& + \frac{C}{n}\sum_{t=1}^{n}\sup_{\boldsymbol{\phi}^\star\in\Phi}\|\dot{q}_{t,\tau_k}(\boldsymbol{\phi}^\star)\|\frac{|\dot{\tilde{q}}_{t,\tau_k}(\boldsymbol{\phi}_0)-\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)} \\
\leqslant & o_p(1)\left\{\sum_{t=1}^{\infty}\frac{|\widetilde{q}_{t,\tau_k}(\boldsymbol{\phi}_0)-q_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)}+\sum_{t=1}^{\infty}\frac{|\dot{\tilde{q}}_{t,\tau_k}(\boldsymbol{\phi}_0)-\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)|}{h_t(\boldsymbol{\vartheta}_0)}\right\} \\
= & o_p(1).
\end{aligned}$$

Therefore, for $\boldsymbol{u}_n = o_p(1)$, we have

$$\widetilde{A}_{k,4n}(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2). \tag{S5.38}$$

Combining (S5.33)-(S5.38), we accomplish the proof of this lemma. □

Proof of Theorem 2. Recall that $L_n(\boldsymbol{\phi}) = n^{-1}\sum_{t=1}^{n}\sum_{k=1}^{K}\rho_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}))$. For $\boldsymbol{u}\in\aleph\equiv\{\boldsymbol{u} : \boldsymbol{\phi}_0+\boldsymbol{u}\in\Phi\}$, define $\widetilde{H}_n(\boldsymbol{u}) = n\left[\widetilde{L}_n(\boldsymbol{\phi}_0+\boldsymbol{u}) - \widetilde{L}_n(\boldsymbol{\phi}_0)\right]$. Denote $\widehat{\boldsymbol{u}}_n = \widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0$. By the consistency of $\widehat{\boldsymbol{\phi}}_n$, it holds that $\widehat{\boldsymbol{u}}_n = o_p(1)$. Note that $\widehat{\boldsymbol{u}}_n$ is the minimizer of $\widetilde{H}_n(\boldsymbol{u})$, since $\widehat{\boldsymbol{\phi}}_n$ minimizes $\widetilde{L}_n(\boldsymbol{\phi})$. This together with Lemmas 1-3, implies that

$$\begin{aligned}
\widetilde{H}_n(\widehat{\boldsymbol{u}}_n) &= -\sqrt{n}\widehat{\boldsymbol{u}}_n'\boldsymbol{T}_n + \sqrt{n}\widehat{\boldsymbol{u}}_n'\Sigma_1\sqrt{n}\widehat{\boldsymbol{u}}_n + o_p(\sqrt{n}\|\widehat{\boldsymbol{u}}_n\| + n\|\widehat{\boldsymbol{u}}_n\|^2) \qquad \text{(S5.39)} \\
&\geqslant -\sqrt{n}\|\widehat{\boldsymbol{u}}_n\|\left[\|\boldsymbol{T}_n\| + o_p(1)\right] + n\|\widehat{\boldsymbol{u}}_n\|^2\left[\lambda_{\min} + o_p(1)\right],
\end{aligned}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $\Sigma_1 = \Sigma/2$ with Σ defined before Theorem 1, and $\boldsymbol{T}_n = n^{-1/2}\sum_{t=1}^{n}\sum_{k=1}^{K}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\psi_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0))$. Denote $\boldsymbol{Z}_t = \sum_{k=1}^{K}\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\psi_{\tau_k}(y_t - q_{t,\tau_k}(\boldsymbol{\phi}_0))$, then $\boldsymbol{T}_n = n^{-1/2}\sum_{t=1}^{n}\boldsymbol{Z}_t$. If Assumptions 1-3 hold, by the Central Limit Theorem, we have

$$\boldsymbol{T}_n \to_{\mathcal{L}} N(\mathbf{0}, \Omega) \quad as \quad n\to\infty,$$

where $\Omega = \sum_{k=1}^{K}\sum_{k'=1}^{K}\Gamma_{kk'}E\left[\dot{q}_{t,\tau_k}(\boldsymbol{\phi}_0)\dot{q}'_{t,\tau_{k'}}(\boldsymbol{\phi}_0)\right]$ with $\Gamma_{kk'} = \min(\tau_k,\tau_{k'})(1-\max(\tau_k,\tau_{k'}))$.

Since $\widetilde{H}_n(\widehat{\boldsymbol{u}}_n) \leqslant 0$, then we have

$$\sqrt{n}\|\widehat{\boldsymbol{u}}_n\| \leqslant [\lambda_{\min} + o_p(1)]^{-1}\left[\|\boldsymbol{T}_n\| + o_p(1)\right] = O_p(1). \tag{S5.40}$$

This together with the consistency of $\widehat{\boldsymbol{\phi}}_n$, verifies the $\sqrt{n}$-consistency of $\widehat{\boldsymbol{\phi}}_n$, i.e. $\sqrt{n}(\widehat{\boldsymbol{\phi}}_n - \boldsymbol{\phi}_0) = O_p(1)$. Let $\sqrt{n}\boldsymbol{u}_n^{\star} = \Sigma_1^{-1}\boldsymbol{T}_n/2 = \Sigma^{-1}\boldsymbol{T}_n$, then we have

$$\sqrt{n}\boldsymbol{u}_n^{\star} \to_{\mathcal{L}} N(\mathbf{0},\Xi) \quad as \quad n\to\infty,$$

where $\Xi = \Sigma^{-1}\Omega\Sigma^{-1}$. As a result, it is sufficient to show that $\sqrt{n}\widehat{\boldsymbol{u}}_n - \sqrt{n}\boldsymbol{u}_n^{\star} = o_p(1)$. By (S5.39) and (S5.40), we have

$$\begin{aligned}\widetilde{H}_n(\widehat{\boldsymbol{u}}_n) &= -\sqrt{n}\widehat{\boldsymbol{u}}_n'\boldsymbol{T}_n + \sqrt{n}\widehat{\boldsymbol{u}}_n'\Sigma_1\sqrt{n}\widehat{\boldsymbol{u}}_n + o_p(1)\\ &= -2\sqrt{n}\widehat{\boldsymbol{u}}_n'\Sigma_1\sqrt{n}\boldsymbol{u}_n^{\star} + \sqrt{n}\widehat{\boldsymbol{u}}_n'\Sigma_1\sqrt{n}\widehat{\boldsymbol{u}}_n + o_p(1),\end{aligned}$$

and

$$\begin{aligned}\widetilde{H}_n(\boldsymbol{u}_n^{\star}) &= -\sqrt{n}\boldsymbol{u}_n^{\star\prime}\boldsymbol{T}_n + \sqrt{n}\boldsymbol{u}_n^{\star\prime}\Sigma_1\sqrt{n}\boldsymbol{u}_n^{\star} + o_p(1)\\ &= -\sqrt{n}\boldsymbol{u}_n^{\star\prime}\Sigma_1\sqrt{n}\boldsymbol{u}_n^{\star} + o_p(1).\end{aligned}$$

It follows that

$$\begin{aligned}\widetilde{H}_n(\widehat{\boldsymbol{u}}_n) - \widetilde{H}_n(\boldsymbol{u}_n^{\star}) &= (\sqrt{n}\widehat{\boldsymbol{u}}_n - \sqrt{n}\boldsymbol{u}_n^{\star})'\Sigma_1(\sqrt{n}\widehat{\boldsymbol{u}}_n - \sqrt{n}\boldsymbol{u}_n^{\star}) + o_p(1)\\ &\geqslant \lambda_{\min}\|\sqrt{n}\widehat{\boldsymbol{u}}_n - \sqrt{n}\boldsymbol{u}_n^{\star}\|^2 + o_p(1).\end{aligned} \tag{S5.41}$$

Since $\widetilde{H}_n(\widehat{\boldsymbol{u}}_n) - H_n(\boldsymbol{u}_n^{\star}) \leqslant 0 \quad a.s.$, then (S5.41) implies that $\|\sqrt{n}\widehat{\boldsymbol{u}}_n - \sqrt{n}\boldsymbol{u}_n^{\star}\| = o_p(1)$. We verify the asymptotic normality of $\widehat{\boldsymbol{\phi}}_n$, and the proof is accomplished. □

S5.3 Proof of Corollary S1

For the consistency in Corollary S1 (i), the proof is the same as that for Theorem 1 and we omit the detailed proof. Similar to the proof of Theorem 2, we introduce Lemmas 5-7 below to show the asymptotic normality in Corollary S1 (ii). Since the proofs of Lemmas 5 and 7 are the same as those of Lemmas 2 and 4, we only verify Lemma 6.

Lemma 5. *If Assumptions 3 and 5-S3 hold, then for any sequence of random variables* $\boldsymbol{u}_n$ *such that* $\boldsymbol{u}_n = o_p(1)$, *it holds that*

$$\pi_{1n}^*(\boldsymbol{u}_n) = o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2),$$

where $\pi_{1n}^*(\boldsymbol{u}) = \boldsymbol{u}' \sum_{k=1}^K \sum_{t=1}^n \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\{X_{t,\tau_k}^*(\boldsymbol{u}) - E\left[X_{t,\tau_k}^*(\boldsymbol{u})|\mathcal{F}_{t-1}\right]\}$ *with*

$$X_{t,\tau_k}^*(\boldsymbol{u}) = \int_0^1 \left[I(y_t \leqslant g_{t,\tau_k}(\boldsymbol{\psi}_0^*) + \Delta_{t,\tau_k}^*(\boldsymbol{u})s) - I(y_t \leqslant g_{t,\tau_k}(\boldsymbol{\psi}_0^*))\right] ds$$

and $\Delta_{t,\tau_k}^*(\boldsymbol{u}) = g_{t,\tau_k}(\boldsymbol{\psi}_0^* + \boldsymbol{u}) - g_{t,\tau_k}(\boldsymbol{\psi}_0^*)$.

Lemma 6. *If Assumptions 3 and 5-S3 hold, then for any sequence of random variables* $\boldsymbol{u}_n$ *such that* $\boldsymbol{u}_n = o_p(1)$, *it holds that*

$$n\left[L_n^*(\boldsymbol{\psi}_0^* + \boldsymbol{u}_n) - L_n^*(\boldsymbol{\psi}_0^*)\right] = -\sqrt{n}\boldsymbol{u}_n'\boldsymbol{T}_n^* + \sqrt{n}\boldsymbol{u}_n'(N_1 - J_1^*)\sqrt{n}\boldsymbol{u}_n + o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2),$$

where $L_n^*(\boldsymbol{\psi}) = n^{-1}\sum_{t=1}^n \sum_{k=1}^K \rho_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}))$, $\boldsymbol{T}_n^* = n^{-1/2}\sum_{t=1}^n \sum_{k=1}^K \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))$, $N_1 = \sum_{k=1}^K f(Q_{\tau_k}(\boldsymbol{\lambda}_0))E[h_t^{-1}(\boldsymbol{\vartheta}_0)\dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\dot{g}_{t,\tau_k}'(\boldsymbol{\psi}_0^*)]/2$ *and* $J_1^* = \sum_{k=1}^K E[\ddot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))]/2$.

Lemma 7. *If Assumptions 3 and 5-S3 hold, then for any sequence of random variables* $\boldsymbol{u}_n$ *such that* $\boldsymbol{u}_n = o_p(1)$, *it holds that*

$$n\left[\widetilde{L}_n^*(\boldsymbol{\psi}_0^* + \boldsymbol{u}_n) - \widetilde{L}_n^*(\boldsymbol{\psi}_0^*)\right] - n\left[L_n^*(\boldsymbol{\psi}_0^* + \boldsymbol{u}_n) - L_n^*(\boldsymbol{\psi}_0^*)\right] = o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2),$$

where $\widetilde{L}_n^*(\boldsymbol{\psi}) = n^{-1}\sum_{t=1}^n \sum_{k=1}^K \rho_{\tau_k}(y_t - \widetilde{g}_{t,\tau_k}(\boldsymbol{\psi}))$ *and* $L_n^*(\boldsymbol{\psi}) = n^{-1}\sum_{t=1}^n \sum_{k=1}^K \rho_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}))$.

Proof of Lemma 6. Denote $\boldsymbol{u} = \boldsymbol{\psi} - \boldsymbol{\psi}_0^*$, where $\boldsymbol{\psi} = (\boldsymbol{\vartheta}', \boldsymbol{\lambda}')'$ and $\boldsymbol{\psi}_0^* = (\boldsymbol{\vartheta}_0^{*\prime}, \boldsymbol{\lambda}_0^{*\prime})'$. Recall that $L_n^*(\boldsymbol{\psi}) = n^{-1}\sum_{t=1}^n \sum_{k=1}^K \rho_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}))$. Let $X_{t,\tau}^*(\boldsymbol{u}) = \int_0^1 \left[I(\xi_{t,\tau}^* \leqslant \Delta_{t,\tau}^*(\boldsymbol{u})s) - I(\xi_{t,\tau}^* \leqslant 0)\right] ds$ with $\Delta_{t,\tau}^*(\boldsymbol{u}) = g_{t,\tau}(\boldsymbol{\psi}_0^* + \boldsymbol{u}) - g_{t,\tau}(\boldsymbol{\psi}_0^*)$ and $\xi_{t,\tau}^* = y_t - g_{t,\tau}(\boldsymbol{\psi}_0^*)$. By the Knight identity (S5.3), it holds that

$$\begin{aligned} n\left[L_n^*(\boldsymbol{\psi}_0^* + \boldsymbol{u}) - L_n^*(\boldsymbol{\psi}_0^*)\right] &= \sum_{t=1}^n \sum_{k=1}^K \left[\rho_{\tau_k}(\xi_{t,\tau_k}^* - \Delta_{t,\tau_k}^*(\boldsymbol{u})) - \rho_{\tau_k}(\xi_{t,\tau_k}^*)\right] \\ &= K_{1n}^*(\boldsymbol{u}) + K_{2n}^*(\boldsymbol{u}), \end{aligned} \tag{S5.42}$$

where $\boldsymbol{u} \in \aleph \equiv \{\boldsymbol{u} \in \mathbb{R}^d : \boldsymbol{u} + \boldsymbol{\psi}_0^* \in \Psi\}$ with d being the dimension of $\boldsymbol{\psi}$,

$$K_{1n}^*(\boldsymbol{u}) = -\sum_{t=1}^n \sum_{k=1}^K \Delta_{t,\tau_k}^*(\boldsymbol{u})\psi_{\tau_k}(\xi_{t,\tau_k}^*), \quad \text{and} \quad K_{2n}^*(\boldsymbol{u}) = \sum_{t=1}^n \sum_{k=1}^K \Delta_{t,\tau_k}^*(\boldsymbol{u}) X_{t,\tau_k}^*(\boldsymbol{u}).$$

By Taylor expansion, we have $\Delta_{t,\tau}^*(\boldsymbol{u}) = \boldsymbol{u}'\dot{g}_{t,\tau}(\boldsymbol{\psi}_0^*) + \boldsymbol{u}'\ddot{g}_{t,\tau}(\boldsymbol{\psi}^\dagger)\boldsymbol{u}/2$, where $\boldsymbol{\psi}^\dagger$ is between $\boldsymbol{\psi}_0^* + \boldsymbol{u}$ and $\boldsymbol{\psi}_0^*$. Then,

$$\begin{aligned} K_{1n}^*(\boldsymbol{u}) &= -\boldsymbol{u}' \sum_{t=1}^n \sum_{k=1}^K \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(\xi_{t,\tau_k}^*) - \boldsymbol{u}' \sum_{k=1}^K \sum_{t=1}^n \ddot{g}_{t,\tau_k}(\boldsymbol{\psi}^\dagger)\psi_{\tau_k}(\xi_{t,\tau_k}^*)\boldsymbol{u} \\ &= -\sqrt{n}\boldsymbol{u}'\boldsymbol{T}_n^* - \sqrt{n}\boldsymbol{u}' R_{1n}^*(\boldsymbol{\psi}^\dagger)\sqrt{n}\boldsymbol{u}, \end{aligned} \tag{S5.43}$$

where

$$\boldsymbol{T}_n^* = \frac{1}{\sqrt{n}} \sum_{t=1}^n \sum_{k=1}^K \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(\xi_{t,\tau_k}^*) \text{ and } R_{1n}^*(\boldsymbol{\psi}^\dagger) = \frac{1}{n} \sum_{k=1}^K \sum_{t=1}^n \ddot{g}_{t,\tau_k}(\boldsymbol{\psi}^\dagger)\psi_{\tau_k}(\xi_{t,\tau_k}^*).$$

Since $E\left[\sup_{\boldsymbol{\psi}^\dagger \in \Psi} \|\ddot{g}_{t,\tau_k}(\boldsymbol{\psi}^\dagger)\|\right] < \infty$ by Assumption 5 (iii) and the fact that $|\psi_{\tau_k}(\xi_{t,\tau_k}^*)| \leqslant 1$, we have

$$E\left[\sup_{\boldsymbol{\psi}^\dagger \in \Psi} \|\ddot{g}_{t,\tau_k}(\boldsymbol{\psi}^\dagger)\psi_{\tau_k}(\xi_{t,\tau_k}^*)\|\right] < \infty.$$

Moreover, since $\ddot{g}_{t,\tau_k}(\boldsymbol{\psi})$ is continuous with respect to $\boldsymbol{\psi} \in \Psi$, then by ergodic theorem for strictly stationary and α-mixing process under Assumption S1, together with $\boldsymbol{\psi}_n = \boldsymbol{\psi}_0^* + \boldsymbol{u}_n =$

$\boldsymbol{\psi}_0^* + o_p(1)$ and $\boldsymbol{\psi}^\dagger$ between $\boldsymbol{\psi}_0^* + \boldsymbol{u}_n$ and $\boldsymbol{\psi}_0^*$, it holds that

$$R_{1n}^*(\boldsymbol{\psi}^\dagger) = J_1^* + o_p(1),$$

where $J_1^* = \sum_{k=1}^K E\left[\ddot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))\right]/2$. This together with (S5.43), implies that

$$K_{1n}^*(\boldsymbol{u}_n) = -\sqrt{n}\boldsymbol{u}_n'\boldsymbol{T}_n^* - \sqrt{n}\boldsymbol{u}_n' J_1^* \sqrt{n}\boldsymbol{u}_n + o_p(n\|\boldsymbol{u}_n\|^2). \tag{S5.44}$$

Then similar to (S5.22)-(S5.32) in the proof of Lemma 3, we can prove that

$$K_{2n}^*(\boldsymbol{u}_n) = \sqrt{n}\boldsymbol{u}_n' N_1 \sqrt{n}\boldsymbol{u}_n + o_p(\sqrt{n}\|\boldsymbol{u}_n\| + n\|\boldsymbol{u}_n\|^2), \tag{S5.45}$$

where $N_1 = \sum_{k=1}^K f(Q_{\tau_k}(\boldsymbol{\lambda}_0^*))E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^*)\dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\dot{g}_{t,\tau_k}'(\boldsymbol{\psi}_0^*)\right]/2$. In view of (S5.42)-(S5.45), we accomplish the proof of Lemma 6. □

Proof of Corollary S1. Recall that $\widetilde{L}_n^*(\boldsymbol{\psi}) = n^{-1}\sum_{t=1}^n \sum_{k=1}^K \rho_{\tau_k}(y_t - \widetilde{g}_{t,\tau_k}(\boldsymbol{\psi}))$. For $\boldsymbol{u} \in \aleph \equiv \{\boldsymbol{u} \in \mathbb{R}^d : \boldsymbol{u} + \boldsymbol{\psi}_0^* \in \Psi\}$, define $\widehat{H}_n^*(\boldsymbol{u}) = n\left[\widetilde{L}_n^*(\boldsymbol{\psi}_0^* + \boldsymbol{u}) - \widetilde{L}_n^*(\boldsymbol{\psi}_0^*)\right]$. Denote $\widehat{\boldsymbol{u}}_n = \widehat{\boldsymbol{\psi}}_n - \boldsymbol{\psi}_0^*$. By the consistency of $\widehat{\boldsymbol{\psi}}_n$, it holds that $\widehat{\boldsymbol{u}}_n = o_p(1)$. Note that $\widehat{\boldsymbol{u}}_n$ is the minimizer of $\widehat{H}_n^*(\boldsymbol{u})$, since $\widehat{\boldsymbol{\psi}}_n$ is the minimizer of $\widetilde{L}_n^*(\boldsymbol{\psi})$. This together with Lemmas 5-7, implies that

$$\begin{aligned}\widetilde{H}_n^*(\widehat{\boldsymbol{u}}_n) &= -\sqrt{n}\widehat{\boldsymbol{u}}_n'\boldsymbol{T}_n^* + \sqrt{n}\widehat{\boldsymbol{u}}_n'(N_1 - J_1^*)\sqrt{n}\widehat{\boldsymbol{u}}_n + o_p(\sqrt{n}\|\widehat{\boldsymbol{u}}_n\| + n\|\widehat{\boldsymbol{u}}_n\|^2) \qquad \text{(S5.46)}\\ &\geqslant -\sqrt{n}\|\widehat{\boldsymbol{u}}_n\|\left[\|\boldsymbol{T}_n^*\| + o_p(1)\right] + n\|\widehat{\boldsymbol{u}}_n\|^2\left[\lambda_{\min} + o_p(1)\right],\end{aligned}$$

where $\lambda_{\min}$ is the smallest eigenvalue of $N_1 - J_1^* = N^*/2$ with N^* defined before Corollary S1, and $\boldsymbol{T}_n^* = n^{-1/2}\sum_{t=1}^n \sum_{k=1}^K \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))$. Denote $\boldsymbol{Z}_t^* = \sum_{k=1}^K \dot{g}_{t,\tau_k}(\boldsymbol{\psi}_0^*)\psi_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}_0^*))$, then $\boldsymbol{T}_n^* = n^{-1/2}\sum_{t=1}^n \boldsymbol{Z}_t^*$. Since $\boldsymbol{\psi}_0^* = \arg\min_{\boldsymbol{\psi}\in\Psi}\sum_{k=1}^K E[\rho_{\tau_k}(y_t - g_{t,\tau_k}(\boldsymbol{\psi}))]$, we have $E(\boldsymbol{Z}_t^*) = 0$. Moreover, by Lemma 2.1 of White and Domowitz (1984) and Assumption S1, for any nonzero vector $\boldsymbol{c} \in \mathbb{R}^d$, we can show that $\boldsymbol{c}'\boldsymbol{Z}_t^*$ is also a strictly stationary and α-mixing process with the mixing coefficient $\alpha(n)$ satisfying $\sum_{n\geqslant 1}[\alpha(n)]^{1-2/\delta} < \infty$ for some $\delta > 2$. As a result, by central limit theorem for α-mixing process given in Theorem 2.21 of Fan and Yao (2003) and the Cramér-Wold device, $\boldsymbol{T}_n^*$ convergences in distribution to a normal random variable with mean zero and variance matrix $M^* = E(\boldsymbol{Z}_t^* \boldsymbol{Z}_t^{*'}) + n^{-1} \sum_{t \neq s}^n E(\boldsymbol{Z}_t^* \boldsymbol{Z}_s^{*'})$ as $n \to \infty$. Then similar to (S5.40)-(S5.41) in the proof of Theorem 2, we can verify the asymptotic normality of $\widehat{\boldsymbol{\psi}}_n$. □

S6 Proofs of Theorems 1-4 for special cases

In this section, we prove that Theorems 1-4 still hold for both CQR estimators in ARMA-GARCH, ALDAR and ESTAR-GARCH models. It is equivalent to verify that Assumptions 4-6 can be implied by Assumptions 4′-5′ (or Assumptions 4″-5″, or Assumptions 4‴-5‴) for ARMA-GARCH models (or ALDAR models, or ESTAR-GARCH models).

S6.1 Proof of Theorems 1-4 for ARMA-GARCH models

Proof. For the ARMA-GARCH model (2.2), recall that

$$\mu_t(\boldsymbol{\vartheta}^{\mathrm{I}}) = \sum_{i=1}^{p} \alpha_i y_{t-i} + \sum_{j=1}^{q} \beta_j \epsilon_{t-j} \quad \text{and} \quad h_t(\boldsymbol{\vartheta}^{\mathrm{I}}) = \sqrt{\omega + \sum_{i=1}^{Q} \gamma_i \epsilon_{t-i}^2 + \sum_{j=1}^{P} \nu_j h_{t-j}^2},$$

where $\boldsymbol{\vartheta}^{\mathrm{I}} = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \omega, \gamma_1, \ldots, \gamma_Q, \nu_1, \ldots, \nu_P)'$. Define the characteristic polynomials by $\alpha(z) = 1 - \sum_{i=1}^{p} \alpha_i z^i$, $\beta(z) = 1 + \sum_{j=1}^{q} \beta_j z^j$, $\gamma(z) = \sum_{i=1}^{Q} \gamma_i z^i$ and $\nu(z) = 1 - \sum_{j=1}^{P} \nu_j z^j$. Denote $\sum_{i=1}^{\infty} c_i z^i = \gamma(z)/\nu(z)$ and $1 + \sum_{i=1}^{\infty} d_i z^i = \alpha(z)/\beta(z)$. Then $\mu_t(\boldsymbol{\vartheta}^{\mathrm{I}})$ and $h_t(\boldsymbol{\vartheta}^{\mathrm{I}})$ of model (2.2) have the autoregressive representions:

$$\mu_t(\boldsymbol{\vartheta}^{\mathrm{I}}) = -\sum_{i=1}^{\infty} d_i y_{t-i} \quad \text{and} \quad h_t(\boldsymbol{\vartheta}^{\mathrm{I}}) = \sqrt{\omega/\nu(1) + \sum_{i=1}^{\infty} c_i \epsilon_{t-i}^2(\boldsymbol{\vartheta}^{\mathrm{I}})},$$

where $\epsilon_t(\boldsymbol{\vartheta}^{\mathrm{I}}) = y_t + \sum_{i=1}^{\infty} d_i y_{t-i}$.

We first verify that $q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})$ and $\widetilde{q}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})$ for model (2.2) with Assumptions 4′-5′ imply Assumptions 4-6. Note that $\omega = 1$ by Assumption 4′ (iii) for the semi-parametric CQR in model (2.2), then we rewrite the model parameter by $\boldsymbol{\vartheta}_*^{\mathrm{I}} = (\alpha_1, \ldots, \alpha_p, \beta_1, \ldots, \beta_q, \gamma_1, \ldots, \gamma_Q, \nu_1, \ldots, \nu_P)'$. Thus it holds that $q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}}) = \mu_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) + b_k h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})$, where $\boldsymbol{\phi}^{\mathrm{I}} = (\boldsymbol{\vartheta}_*^{\mathrm{I}\prime}, b_1, \ldots, b_K)'$. Obviously, $q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})$ is continuous in $\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}$ and Assumption 4 (i) holds, where $\Phi^{\mathrm{I}} = \mathbb{R}^{p+q} \times \mathbb{R}^{+P+Q} \times \mathbb{R}^K$ is the parameter space. Denote the true parameter by $\boldsymbol{\phi}_0^{\mathrm{I}} = (\boldsymbol{\vartheta}_{*0}^{\mathrm{I}\prime}, b_{10}, \ldots, b_{K0})'$, where $\boldsymbol{\vartheta}_{*0}^{\mathrm{I}} = (\alpha_{10}, \ldots, \alpha_{p0},\ \beta_{10}, \ldots, \beta_{q0},\ \gamma_{10}, \ldots, \gamma_{Q0}, \nu_{10}, \ldots, \nu_{P0})'$. Then we can get $\alpha_0(z) = 1 - \sum_{i=1}^p \alpha_{i0} z^i$, $\beta_0(z) = 1 + \sum_{j=1}^q \beta_{j0} z^j$, $\gamma_0(z) = \sum_{i=1}^Q \gamma_{i0} z^i$ and $\nu_0(z) = 1 - \sum_{j=1}^P \nu_{j0} z^j$. For Assumption 4 (ii), we can write $q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}}) = y_t - \epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) + b_k h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})$. Then if $q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}}) = q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}_0^{\mathrm{I}})$, it holds that

$$\epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \epsilon_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) - b_k h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) + b_{k0} h_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) = 0. \tag{S6.1}$$

Denote $\epsilon_t = \epsilon_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}})$, then by Assumption 1, we have

$$\epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \epsilon_t = \sum_{i=1}^{\infty} a_i \epsilon_{t-i},$$

where $1 + \sum_{i=1}^{\infty} a_i z^i = \alpha(z)\beta_0(z)/[\alpha_0(z)\beta(z)]$. Therefore, (S6.1) can be written as

$$a_1 \epsilon_{t-1} + H_{1,t-2} - b_k \sqrt{c_1(\epsilon_{t-1} + H_{2,t-2})^2 + H_{3,t-2}} + b_{k0} \sqrt{c_{10}\epsilon_{t-1}^2 + H_{4,t-2}} = 0, \tag{S6.2}$$

where $\sum_{i=1}^{\infty} c_{i0} z^i = \gamma_0(z)/\nu_0(z)$,

$$\begin{aligned} H_{1,t-j} &= \sum_{i=j}^{\infty} a_i \epsilon_{t-i}, \quad H_{2,t-j} = \sum_{i=1}^{\infty} a_i \epsilon_{t-j-i}, \\ H_{3,t-j} &= 1/\nu(1) + \sum_{i=j}^{\infty} c_i \epsilon_{t-i}^2(\boldsymbol{\vartheta}_*^{\mathrm{I}}), \quad \text{and} \quad H_{4,t-j} = 1/\nu_0(1) + \sum_{i=j}^{\infty} c_{i0} \epsilon_{t-i}^2. \end{aligned} \tag{S6.3}$$

Since ϵ_{t-1} is independent of all the others given $\mathcal{F}_{t-2}$, it holds that

$$a_i - b_k c_i^{1/2} + b_{k0} c_{i0}^{1/2} = 0. \tag{S6.4}$$

Let $\varsigma(x) = ax + H_1 - b_k[c_1(x + H_2)^2 + H_3]^{1/2} + b_{k0}(c_{10}x^2 + H_4)^{1/2}$, then we can get from (S6.2) that $\varsigma(x) = 0$ for all $x \in \mathbb{R}$. Since $H_{4,t-j} > 0$ for any $t - j \in \mathbb{Z}$ and $d^3\varsigma(x)/dx^3 = 0$

when $x = -H_2$, it holds that $H_2 = 0$, which together with (S6.2) yields $H_{2,t-2} = 0$. Hence, $\epsilon_{t-1}(\boldsymbol{\vartheta}_*^{\mathrm{I}}) = \epsilon_{t-1}$, then together with Assumption 4′ (i), we obtain $\alpha_i = \alpha_{i0}$ for all $1 \leqslant i \leqslant p$ and $\beta_j = \beta_{j0}$ for all $1 \leqslant j \leqslant q$. Then combining with (S6.3) and (S6.4), it can be further verified that $H_{1,t-j} = 0$ for any $t-j \in \mathbb{Z}$ and $b_k c_i^{1/2} = b_{k0} c_{i0}^{1/2}$ for all $i \geqslant 1$. Therefore, it holds that

$$b_k^2 \frac{\gamma(z)}{\nu(z)} = b_{k0}^2 \frac{\gamma_0(z)}{\nu_0(z)}, \ |z| \leqslant 1. \tag{S6.5}$$

This together with Assumption 4′ (ii) implies that $b_k = b_{k0}$ for $1 \leqslant k \leqslant K$, and then $\gamma_i = \gamma_{i0}$ for all $1 \leqslant i \leqslant Q$ and $\nu_j = \nu_{j0}$ for $1 \leqslant j \leqslant P$. Therefore, $\boldsymbol{\phi}^{\mathrm{I}} = \boldsymbol{\phi}_0^{\mathrm{I}}$ and Assumption 4 (ii) holds.

Denote $\dot{c}_i = \partial c_i / \partial \boldsymbol{\phi}^{\mathrm{I}}$, $\ddot{c}_i = \partial^2 c_i / (\partial \boldsymbol{\phi}^{\mathrm{I}} \partial \boldsymbol{\phi}^{\mathrm{I}'})$, and the same as $\dot{d}_i$ and $\ddot{d}_i$ for each $i \geqslant 1$. Then under Assumptions 1, 2 and 5′ (i), it holds that

$$\begin{aligned} &(i) \sup_{\boldsymbol{\phi}^{\mathrm{I}}} |c_i|, \ \sup_{\boldsymbol{\phi}^{\mathrm{I}}} \|\dot{c}_i\|, \ \sup_{\boldsymbol{\phi}^{\mathrm{I}}} \|\ddot{c}_i\| \ \sup_{\boldsymbol{\phi}^{\mathrm{I}}} |d_i|, \ \sup_{\boldsymbol{\phi}^{\mathrm{I}}} \|\dot{d}_i\|, \ \sup_{\boldsymbol{\phi}^{\mathrm{I}}} \|\ddot{d}_i\| \leqslant C\rho^i; \\ &(ii) c_i \geqslant C\rho^i, \end{aligned} \tag{S6.6}$$

for some constants $C > 0$ and $0 < \rho < 1$; see Francq and Zakoïan (2004). Since $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5′ (i), $E(|y_t|) < \infty$ by Assumption 5′ (ii) and (S6.6), it follows that

$$\begin{aligned} E \sup_{\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}} |q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})| &\leqslant E \sup_{\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}} \left[\sum_{i=1}^{\infty} |d_i||y_{t-i}| + |\overline{b}| + |\overline{b}| \sum_{i=1}^{\infty} \sqrt{|c_i|} |\epsilon_{t-i}| \right] \\ &\leqslant CE(|y_t|) \sum_{i=1}^{\infty} (\rho^i + \rho^{i/2}) < \infty. \end{aligned} \tag{S6.7}$$

Since

$$\left\| \frac{\partial h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}} \right\| \leqslant C + \sum_{i=1}^{\infty} \|\dot{c}_i / \sqrt{c_i}\| |\epsilon_{t-i}| + \sum_{i=1}^{\infty} \sqrt{c_i} \left(\sum_{l=1}^{\infty} \|\dot{d}_l\| |y_{t-i-l}| \right) \tag{S6.8}$$

and

$$\left\|\frac{\partial^2 h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}} \partial \boldsymbol{\vartheta}_*^{\mathrm{I}'}}\right\| \leqslant C + C\sum_{i=1}^{\infty}\left(\|\dot{c}_i/\sqrt{c_i}\| + \sqrt{c_i}\right)\left(\sum_{l=1}^{\infty}\|\dot{d}_l\||y_{t-i-l}|\right) + C\sum_{i=1}^{\infty}\left(\|\ddot{c}_i/\sqrt{c_i}\| + \|\dot{c}_i/\sqrt{c_i}\|\right)|\epsilon_{t-i}| + \sum_{i=1}^{\infty}\sqrt{c_i}\left(\sum_{l=1}^{\infty}\|\ddot{d}_l\||y_{t-i-l}|\right), \quad \text{(S6.9)}$$

together with $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5′ (i), $E(y_t^2) < \infty$ by Assumption 5′ (ii), and (S6.6), we have

$$\begin{aligned}
& E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}} \|\dot{q}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})\|^3\right] \\
\leqslant & E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}} \left(\left\|\frac{\partial \mu_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}}\right\|^3 + \overline{b}^3\left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}}\right\|^3 + |h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})|^3\right)\right] \\
\leqslant & CE \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}}\left[\sum_{i=1}^{\infty}\|\dot{d}_i\|^2 y_{t-i}^2 + \sum_{i=1}^{\infty}(\|\dot{c}_i/\sqrt{c_i}\|^2 + c_i)\epsilon_{t-i}^2 + \sum_{i=1}^{\infty}c_i\left(\sum_{l=1}^{\infty}\|\dot{d}_l\|^2 y_{t-i-l}^2\right)\right] \\
\leqslant & CE(y_t^2)\sum_{i=1}^{\infty}(\rho^{2i} + \rho^i) < \infty
\end{aligned}$$

and

$$\begin{aligned}
& E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}} \|\ddot{q}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})\|^2\right] \\
\leqslant & E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}} \left(\left\|\frac{\partial^2 \mu_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}\partial \boldsymbol{\vartheta}_*^{\mathrm{I}'}}\right\|^2 + \overline{b}^2\left\|\frac{\partial^2 h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}\partial \boldsymbol{\vartheta}_*^{\mathrm{I}'}}\right\|^2 + 4\left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}})\right\|^2\right)\right] \\
\leqslant & CE \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}}\left[\sum_{i=1}^{\infty}\|\ddot{d}_i\||y_{t-i}| + \sum_{i=1}^{\infty}\left(\|\dot{c}_i/\sqrt{c_i}\| + \sqrt{c_i}\right)\left(\sum_{l=1}^{\infty}\|\dot{d}_l\||y_{t-i-l}|\right)\right] \\
& + CE \sup_{\boldsymbol{\phi}^{\mathrm{I}}\in\Phi^{\mathrm{I}}}\left[\sum_{i=1}^{\infty}\left(\|\ddot{c}_i/\sqrt{c_i}\| + \|\dot{c}_i/\sqrt{c_i}\|\right)|\epsilon_{t-i}| + \sum_{i=1}^{\infty}\sqrt{c_i}\left(\sum_{l=1}^{\infty}\|\ddot{d}_l\||y_{t-i-l}|\right)\right] \\
\leqslant & CE(|y_t|)\sum_{i=1}^{\infty}(\rho^i + \rho^{i/2}) < \infty,
\end{aligned}$$

then Assumption 5 holds.

For Assumption 6, since $\mu_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \widetilde{\mu}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) = -\sum_{i=t}^{\infty} d_i y_{t-i}$, $\dot{\mu}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \dot{\widetilde{\mu}}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) = -\sum_{i=t}^{\infty} \dot{d}_i y_{t-i}$,

$$h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) = \frac{h_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \widetilde{h}_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) + \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})} \leqslant \sum_{i=t}^{\infty}\sqrt{c_i}|\epsilon_{t-i}| + \sum_{i=1}^{t-1}\sqrt{c_i}|\epsilon_{t-i} - \widetilde{\epsilon}_{t-i}|, \quad \text{(S6.10)}$$

and

$$
\begin{aligned}
\dot{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \dot{\widetilde{h}}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) =& \frac{1}{2h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})} \frac{\partial h_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}} - \frac{1}{2\widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})} \frac{\partial \widetilde{h}_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}}} \\
=& \frac{1}{2h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})} \left[\frac{\partial h_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}} - \frac{\partial \widetilde{h}_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}} \right] + \left[\frac{1}{2h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})} - \frac{1}{2\widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})} \right] \frac{\partial \widetilde{h}_t^2(\boldsymbol{\vartheta}_*^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{I}}} \\
\leqslant& C + \sum_{i=1}^{\infty} \dot{c}_i |\epsilon_{t-i}| / \sqrt{c_i} + \sum_{i=t}^{\infty} \sqrt{c_i} \left(\sum_{l=1}^{\infty} \dot{d}_l y_{t-i-l} \right) \\
& + \sum_{i=1}^{t-1} \sqrt{c_i} \left[\left(\sum_{l=t-i}^{\infty} \dot{d}_l y_{t-i-l} \right) + \frac{1}{2} \left(\sum_{l=1}^{t-i-1} \dot{d}_l y_{t-i-l} \right) \right],
\end{aligned}
\tag{S6.11}
$$

together with $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption $5'$ (i), $E(|y_t|) < \infty$ by Assumption $5'$ (ii) and (S6.6), it holds that

$$
\begin{aligned}
& \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}} |q_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}}) - \widetilde{q}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})|^2 \\
\leqslant & 2 \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}} \left(|\mu_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \widetilde{\mu}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})|^2 + \overline{b}^2 |h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})|^2 \right) \\
\leqslant & C \sum_{t=1}^{\infty} \rho^t \left[\sum_{i=t}^{\infty} \rho^i |y_{t-i}| + \sum_{i=t}^{\infty} \rho^{i/2} |\epsilon_{t-i}| + \sum_{i=t}^{\infty} \rho^{i/2} \left(\sum_{l=t-i}^{\infty} \rho^l |y_{t-i-l}| \right) \right] \\
\leqslant & C \sum_{t=1}^{\infty} \rho^t \varsigma_\rho < \infty,
\end{aligned}
$$

and

$$
\begin{aligned}
& \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}} \|\dot{q}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}}) - \dot{\widetilde{q}}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})\|^2 \\
\leqslant & C \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{I}}) \sup_{\boldsymbol{\phi}^{\mathrm{I}} \in \Phi^{\mathrm{I}}} \left(\|\dot{\mu}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \dot{\widetilde{\mu}}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})\|^2 + \overline{b}^2 \|\dot{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \dot{\widetilde{h}}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})\|^2 + |h_t(\boldsymbol{\vartheta}_*^{\mathrm{I}}) - \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{I}})|^2 \right) \\
\leqslant & C \sum_{t=1}^{\infty} \rho^t \left[\sum_{i=t}^{\infty} \rho^i |y_{t-i}| + \sum_{i=1}^{\infty} \rho^{i/2} |\epsilon_{t-i}| \right] + C \sum_{i=t}^{\infty} \rho^{i/2} \left(\sum_{l=1}^{\infty} \rho^l |y_{t-i-l}| \right) + C \sum_{i=1}^{t-1} \rho^{i/2} \left(\sum_{l=t-i}^{\infty} \rho^l |y_{t-i-l}| \right) \\
\leqslant & C \sum_{t=1}^{\infty} \rho^t \varsigma_\rho < \infty,
\end{aligned}
$$

where $0 < \rho < 1$ and $\varsigma_\rho = \sum_{t=0}^{\infty} \rho^t |y_{-t}|$. Then Assumption 6 holds.

Then we prove that $g_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}})$ and $\widetilde{g}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}})$ of the parametric CQR for model (2.2) with

Assumptions 4′-5′ imply Assumptions 4-6. For model (2.2), $g^{\mathrm{I}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{I}})$ has the form of

$$g^{\mathrm{I}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{I}}) = \mu_t(\boldsymbol{\vartheta}^{\mathrm{I}}) + Q_{\tau_k}(\lambda)h_t(\boldsymbol{\vartheta}^{\mathrm{I}}),$$

where $\boldsymbol{\psi}^{\mathrm{I}} = (\boldsymbol{\vartheta}^{\mathrm{I}'}, \lambda)'$. Obviously, $g^{\mathrm{I}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{I}})$ is continuous in $\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}$ and Assumption 4 (i) holds, where $\Psi^{\mathrm{I}} = \mathbb{R}^{p+q} \times \mathbb{R}^{+^{1+P+Q}} \times \mathbb{R}$. For Assumption 4 (ii), similar to (S6.1)-(S6.5), it holds that $\boldsymbol{\vartheta}^{\mathrm{I}}_{*} = \boldsymbol{\vartheta}^{\mathrm{I}}_{*0}$ and

$$\sqrt{\omega}Q_{\tau_k}(\lambda) = \sqrt{\omega_0}Q_{\tau_k}(\lambda_0)$$

if $g^{\mathrm{I}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{I}}) = g^{\mathrm{I}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{I}}_0)$. Then we prove that $\omega = \omega_0$ and $\lambda = \lambda_0$ under Assumption 4′ (iv). Consider four arbitrary quantile levels $0 < \tau_1 < \tau_2 < \tau_3 < \tau_4 < 1$, and two arbitrary shape parameter $\lambda, \widetilde{\lambda} < 1$ such that

$$Q_{\tau_j}(\lambda) = cQ_{\tau_j}(\widetilde{\lambda}) \text{ for all } 1 \leqslant j \leqslant 4, \tag{S6.12}$$

where $c > 0$. We show (S6.12) holds if and only if $\lambda = \widetilde{\lambda}$ and $c = 1$ in the following. Define $G(\tau) = Q_\tau(\lambda) - cQ_\tau(\widetilde{\lambda})$, it follows that $G(\tau_1) = G(\tau_2) = G(\tau_3) = G(\tau_4) = 0$ and thus $G(\tau) = 0$ has at least four different solutions. The first derivative of $G(\tau)$ is $\dot{G}(\tau) = \tau^{\lambda-1} + (1-\tau)^{\lambda-1} - c[\tau^{\widetilde{\lambda}-1} + (1-\tau)^{\widetilde{\lambda}-1}]$. Then $\dot{G}(\tau) = 0$ if and only if $S(\tau) = c$, where $S(\tau) = \frac{\tau^{\lambda-1}+(1-\tau)^{\lambda-1}}{\tau^{\widetilde{\lambda}-1}+(1-\tau)^{\widetilde{\lambda}-1}}$. Then $S(\tau) = c$ has at least three different solutions. It can be simply verified that: (i) when $\lambda < \widetilde{\lambda} < 1$, $S(\tau)$ is strictly increasing for $\tau > 0.5$ and strictly decreasing for $\tau < 0.5$, which implies that $S(\tau) \geqslant S(0.5) = 1$; (ii) when $\widetilde{\lambda} < \lambda < 1$, $S(\tau)$ is strictly decreasing for $\tau > 0.5$ and strictly increasing for $\tau < 0.5$, which implies that $S(\tau) \leqslant S(0.5) = 1$; (iii) when $\lambda = \widetilde{\lambda} < 1$, $S(\tau) = 1$ for all τ. Then it holds that: (a) when $c \neq 1$, the equation $S(\tau) = c$ has at most two different solutions; (b) when $c = 1$ and $\lambda \neq \widetilde{\lambda}$, the equation $S(\tau) = 1$ has at most two different solutions; (c) when $c = 1$ and $\lambda = \widetilde{\lambda}$, $S(\tau) = 1$ holds for all τ. Since $G(\tau) = 0$ has at least four different solutions, we prove that

Case (c) holds, then $\lambda = \lambda_0$ and $\omega = \omega_0$. Therefore, $\boldsymbol{\psi}^{\mathrm{I}} = \boldsymbol{\psi}_0^{\mathrm{I}}$ and Assumption 4 (ii) holds.

For Assumption 5, similar to (S6.8)-(S6.9), we have

$$\left\| \frac{\partial h_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}}} \right\| \leqslant C + \sum_{i=1}^{\infty} \|\dot{c}_i/\sqrt{c_i}\| |\epsilon_{t-i}| + \sum_{i=1}^{\infty} \sqrt{c_i} \left(\sum_{l=1}^{\infty} \|\dot{d}_l\| |y_{t-i-l}| \right) \quad \text{and}$$

$$\left\| \frac{\partial^2 h_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}} \partial \boldsymbol{\vartheta}^{\mathrm{I}\prime}} \right\| \leqslant C + C \sum_{i=1}^{\infty} \left(\|\dot{c}_i/\sqrt{c_i}\| + \sqrt{c_i} \right) \left(\sum_{l=1}^{\infty} \|\dot{d}_l\| |y_{t-i-l}| \right)$$
$$+ C \sum_{i=1}^{\infty} \left(\|\ddot{c}_i/\sqrt{c_i}\| + \|\dot{c}_i/\sqrt{c_i}\| \right) |\epsilon_{t-i}| + \sum_{i=1}^{\infty} \sqrt{c_i} \left(\sum_{l=1}^{\infty} \|\ddot{d}_l\| |y_{t-i-l}| \right).$$

Since $\underline{\lambda} \leqslant \lambda \leqslant \overline{\lambda}$ by Assumption 5′ (i), there exist positive constants $\overline{Q}_1$ and $\overline{Q}_2$ such that

$$|Q_\tau(\lambda)| \leqslant \overline{Q}_1 \quad \text{and} \quad |\dot{Q}_\tau(\lambda)| \leqslant \overline{Q}_2, \tag{S6.13}$$

where $\dot{Q}_\tau(\lambda)$ denotes the first derivative of $Q_\tau(\lambda)$. This together with $E(y_t^2) < \infty$ by Assumption 5′ (ii), and (S6.6) yields that

$$E \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} |g_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}})| \leqslant E \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \left[\sum_{i=1}^{\infty} |d_i||y_{t-i}| + \overline{Q}_1 \sqrt{\overline{\omega}} + \overline{Q}_1 \sum_{i=1}^{\infty} \sqrt{|c_i|} |\epsilon_{t-i}| \right]$$
$$\leqslant C E(|y_t|) \sum_{i=1}^{\infty} (\rho^i + \rho^{i/2}) < \infty,$$

$$E \left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \|\dot{g}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}})\|^3 \right]$$
$$\leqslant E \left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \left(\left\| \frac{\partial \mu_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}}} \right\|^3 + \overline{Q}_1^3 \left\| \frac{\partial h_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}}} \right\|^3 + |h_t(\boldsymbol{\vartheta}^{\mathrm{I}})|^3 \right) \right]$$
$$\leqslant C E(y_t^2) \sum_{i=1}^{\infty} (\rho^{2i} + \rho^i) < \infty,$$

and

$$E \left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \|\ddot{g}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\phi}^{\mathrm{I}})\|^2 \right]$$
$$\leqslant E \left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \left(\left\| \frac{\partial^2 \mu_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}} \partial \boldsymbol{\vartheta}^{\mathrm{I}\prime}} \right\|^2 + \overline{Q}_1^2 \left\| \frac{\partial^2 h_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}} \partial \boldsymbol{\vartheta}^{\mathrm{I}\prime}} \right\|^2 + \overline{Q}_2^2 \left\| \frac{\partial h_t(\boldsymbol{\vartheta}^{\mathrm{I}})}{\partial \boldsymbol{\vartheta}^{\mathrm{I}}} \right\|^2 \right) \right]$$
$$\leqslant C E(|y_t|) \sum_{i=1}^{\infty} (\rho^i + \rho^{i/2}) < \infty,$$

then Assumption 5 holds.

For Assumption 6, similar to (S6.10)-(S6.11), we have $\mu_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \widetilde{\mu}_t^{\mathrm{I}}(\boldsymbol{\vartheta}^{\mathrm{I}}) = -\sum_{i=t}^{\infty} d_i y_{t-i}$, $\dot{\mu}_t^{\mathrm{I}}(\boldsymbol{\vartheta}^{\mathrm{I}}) - \dot{\widetilde{\mu}}_t^{\mathrm{I}}(\boldsymbol{\vartheta}^{\mathrm{I}}) = -\sum_{i=t}^{\infty} \dot{d}_i y_{t-i}$,

$$
\begin{aligned}
h_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{I}}) \leqslant & \sum_{i=t}^{\infty} \sqrt{c}_i |\epsilon_{t-i}| + \sum_{i=1}^{t-1} \sqrt{c}_i |\epsilon_{t-i} - \widetilde{\epsilon}_{t-i}|, \quad \text{and} \\
\dot{h}_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \dot{\widetilde{h}}_t(\boldsymbol{\vartheta}^{\mathrm{I}}) \leqslant & C + \sum_{i=1}^{\infty} \dot{c}_i |\epsilon_{t-i}| / \sqrt{c_i} + \sum_{i=t}^{\infty} \sqrt{c_i} \left(\sum_{l=1}^{\infty} \dot{d}_l y_{t-i-l} \right) \\
& + \sum_{i=1}^{t-1} \sqrt{c_i} \left[\left(\sum_{l=t-i}^{\infty} \dot{d}_l y_{t-i-l} \right) + \frac{1}{2} \left(\sum_{l=1}^{t-i-1} \dot{d}_l y_{t-i-l} \right) \right].
\end{aligned}
$$

These together with $E(y_t^2) < \infty$ by Assumption $5'$ (ii) and (S6.13) imply that

$$
\begin{aligned}
& \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} |g_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}}) - \widetilde{g}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}})|^2 \\
\leqslant & 2 \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \left(|\mu_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \widetilde{\mu}_t(\boldsymbol{\vartheta}^{\mathrm{I}})|^2 + \overline{Q}_1^2 |h_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{I}})|^2 \right) \\
\leqslant & C \sum_{t=1}^{\infty} \rho^t \varsigma_\rho < \infty,
\end{aligned}
$$

and

$$
\begin{aligned}
& \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \|\dot{g}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}}) - \dot{\widetilde{g}}_{t,\tau_k}^{\mathrm{I}}(\boldsymbol{\psi}^{\mathrm{I}})\|^2 \\
\leqslant & C \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{I}}) \sup_{\boldsymbol{\psi}^{\mathrm{I}} \in \Psi^{\mathrm{I}}} \left(\|\dot{\mu}_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \dot{\widetilde{\mu}}_t(\boldsymbol{\vartheta}^{\mathrm{I}})\|^2 + \overline{Q}_1^2 \|\dot{h}_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \dot{\widetilde{h}}_t(\boldsymbol{\vartheta}^{\mathrm{I}})\|^2 + \overline{Q}_2^2 |h_t(\boldsymbol{\vartheta}^{\mathrm{I}}) - \widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{I}})|^2 \right) \\
\leqslant & C \sum_{t=1}^{\infty} \rho^t \varsigma_\rho < \infty,
\end{aligned}
$$

where $0 < \rho < 1$ and $\varsigma_\rho = \sum_{t=0}^{\infty} \rho^t |y_{-t}|$. Then Assumption 6 holds and the proof is accomplished.

□

S6.2 Proof of Theorems 1-4 in ALDAR models

Similar to the previous subsection S6.1, we only verify that $q^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}})$ and $g^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{II}})$ for ALDAR model (2.3) with Assumptions 4″-5″ imply Assumptions 4-5.

Proof. For ALDAR model (2.3), recall that

$$\mu_t(\boldsymbol{\vartheta}^{\mathrm{II}}) = \sum_{i=1}^{p} \varphi_i y_{t-i} \quad \text{and} \quad h_t(\boldsymbol{\vartheta}^{\mathrm{II}}) = \omega + \sum_{j=1}^{q} (\alpha_j^+ y_{t-j}^+ - \alpha_j^- y_{t-j}^-),$$

where $\boldsymbol{\vartheta}^{\mathrm{II}} = (\varphi_1, \varphi_2, \ldots, \varphi_p, \omega, \alpha_1^+, \alpha_2^+, \ldots, \alpha_q^+, \alpha_1^-, \alpha_2^-, \ldots, \alpha_q^-)'$.

We first verify that $q^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}})$ with Assumptions 4″-5″ satisfies Assumptions 4-5. Note that $\omega = 1$ by Assumption 4″ (i) for semi-parametric CQR in ALDAR models, then we denote the model parameter by $\boldsymbol{\vartheta}^{\mathrm{II}}_* = (\varphi_1, \varphi_2, \ldots, \varphi_p, \alpha_1^+, \alpha_2^+, \ldots, \alpha_q^+, \alpha_1^-, \alpha_2^-, \ldots, \alpha_q^-)'$. Thus it holds that $q^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}}) = \mu_t(\boldsymbol{\vartheta}^{\mathrm{II}}_*) + b_k h_t(\boldsymbol{\vartheta}^{\mathrm{II}}_*)$, where $\boldsymbol{\phi}^{\mathrm{II}} = (\boldsymbol{\vartheta}^{\mathrm{II}\prime}_*, b_1, \ldots, b_K)'$. Obviously, $q^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}})$ is continuous in $\boldsymbol{\phi}^{\mathrm{II}} \in \Phi^{\mathrm{II}}$ and then Assumption 4 (i) holds, where $\Phi^{\mathrm{II}} = \mathbb{R}^p \times \mathbb{R}^{+2q} \times \mathbb{R}^K$ is the parameter space. Denote the true parameter $\boldsymbol{\phi}^{\mathrm{II}}_0 = (\boldsymbol{\vartheta}^{\mathrm{II}\prime}_{*0}, b_{10}, \ldots, b_{K0})$ with $\boldsymbol{\vartheta}^{\mathrm{II}}_{*0} = (\varphi_{10}, \ldots, \varphi_{p0}, \alpha_{10}^+, \ldots, \alpha_{q0}^+, \alpha_{10}^-, \ldots, \alpha_{q0}^-)$. For Assumption 4 (ii), if $q^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}}) = q_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}}_0)$, it holds that

$$\begin{aligned} &(\varphi_1 - \varphi_{10}) y_{t-1} + (b_k \alpha_1^+ - b_{k0}\alpha_{10}^+) y_{t-1}^+ - (b_k \alpha_1^- - b_{k0}\alpha_{10}^-) y_{t-1}^- + (b_k - b_{k0}) \\ =& \sum_{i=2}^{p} (\varphi_{i0} - \varphi_i) y_{t-i} + \sum_{i=2}^{q} \left[(b_{k0}\alpha_{i0}^+ - b_k \alpha_i^+) y_{t-i}^+ - (b_{k0}\alpha_{i0}^- - b_k \alpha_i^-) y_{t-1}^- \right]. \end{aligned} \tag{S6.14}$$

Since $(y_{t-1}, y_{t-1}^+, y_{t-1}^-)$ are independent of all the others given $\mathcal{F}_{t-2}$, we have $b_k = b_{k0}$ and then it holds that $\alpha_i^+ = \alpha_{i0}^+$ and $\alpha_i^- = \alpha_{i0}^-$ under Assumption 4″ (i), thus $\varphi_i = \varphi_{i0}$ follows. Therefore, $\boldsymbol{\phi}^{\mathrm{II}} = \boldsymbol{\phi}^{\mathrm{II}}_0$ and Assumption 4 (ii) holds.

Then we consider Assumption 5. In model (2.3), it holds that

$$\dot{q}^{\mathrm{II}}_{t,\tau_k}(\boldsymbol{\phi}^{\mathrm{II}}) = (\boldsymbol{Y}'_{t-1}, b_k \boldsymbol{X}'_{t-1}, h_t(\boldsymbol{\vartheta}^{\mathrm{II}}_*) \boldsymbol{e}'_k)',$$

where $\boldsymbol{Y}_{t-1} = (y_{t-1}, y_{t-2}, \ldots, y_{t-p})'$, $\boldsymbol{X}_{t-1} = (y_{t-1}^+, \ldots, y_{t-q}^+, -y_{t-1}^-, \ldots, -y_{t-q}^-)'$ and $\boldsymbol{e}_k$ is a K-dimensional vector with the kth element being 1 and others being 0. Since $\underline{\varphi} \leqslant \varphi_i \leqslant \overline{\varphi}$ for $1 \leqslant i \leqslant p$, $0 \leqslant \underline{\alpha} \leqslant \alpha_i^+, \alpha_i^- \leqslant \overline{\alpha}$ for $1 \leqslant i \leqslant q$ and $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption $5''$ (i) and $E(|y_t|^3) < \infty$ by Assumption $5''$ (ii), then by Taylor expansion, it holds that

$$E \sup_{\boldsymbol{\phi}^{\mathrm{II}} \in \Phi^{\mathrm{II}}} |q_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\phi}^{\mathrm{II}})| \leqslant E\left[|\overline{\varphi}| \sum_{i=1}^{p} |y_{t-i}| + \overline{b}\left(1 + \overline{\alpha} \sum_{i=1}^{q} |y_{t-i}|\right)\right] < \infty,$$

$$\begin{aligned} E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{II}}) \sup_{\boldsymbol{\phi}^{\mathrm{II}} \in \Phi^{\mathrm{II}}} \|\dot{q}_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\phi}^{\mathrm{II}})\|^3\right] &\leqslant CE\left[\sup_{\boldsymbol{\phi} \in \Phi^{\mathrm{II}}} \|\dot{q}_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\phi}^{\mathrm{II}})\|^3\right] \\ &\leqslant CE\left[\|\boldsymbol{Y}_{t-1}\|^3 + \overline{b}^3 \|\boldsymbol{X}_{t-1}\|^3 + \left(1 + \overline{\alpha} \sum_{i=1}^{q} |y_{t-i}|\right)^3\right] \\ &\leqslant CE(|y_t|^3) < \infty, \end{aligned}$$

and

$$E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{II}}) \sup_{\boldsymbol{\phi}^{\mathrm{II}} \in \Phi^{\mathrm{II}}} \|\ddot{q}_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\phi}^{\mathrm{II}})\|^2\right] \leqslant 2E\left[\|\boldsymbol{X}_{t-1}\|^2\right] \leqslant CE(y_t^2) < \infty.$$

Then Assumption 5 holds.

We next prove that $g_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}})$ with Assumptions $4''$-$5''$ for model (2.3) implies Assumptions 4-5. For model (2.3), $g_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}})$ has the form $g_{t,\tau}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}}) = \mu_t(\boldsymbol{\vartheta}^{\mathrm{II}}) + Q_\tau(\lambda) h_t(\boldsymbol{\vartheta}^{\mathrm{II}})$, where $\boldsymbol{\psi}^{\mathrm{II}} = (\boldsymbol{\vartheta}^{\mathrm{II}\prime}, \lambda)'$. Clearly, $g_{t,\tau}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}})$ is continuous in $\boldsymbol{\psi}^{\mathrm{II}} \in \Psi^{\mathrm{II}}$ and then Assumption 4 (i) holds, where $\Psi^{\mathrm{II}} = \mathbb{R}^p \times \mathbb{R}^{+^{2q+1}} \times \mathbb{R}$ is the parameter space. For Assumption 4 (ii), similar to (S6.14) and the proof of (S6.12), we can show that $\boldsymbol{\psi}^{\mathrm{II}} = \boldsymbol{\psi}_0^{\mathrm{II}}$ under the Assumption $4''$ (ii) if $g_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}}) = g_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}_0^{\mathrm{II}})$. Then Assumption 4 (ii) holds.

For Assumption 5, since

$$\dot{g}_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}}) = (\boldsymbol{Y}_{t-1}', Q_{\tau_k}(\lambda), Q_{\tau_k}(\lambda)\boldsymbol{X}_{t-1}', \dot{Q}_{\tau_k}(\lambda) h_t(\boldsymbol{\vartheta}^{\mathrm{II}}))',$$

$\underline{\varphi} \leqslant \varphi_i \leqslant \overline{\varphi}$ for $1 \leqslant i \leqslant p$, $\underline{\omega} \leqslant \omega \leqslant \overline{\omega}$ and $0 \leqslant \underline{\alpha} \leqslant \alpha_i^+, \alpha_i^- \leqslant \overline{\alpha}$ for $1 \leqslant i \leqslant q$ by Assumption

5″ (i), $E(|y_t|^3) < \infty$ by Assumption 5″ (ii), together with (S6.13), it holds that

$$E \sup_{\boldsymbol{\psi}^{\mathrm{II}} \in \Psi^{\mathrm{II}}} |g_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}})| \leqslant E\left[|\overline{\varphi}| \sum_{i=1}^{p} |y_{t-i}| + \overline{Q}_1 \left(\overline{\omega} + \overline{\alpha} \sum_{i=1}^{q} |y_{t-i}| \right) \right] < \infty,$$

$$\begin{aligned} E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{II}}) \sup_{\boldsymbol{\psi}^{\mathrm{II}} \in \Psi^{\mathrm{II}}} \|\dot{g}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{II}})\|^3 \right] &\leqslant CE \sup_{\boldsymbol{\psi}^{\mathrm{II}} \in \Psi^{\mathrm{II}}} \|\dot{g}_{t,\tau_k}^{\mathrm{II}}(\boldsymbol{\psi}^{\mathrm{II}})\|^3 \\ &\leqslant CE\left[\|\boldsymbol{Y}_{t-1}\|^3 + \overline{Q}_1^3 \|\boldsymbol{X}_{t-1}\|^3 + \overline{Q}_2^3 \left(\overline{\omega} + \overline{\alpha} \sum_{i=1}^{q} |y_{t-i}| \right)^3 \right] \\ &\leqslant CE(|y_t|^3) < \infty, \end{aligned}$$

and

$$E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{II}}) \sup_{\boldsymbol{\psi}^{\mathrm{II}} \in \Psi^{\mathrm{II}}} \|\ddot{g}_{t,\tau_k}(\boldsymbol{\psi}^{\mathrm{II}})\|^2 \right] \leqslant CE\left[\|\boldsymbol{X}_{t-1}\|^2 \right] \leqslant CE(y_t^2) < \infty.$$

Then Assumption 5 holds and the proof is accomplished. □

S6.3 Proof of Theorems 1-4 in ESTAR-GARCH models

Similar to Section S6.1, we only verify that $q_{t,\tau_k}(\boldsymbol{\phi})$ and $g_{t,\tau_k}(\boldsymbol{\psi})$ for ESTAR-GARCH model (2.4) with Assumptions 4‴-5‴ imply Assumptions 4-5.

Proof. For ESTAR-GARCH model (2.4), recall that $\mu_t(\boldsymbol{\vartheta}^{\mathrm{III}}) = \alpha_{00} + \alpha_{10} G(y_{t-d}; \gamma, c) + \sum_{i=1}^{p} [\alpha_{0i} + \alpha_{1i} G(y_{t-d}; \gamma, c)] y_{t-i}$ and $h_t(\boldsymbol{\vartheta}^{\mathrm{III}}) = (\omega + a\epsilon_{t-1}^2 + b h_{t-1}^2)^{1/2} = \left[\omega/(1-b) + a \sum_{i=1}^{\infty} b^{i-1} \epsilon_{i-1}^2 \right]^{1/2}$, where $\boldsymbol{\vartheta}^{\mathrm{III}} = (\alpha_{00}, \alpha_{01}, \ldots, \alpha_{0p}, \alpha_{10}, \alpha_{11}, \ldots, \alpha_{1p}, \gamma, c, \omega, a, b)'$.

We first verify that $q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})$ and $\widetilde{q}_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})$ for model (2.4) with Assumptions 4‴-5‴ imply Assumptions 4-6. Note that $\omega = 1$ by Assumption 4‴ (i) for semi-parametric CQR in model (2.4), then we rewrite the model parameter by $\boldsymbol{\vartheta}_*^{\mathrm{III}} = (\alpha_{00}, \alpha_{01}, \ldots, \alpha_{0p}, \alpha_{10}, \alpha_{11}, \ldots, \alpha_{1p}, \gamma, c, a, b)'$. Thus it holds that $q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}}) = \mu_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) + b_k h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})$, where $\boldsymbol{\phi}^{\mathrm{III}} = (\boldsymbol{\vartheta}_*^{\mathrm{III}\prime}, b_1, \ldots, b_K)'$. Obviously, $q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})$ is continuous in $\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}$ and Assumption 4 (i) holds, where

$\Phi^{\mathrm{III}} = \mathbb{R}^{2p+2} \times \mathbb{R}^{+} \times \mathbb{R} \times \mathbb{R}^{+^2} \times \mathbb{R}^K$ is the parameter space. Denote the true parameter by $\boldsymbol{\phi}_0^{\mathrm{III}} = (\boldsymbol{\vartheta}_{*0}^{\mathrm{III}\prime}, b_{10}, \ldots, b_{K0})'$ with $\boldsymbol{\vartheta}_{*0}^{\mathrm{III}} = (\alpha_{00}^*, \alpha_{01}^*, \ldots, \alpha_{0p}^*, \alpha_{10}^*, \alpha_{11}^*, \ldots, \alpha_{1p}^*, \gamma_0, c_0, a_0, b_0)'$. Define $f(z) = \alpha_{10}e^{-(z^d-c)/\gamma} + \sum_{i=1}^p [\alpha_{0i} + \alpha_{1i}e^{-(z^d-c)/\gamma}]z^i$ and $f_0(z) = \alpha_{10}^* e^{-(z^d-c_0)/\gamma_0} + \sum_{i=1}^p [\alpha_{0i}^* + \alpha_{1i}^* e^{-(z^d-c_0)/\gamma_0}]z^i$, then we have $\epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) = -\alpha_{00} + [1 - f(B)]y_t$ and $\epsilon_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}}) = -\alpha_{00}^* + [1 - f_0(B)]y_t$. For Assumption 4 (ii), if $q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}}) = q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}_0^{\mathrm{III}})$, we can write $q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}}) = y_t - \epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) + b_k h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})$. Then if $q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}}) = q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}_0^{\mathrm{III}})$, it holds that

$$\epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) - \epsilon_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}}) - b_k h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) + b_{k0} h_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}}) = 0. \tag{S6.15}$$

Denote $\epsilon_t = \epsilon_t(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}})$, then by Assumption 1, we have

$$\epsilon_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) - \epsilon_t = m_0 + \sum_{i=1}^{\infty} m_i \epsilon_{t-i},$$

where $m_0 = -\alpha_{00} + \alpha_{00}^*[1 - f(z)]/[1 - f_0(z)]$ and $1 + \sum_{i=1}^{\infty} m_i z^i = [1 - f(z)]/[1 - f_0(z)]$. Then similar to (S6.2)-(S6.5), we can prove that $\alpha_{ij} = \alpha_{ij}^*$ for $i = 0, 1$ and $0 \leqslant j \leqslant p$, $\gamma = \gamma_0$, $c = c_0$, $a = a_0$, $b = b_0$ and $b_k = b_{k0}$ for $1 \leqslant k \leqslant K$ and then Assumption 4 (ii) holds.

Denote $c_i = ab^{i-1}$, $\dot{c}_i = \partial c_i / \partial \boldsymbol{\phi}^{\mathrm{III}}$ and $\ddot{c}_i = \partial^2 c_i / (\partial \boldsymbol{\phi}^{\mathrm{III}} \partial \boldsymbol{\phi}^{\mathrm{III}\prime})$ for each $i \geqslant 1$. Then under Assumptions 1, 2 and 5‴ (i), together with (S6.6), we have

$$\begin{aligned} &(i) \sup_{\boldsymbol{\phi}^{\mathrm{III}}} c_i \leqslant C\rho^i, \sup_{\boldsymbol{\phi}^{\mathrm{III}}} \|\dot{c}_i\| \leqslant C\rho^i, \sup_{\boldsymbol{\phi}^{\mathrm{III}}} \|\ddot{c}_i\| \leqslant C\rho^i; \\ &(ii) c_i \geqslant C\rho^i, \end{aligned} \tag{S6.16}$$

for some constants $C > 0$ and $0 < \rho < 1$. Since $\underline{\alpha} \leqslant \alpha_{ij} \leqslant \overline{\alpha}$ for $i = 0, 1$ and $0 \leqslant j \leqslant p$, $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5‴ (i), $E(|y_t|) < \infty$ by Assumption 5‴ (ii), together with the fact that $|G(y_{t-d}; \gamma, c)| < 1$ and (S6.16), it follows that

$$\begin{aligned} E \sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left| q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}}) \right| &\leqslant E \left[2\overline{\alpha}(1 + \sum_{i=1}^{p} |y_{t-i}|) + \overline{b} + \overline{b} \sum_{i=1}^{\infty} \sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \sqrt{c_i} |\epsilon_{t-i}| \right] \\ &\leqslant CE(|y_t|)(1 + \sum_{i=1}^{\infty} \rho^{i/2}) < \infty. \end{aligned} \tag{S6.17}$$

Since $0 < \underline{\gamma} \leqslant \gamma \leqslant \overline{\gamma}$ by Assumption 5‴ (i), together with the facts that $x, x^2 < e^x$ for $x > 0$, $|xe^{-x^2/\gamma}| \leqslant \gamma^{-1/2}$ and $|x^3 e^{-x^2/\gamma}| \leqslant \gamma^{3/2}$, it holds that

$$\begin{aligned}
\left|\frac{\partial G(y_{t-d};\gamma,c)}{\partial \gamma}\right| &= \frac{(y_{t-d}-c)^2}{\gamma^2} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} \leqslant \frac{1}{\gamma} \leqslant \frac{1}{\underline{\gamma}}, \\
\left|\frac{\partial G(y_{t-d};\gamma,c)}{\partial c}\right| &= \frac{|2(y_{t-d}-c)|}{\gamma} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} \leqslant \underline{\gamma}^{-1/2}, \\
\left|\frac{\partial^2 G(y_{t-d};\gamma,c)}{\partial \gamma^2}\right| &\leqslant \frac{2(y_{t-d}-c)^2}{\gamma^3} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} + \frac{(y_{t-d}-c)^4}{\gamma^4} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} \leqslant \frac{3}{\underline{\gamma}^2}, \\
\left|\frac{\partial^2 G(y_{t-d};\gamma,c)}{\partial c^2}\right| &\leqslant \frac{2}{\gamma} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} + \frac{|4(y_{t-d}-c)^2|}{\gamma^2} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} \leqslant \frac{6}{\underline{\gamma}}, \\
\left|\frac{\partial^2 G(y_{t-d};\gamma,c)}{\partial \gamma \partial c}\right| &\leqslant \frac{2(y_{t-d}-c)}{\gamma^2} e^{-\frac{(y_{t-d}-c)^2}{\gamma}} + \frac{2(y_{t-d}-c)^3}{\gamma^3} e^{-\frac{(y_{t-d}-c)^3}{\gamma}} \leqslant 2\underline{\gamma}^{-1} + \underline{\gamma}^{-3/2}. \quad \text{(S6.18)}
\end{aligned}$$

These together with $\underline{\alpha} \leqslant \alpha_{ij} \leqslant \overline{\alpha}$ for $i = 0, 1$ and $0 \leqslant j \leqslant p$, $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5‴ (i) and (S6.16), imply that

$$\begin{aligned}
\sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left|h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})\right| \leqslant & C\left(1 + \sum_{i=1}^{\infty} \rho^{i/2}|y_{t-i}|\right), \\
\sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left\|\frac{\partial \mu_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\| \leqslant & \sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left[1 + G(y_{t-d};\gamma,c) + \overline{\alpha}\left(\left|\frac{\partial G(y_{t-d};\gamma,c)}{\partial \gamma}\right| + \left|\frac{\partial G(y_{t-d};\gamma,c)}{\partial c}\right|\right)\right]\left(1 + \sum_{i=1}^{p}|y_{t-i}|\right) \\
\leqslant & C\left(1 + \sum_{i=1}^{p}|y_{t-i}|\right), \\
\sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\| \leqslant & C + \sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \sum_{i=1}^{\infty} \|\dot{c}_i/\sqrt{c_i}\||\epsilon_{t-i}| + \sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \sum_{i=1}^{\infty} \sqrt{c_i}\left\|\frac{\partial \mu_{t-i}(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\| \\
\leqslant & C\left(1 + \sum_{i=1}^{\infty} \rho^{i/2}|y_{t-i}|\right), \\
\sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left\|\frac{\partial^2 \mu_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}_*^{\mathrm{III}} \partial \boldsymbol{\vartheta}_*^{\mathrm{III}\prime}}\right\| \leqslant & \sup_{\boldsymbol{\phi}^{\mathrm{III}} \in \Phi^{\mathrm{III}}} \left[\overline{\alpha}\left(\left|\frac{\partial^2 G(y_{t-d};\gamma,c)}{\partial \gamma^2}\right| + 2\left|\frac{\partial^2 G(y_{t-d};\gamma,c)}{\partial \gamma \partial c}\right| + \left|\frac{\partial^2 G(y_{t-d};\gamma,c)}{\partial c^2}\right|\right)\right. \\
& \left. + 2\left(\left|\frac{\partial G(y_{t-d};\gamma,c)}{\partial \gamma}\right| + \left|\frac{\partial G(y_{t-d};\gamma,c)}{\partial c}\right|\right)\right]\left(1 + \sum_{i=1}^{p}|y_{t-i}|\right) \\
\leqslant & C\left(1 + \sum_{i=1}^{p}|y_{t-i}|\right),
\end{aligned}$$

$$\begin{aligned}
\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left\|\frac{\partial^2 h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}\partial\boldsymbol{\vartheta}_*^{\mathrm{III}\prime}}\right\| \leqslant & C + C\sum_{i=1}^{\infty}\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left(\|\dot{c}_i/\sqrt{c_i}\| + \sqrt{c_i}\right)\left\|\frac{\partial\mu_{t-i}(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\| \\
& + C\sum_{i=1}^{\infty}\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left(\|\ddot{c}_i/\sqrt{c_i}\| + \|\dot{c}_i/\sqrt{c_i}\|\right)|\epsilon_{t-i}| \\
& + C\sum_{i=1}^{\infty}\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\sqrt{c_i}\left\|\frac{\partial^2\mu_{t-i}(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}\partial\boldsymbol{\vartheta}_*^{\mathrm{III}\prime}}\right\| \leqslant C\left(1+\sum_{i=1}^{\infty}\rho^{i/2}|y_{t-i}|\right). \quad \text{(S6.19)}
\end{aligned}$$

Since $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5‴ (i) and $E(y_t^2) < \infty$ by Assumption 5‴ (ii), together with (S6.19), it holds that

$$\begin{aligned}
& E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}})\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left\|\dot{q}_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})\right\|^3\right] \\
\leqslant & E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}})\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left(\left\|\frac{\partial\mu_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\|^3 + \overline{b}^3\left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\|^3 + |h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})|^3\right)\right] \\
\leqslant & E\left[C\left(1+\sum_{i=1}^{p}y_{t-i}^2\right) + C\left(1+\sum_{i=1}^{\infty}\rho^i y_{t-i}^2\right)\right] \\
\leqslant & C\left[1+E(y_t^2)\right] < \infty.
\end{aligned}$$

and

$$\begin{aligned}
& E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}})\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left\|\ddot{q}_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})\right\|^2\right] \\
\leqslant & E\left[h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\mathrm{III}})\sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}}\left(\left\|\frac{\partial^2\mu_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}\partial\boldsymbol{\vartheta}_*^{\mathrm{III}\prime}}\right\|^2 + \overline{b}^2\left\|\frac{\partial^2 h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}\partial\boldsymbol{\vartheta}_*^{\mathrm{III}\prime}}\right\|^2 + 4\left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})}{\partial\boldsymbol{\vartheta}_*^{\mathrm{III}}}\right\|^2\right)\right] \\
\leqslant & E\left[C\left(1+\sum_{i=1}^{p}|y_{t-i}|\right) + C\left(1+\sum_{i=1}^{\infty}\rho^{i/2}|y_{t-i}|\right)\right] \\
\leqslant & C\left(1+E|y_t|\right) < \infty,
\end{aligned}$$

then Assumption 5 holds.

For Assumption 6, similar to (S6.10)-(S6.11), it holds that $\mu_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) = \widetilde{\mu}_t(\boldsymbol{\vartheta}_*^{\mathrm{III}})$,

$$h_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) - \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\mathrm{III}}) \leqslant \sum_{i=t}^{\infty}\sqrt{c_i}|\epsilon_{t-i}|,$$

and

$$\left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\rm III})}{\partial \boldsymbol{\vartheta}_*^{\rm III}} - \frac{\partial \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\rm III})}{\partial \boldsymbol{\vartheta}_*^{\rm III}}\right\| \leqslant C + \sum_{i=1}^{\infty} \|\dot{c}_i\| \, |\epsilon_{t-i}|/\sqrt{c_i} + \sum_{i=1}^{\infty} \sqrt{c_i} \left\|\frac{\partial \mu_{t-i}(\boldsymbol{\vartheta}_*^{\rm III})}{\partial \boldsymbol{\vartheta}_*^{\rm III}}\right\| + \sum_{i=1}^{t-1} \|\dot{c}_i\| \, |\epsilon_{t-i}|/\sqrt{c_i} + \sum_{i=1}^{t-1} \sqrt{c_i} \left\|\frac{\partial \mu_{t-i}(\boldsymbol{\vartheta}_*^{\rm III})}{\partial \boldsymbol{\vartheta}_*^{\rm III}}\right\|.$$

These together with $E(y_t^2) < \infty$ by Assumption 5‴ (ii) and $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5‴ (i), imply that

$$\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\rm III}) \sup_{\boldsymbol{\phi}^{\rm III} \in \Phi^{\rm I}} |q_{t,\tau_k}^{\rm III}(\boldsymbol{\phi}^{\rm III}) - \widetilde{q}_{t,\tau_k}^{\rm III}(\boldsymbol{\phi}^{\rm III})|^2 \leqslant C \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\rm III}) \sup_{\boldsymbol{\phi}^{\rm III} \in \Phi^{\rm III}} \overline{b}^2 |h_t(\boldsymbol{\vartheta}_*^{\rm III}) - \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\rm III})|^2 \leqslant C \sum_{t=1}^{\infty} \rho^t \varsigma_\rho < \infty,$$

and

$$\begin{aligned}
&\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\rm III}) \sup_{\boldsymbol{\phi}^{\rm III} \in \Phi^{\rm III}} \|\dot{q}_{t,\tau_k}^{\rm III}(\boldsymbol{\phi}^{\rm III}) - \dot{\widetilde{q}}_{t,\tau_k}^{\rm III}(\boldsymbol{\phi}^{\rm III})\|^2 \\
\leqslant &C \sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_{*0}^{\rm III}) \sup_{\boldsymbol{\phi}^{\rm III} \in \Phi^{\rm III}} \left(\overline{b}^2 \left\|\frac{\partial h_t(\boldsymbol{\vartheta}_*^{\rm III})}{\partial \boldsymbol{\vartheta}_*^{\rm III}} - \frac{\partial \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\rm III})}{\partial \boldsymbol{\vartheta}_*^{\rm III}}\right\|^2 + |h_t(\boldsymbol{\vartheta}_*^{\rm III}) - \widetilde{h}_t(\boldsymbol{\vartheta}_*^{\rm III})|^2 \right) \\
\leqslant &C \sum_{t=1}^{\infty} \rho^t \varsigma_\rho < \infty,
\end{aligned}$$

where $0 < \rho < 1$ and $\varsigma_\rho = \sum_{t=0}^{\infty} \rho^t |y_{-t}|$. Then Assumption 6 holds.

Then we prove that $g_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}^{\rm III})$ and $\widetilde{g}_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}^{\rm III})$ for model (2.4) with Assumptions 4‴-5‴ imply Assumptions 4-6. For model (2.4), $g_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}^{\rm III})$ has the form of

$$g_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}^{\rm III}) = \mu_t(\boldsymbol{\vartheta}^{\rm III}) + Q_{\tau_k}(\lambda) h_t(\boldsymbol{\vartheta}^{\rm III}),$$

where $\boldsymbol{\psi}^{\rm III} = (\boldsymbol{\vartheta}^{\rm III\prime}, \lambda)'$. Obviously, $g_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}^{\rm III})$ is continuous in $\boldsymbol{\psi}^{\rm III} \in \Psi^{\rm III}$ and Assumption 4 (i) holds, where $\Psi^{\rm III} = \mathbb{R}^{2p+2} \times \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R}^{+^2} \times \mathbb{R}$. For Assumption 4 (ii), if $g_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}^{\rm III}) = g_{t,\tau_k}^{\rm III}(\boldsymbol{\psi}_0^{\rm III})$, similar to the proof of (S6.15) and (S6.12), it holds that $\boldsymbol{\psi}^{\rm III} = \boldsymbol{\psi}_0^{\rm III}$ under Assumption 4‴ (ii). Then Assumption 4 (ii) holds.

For Assumption 5, similar to (S6.19), we have $\left|h_t(\boldsymbol{\vartheta}^{\text{III}})\right| \leqslant C\left(1+\sum_{i=1}^{\infty}\rho^{i/2}|y_{t-i}|\right)$,

$$\sup_{\boldsymbol{\phi}^{\text{III}}\in\Phi^{\text{III}}}\left\|\frac{\partial\mu_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}}\right\| \leqslant C\left(1+\sum_{i=1}^{p}|y_{t-i}|\right), \quad \sup_{\boldsymbol{\phi}^{\text{III}}\in\Phi^{\text{III}}}\left\|\frac{\partial h_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}}\right\| \leqslant C\left(1+\sum_{i=1}^{\infty}\rho^{i/2}|y_{t-i}|\right),$$
$$\sup_{\boldsymbol{\phi}^{\text{III}}\in\Phi^{\text{III}}}\left\|\frac{\partial^2\mu_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}\partial\boldsymbol{\vartheta}^{\text{III}\prime}}\right\| \leqslant C\left(1+\sum_{i=1}^{p}|y_{t-i}|\right), \text{ and } \sup_{\boldsymbol{\phi}^{\text{III}}\in\Phi^{\text{III}}}\left\|\frac{\partial^2 h_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}\partial\boldsymbol{\vartheta}^{\text{III}\prime}}\right\| \leqslant C\left(1+\sum_{i=1}^{\infty}\rho^{i/2}|y_{t-i}|\right).$$

These together with $\underline{b} \leqslant b_k \leqslant \overline{b}$ for $1 \leqslant k \leqslant K$ by Assumption 5‴ (i), $E(y_t^2) < \infty$ by Assumption 5‴ (ii) and (S6.13), imply that

$$E\sup_{\boldsymbol{\psi}^{\text{III}}\in\Psi^{\text{III}}}|g_{t,\tau_k}^{\text{III}}(\boldsymbol{\psi}^{\text{III}})| \leqslant E\sup_{\boldsymbol{\psi}^{\text{III}}\in\Psi^{\text{III}}}\left[|\mu_t(\boldsymbol{\vartheta}^{\text{III}})| + \overline{Q}_1|h_t(\boldsymbol{\vartheta}^{\text{III}})|\right] \leqslant CE(|y_t|)\sum_{i=1}^{\infty}(1+\rho^{i/2}) < \infty,$$
$$\begin{aligned}&E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\text{III}})\sup_{\boldsymbol{\psi}^{\text{III}}\in\Psi^{\text{III}}}\|\dot{g}_{t,\tau_k}^{\text{III}}(\boldsymbol{\psi}^{\text{III}})\|^3\right]\\ \leqslant&E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\text{III}})\sup_{\boldsymbol{\psi}^{\text{III}}\in\Psi^{\text{III}}}\left(\left\|\frac{\partial\mu_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}}\right\|^3 + \overline{Q}_1^3\left\|\frac{\partial h_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}}\right\|^3 + |h_t(\boldsymbol{\vartheta}^{\text{III}})|^3\right)\right]\\ \leqslant&CE(y_t^2)\sum_{i=1}^{\infty}(1+\rho^{i}) < \infty,\end{aligned}$$

and

$$\begin{aligned}&E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\text{III}})\sup_{\boldsymbol{\psi}^{\text{III}}\in\Psi^{\text{III}}}\|\ddot{g}_{t,\tau_k}^{\text{III}}(\boldsymbol{\phi}^{\text{III}})\|^2\right]\\ \leqslant&E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{\text{III}})\sup_{\boldsymbol{\psi}^{\text{III}}\in\Psi^{\text{III}}}\left(\left\|\frac{\partial^2\mu_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}\partial\boldsymbol{\vartheta}^{\text{I}\prime}})\right\|^2 + \overline{Q}_1^2\left\|\frac{\partial^2 h_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}\partial\boldsymbol{\vartheta}^{\text{I}\prime}})\right\|^2 + \overline{Q}_2^2\left\|\frac{\partial h_t(\boldsymbol{\vartheta}^{\text{III}})}{\partial\boldsymbol{\vartheta}^{\text{III}}}\right\|^2\right)\right]\\ \leqslant&CE(|y_t|)\sum_{i=1}^{\infty}(1+\rho^{i/2}) < \infty,\end{aligned}$$

then Assumption 5 holds.

For Assumption 6, similar to (S6.10)-(S6.11), it holds that $\mu_t(\boldsymbol{\vartheta}^{\text{III}}) = \widetilde{\mu}_t(\boldsymbol{\vartheta}^{\text{III}})$,

$$h_t(\boldsymbol{\vartheta}^{\text{III}}) - \widetilde{h}_t(\boldsymbol{\vartheta}^{\text{III}}) \leqslant \sum_{i=t}^{\infty}\sqrt{c}_i|\epsilon_{t-i}|,$$

and

$$\left\|\frac{\partial h_t(\boldsymbol{\vartheta}^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}^{\mathrm{III}}}-\frac{\partial \widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}^{\mathrm{III}}}\right\| \leqslant C+\sum_{i=1}^{\infty}\|\dot{c}_i\|\,|\epsilon_{t-i}|/\sqrt{c_i}+\sum_{i=1}^{\infty}\sqrt{c_i}\left\|\frac{\partial \mu_{t-i}(\boldsymbol{\vartheta}^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}^{\mathrm{III}}}\right\| \\ +\sum_{i=1}^{t-1}\|\dot{c}_i\|\,|\epsilon_{t-i}|/\sqrt{c_i}+\sum_{i=1}^{t-1}\sqrt{c_i}\left\|\frac{\partial \mu_{t-i}(\boldsymbol{\vartheta}^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}^{\mathrm{III}}}\right\|.$$

These together with $E(y_t^2)<\infty$ by Assumption 5‴ (ii) and (S6.13), imply that

$$\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{III}}) \sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{I}}} |q_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})-\widetilde{q}_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})|^2 \leqslant C\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{III}}) \sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}} \overline{Q}_1^2|h_t(\boldsymbol{\vartheta}^{\mathrm{III}})-\widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{III}})|^2 \\ \leqslant C\sum_{t=1}^{\infty}\rho^t\varsigma_\rho<\infty,$$

and

$$\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{III}}) \sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}} \|\dot{q}_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})-\dot{\widetilde{q}}_{t,\tau_k}^{\mathrm{III}}(\boldsymbol{\phi}^{\mathrm{III}})\|^2 \\ \leqslant C\sum_{t=1}^{\infty} h_t^{-1}(\boldsymbol{\vartheta}_0^{\mathrm{III}}) \sup_{\boldsymbol{\phi}^{\mathrm{III}}\in\Phi^{\mathrm{III}}} \left(\overline{Q}_1^2\left\|\frac{\partial h_t(\boldsymbol{\vartheta}^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}^{\mathrm{III}}}-\frac{\partial \widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{III}})}{\partial \boldsymbol{\vartheta}^{\mathrm{III}}}\right\|^2+\overline{Q}_2^2|h_t(\boldsymbol{\vartheta}^{\mathrm{III}})-\widetilde{h}_t(\boldsymbol{\vartheta}^{\mathrm{III}})|^2\right) \\ \leqslant C\sum_{t=1}^{\infty}\rho^t\varsigma_\rho<\infty,$$

where $0<\rho<1$ and $\varsigma_\rho=\sum_{t=0}^{\infty}\rho^t|y_{-t}|$. Then Assumption 6 holds and the proof is accomplished. □

S7 The selection consistency for the proposed BIC

Theorem 1. *Let (p_0,q_0) be the true order and $m_{\max}$ be a predetermined positive integer. For $(\widehat{p}_n,\widehat{q}_n)=\arg\min_{1\leqslant p,q\leqslant m_{\max}} BIC(p,q)$, if Assumptions 1-6 hold and $p_0,q_0\leqslant m_{\max}$, then $Pr(\widehat{p}_n=p_0,\widehat{q}_n=q_0)\to 1$ as $n\to\infty$.*

Proof. Recall that the BIC proposed in Section S4.2 has the form of

$$\mathrm{BIC}(p,q)=2(n-m_{\max})\log\widetilde{L}_n(\boldsymbol{v})+d\log(n-m_{\max}), \tag{S7.1}$$

where $\widetilde{L}_n(\boldsymbol{\upsilon}) = (n - m_{\max})^{-1} \sum_{t=m_{\max}+1}^{n} \widetilde{l}_t(\boldsymbol{\upsilon})$ with $\widetilde{l}_t(\boldsymbol{\upsilon}) = \sum_{k=1}^{K} \rho_{\tau_k}(y_t - \widetilde{\zeta}_{t,\tau_k}(\boldsymbol{\upsilon}))$, $\widetilde{\zeta}_{t,\tau_k}(\cdot)$ is the feasible conditional quantile function of both CQRs for ARMA-GARCH, ALDAR and ESTAR-GARCH models. Denote the true order by (p_0, q_0), then it suffices to show that the following result holds for any $(p, q) \neq (p_0, q_0)$:

$$\lim_{n\to\infty} (\text{BIC}(p,q) - \text{BIC}(p_0,q_0) > 0) = 1. \tag{S7.2}$$

Denote $\widehat{\boldsymbol{\upsilon}}_n^{p,q}$ (or $\widehat{\boldsymbol{\upsilon}}_n^{p_0,q_0}$) as the semi-parametric or parametric CQR estimator for ARMA-GARCH, ALDAR or ESTAR-GARCH models with the order (p,q) (or (p_0,q_0)). Then by (S7.1), it holds that

$$\text{BIC}(p,q) - \text{BIC}(p_0,q_0) = 2(n - m_{\max}) \left[\log \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p,q}) - \log \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p_0,q_0})\right] + (d - d_0)\log(n - m_{\max}),$$

where d is the dimension of $\widehat{\boldsymbol{\upsilon}}_n^{p,q}$, d_0 is the dimension of $\widehat{\boldsymbol{\upsilon}}_n^{p_0,q_0}$, and $m_{\max}$ is a predetermined positive integer. Moreover, denote $\Upsilon^{p,q}$ and $\boldsymbol{\upsilon}_0^{p,q}$ as the parameter space and true parameter of $\boldsymbol{\upsilon}$ with the order set to (p,q), respectively. Below we prove the selection consistency of BIC for the semi-parametric CQR, and the proof also applies to the parametric CQR. To verify (S7.2), we next consider two cases.

Case I (overfitting): $p \geqslant p_0$ and $q \geqslant q_0$, and at least one inequality holds. In this case, it holds that $d - d_0 > 0$, which implies that $(d - d_0)\log(n - m_{\max}) \to \infty$ as $n \to \infty$. Then by Taylor expansion, we have

$$\log \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p,q}) - \log \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p_0,q_0}) = \frac{1}{\widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}^*)} \left[\widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p,q}) - \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p_0,q_0})\right],$$

where $\boldsymbol{\upsilon}^*$ is between $\boldsymbol{\upsilon}_n^{p,q}$ and $\boldsymbol{\upsilon}_n^{p_0,q_0}$. By ergodic theorem, together with the stationarity and ergodicity of $\{y_t\}$ by Assumption 1 and Lemma 1, it holds that

$$\sup_{\boldsymbol{\upsilon}^* \in \Upsilon} |\widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}^*)| = \sup_{\boldsymbol{\upsilon}^* \in \Upsilon} |L_n(\widehat{\boldsymbol{\upsilon}}^*)| + o_p(1) \leqslant E\left[\sup_{\boldsymbol{\upsilon}^* \in \Upsilon} |l_t(\boldsymbol{\upsilon}^*)|\right] + o_p(1) < \infty,$$

which implies that

$$\widetilde{L}_n^{-1}(\widehat{\boldsymbol{v}}^*) = O_p(1). \tag{S7.3}$$

Rewrite $\widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p,q}) - \widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p_0,q_0})$ as follows:

$$\begin{aligned}\widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p,q}) - \widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p_0,q_0}) &= \left[\widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p,q}) - L_n(\boldsymbol{v}_0^{p,q})\right] + \left[L_n(\boldsymbol{v}_0^{p,q}) - L_n(\boldsymbol{v}_0^{p_0,q_0})\right] \\ &\quad + \left[L_n(\boldsymbol{v}_0^{p,q}) - \widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p_0,q_0})\right],\end{aligned} \tag{S7.4}$$

where $L_n(\boldsymbol{v}) = (n - m_{\max})^{-1}\sum_{t=m_{\max}+1}^{n} l_t(\boldsymbol{v})$ with $l_t(\boldsymbol{v}) = \sum_{k=1}^{K}\rho_{\tau_k}(y_t - \zeta_{t,\tau_k}(\boldsymbol{v}))$. Note that the model with order (p,q) in Case I corresponds to a bigger model, and then it holds that $l_t(\boldsymbol{v}_0^{p,q}) = l_t(\boldsymbol{v}_0^{p_0,q_0})$ and

$$L_n(\boldsymbol{v}_0^{p,q}) = L_n(\boldsymbol{v}_0^{p_0,q_0}). \tag{S7.5}$$

Moreover, by Theorem 1, it holds that $\widehat{\boldsymbol{v}}_n^{p,q} \to \boldsymbol{v}_0^{p,q}$ and $\widehat{\boldsymbol{v}}_n^{p_0,q_0} \to \boldsymbol{v}_0^{p_0,q_0}$ as $n \to \infty$. From (S5.39) in the proof of Theorem 2, we have

$$n\left[\widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p,q}) - \widetilde{L}_n(\boldsymbol{v}_0^{p,q})\right] = -\sqrt{n}\widehat{\boldsymbol{u}}_n^{p,q\prime}\boldsymbol{T}_n^{p,q} + \sqrt{n}\widehat{\boldsymbol{u}}_n^{p,q\prime}\Sigma_1^{p,q}\sqrt{n}\widehat{\boldsymbol{u}}_n + o_p(\sqrt{n}\|\widehat{\boldsymbol{u}}_n^{p,q}\| + n\|\widehat{\boldsymbol{u}}_n^{p,q}\|^2),$$

where $\widehat{\boldsymbol{u}}_n^{p,q} = \widehat{\boldsymbol{v}}_n^{p,q} - \boldsymbol{v}_0^{p,q}$, $\boldsymbol{T}_n^{p,q} = n^{-1/2}\sum_{t=1}^{n}\sum_{k=1}^{K}\dot{\zeta}_{t,\tau_k}(\boldsymbol{v}_0^{p,q})\psi_{\tau_k}(y_t - \zeta_{t,\tau_k}(\boldsymbol{v}_0^{p,q}))$, and $\Sigma_1^{p,q} = \sum_{k=1}^{K} f(b_{\tau_k})E\left[h_t^{-1}(\boldsymbol{\vartheta}_0^{p,q})\dot{\zeta}_{t,\tau_k}(\boldsymbol{v}_0^{p,q})\dot{\zeta}_{t,\tau_k}'(\boldsymbol{v}_0^{p,q})\right]/2$ with $\boldsymbol{\vartheta}_0^{p,q}$ being the true model parameter vector with the order (p,q). From the proof of Theorem 2, we have $\sqrt{n}\|\widehat{\boldsymbol{u}}_n^{p,q}\| = O_p(1)$, $\|\boldsymbol{T}_n^{p,q}\| = O_p(1)$ and $\|\Sigma_1^{p,q}\| = O_p(1)$, which imply that $n\left[\widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p,q}) - \widetilde{L}_n(\boldsymbol{v}_0^{p,q})\right] = O_p(1)$. This together with Lemma 1 (3) implies that

$$\widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p,q}) - L_n(\boldsymbol{v}_0^{p,q}) = O_p(n^{-1}) \quad \text{and} \quad L_n(\boldsymbol{v}_0^{p_0,q_0}) - \widetilde{L}_n(\widehat{\boldsymbol{v}}_n^{p_0,q_0}) = O_p(n^{-1}). \tag{S7.6}$$

Then combining (S7.3)-(S7.6), we have

$$\text{BIC}(p,q) - \text{BIC}(p_0,q_0) = O_p(1) + (d - d_0)\log(n - m_{\max}) \to \infty \quad \text{as} \quad n \to \infty.$$

As a result, (S7.2) holds for Case I.

Case II (underfitting): $p < p_0$ or $q < q_0$. Let $p_* = \max(p, p_0)$ and $q_* = \max(q, q_0)$. Denote $\boldsymbol{\upsilon}_{0*}^{p,q}$ (or $\boldsymbol{\upsilon}_{0*}^{p_0,q_0}$) as the parameter vector with the order set to (p_*, q_*), including $\boldsymbol{\upsilon}_{0}^{p,q}$ (or $\boldsymbol{\upsilon}_{0}^{p_0,q_0}$) as its subvector at the corresponding locations and zeroes at the remaining locations. Since $E\left[l_t(\boldsymbol{\upsilon})\right]$ has a unique minimum at $\boldsymbol{\upsilon}_{0}^{p_0,q_0}$ by Lemma 1, the following result holds for some constant $\delta > 0$:

$$E\left[l_t(\boldsymbol{\upsilon}_{0}^{p,q})\right] - E\left[l_t(\boldsymbol{\upsilon}_{0}^{p_0,q_0})\right] = E\left[l_t(\boldsymbol{\upsilon}_{0*}^{p,q})\right] - E\left[l_t(\boldsymbol{\upsilon}_{0*}^{p_0,q_0})\right] > \delta.$$

This together with ergodic theorem, implies that

$$L_n(\boldsymbol{\upsilon}_{0}^{p,q}) - L_n(\boldsymbol{\upsilon}_{0}^{p_0,q_0}) > o_p(1) + \delta. \tag{S7.7}$$

Similar to the proof of (S7.6), it holds that

$$L_n(\boldsymbol{\upsilon}_{0}^{p_0,q_0}) - \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_{n}^{p_0,q_0}) = O_p(n^{-1}). \tag{S7.8}$$

Assume that $E\left[l_t(\boldsymbol{\upsilon}^{p,q})\right]$ has a unique minimum at $\boldsymbol{\upsilon}_{0}^{p,q}$ on $\Upsilon^{p,q}$. Similar to the proof of Theorem 1, we can prove that $\widehat{\boldsymbol{\upsilon}}_{n}^{p,q} \to \boldsymbol{\upsilon}_{0}^{p,q}$ as $n \to \infty$, which implies that $\|\widehat{\boldsymbol{\upsilon}}_{n}^{p,q} - \boldsymbol{\upsilon}_{0}^{p,q}\| = o_p(1)$. Then by Taylor expansion and ergodic theorem, together with the fact that $|\rho_\tau(x)| \leqslant |x|$ and $E\left(\sup_{\upsilon^{p,q}\in\Upsilon^{p,q}} \|\dot{\tilde{\zeta}}_{t,\tau_k}(\boldsymbol{\upsilon}^{p,q})\|\right) < \infty$ by Assumption 5 (ii), it holds that

$$\begin{aligned}
\widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_{n}^{p,q}) - \widetilde{L}_n(\boldsymbol{\upsilon}_{0}^{p,q}) &\leqslant \frac{1}{n - m_{\max}} \sum_{t=m_{\max}+1}^{n} \sum_{k=1}^{K} \left[\rho_{\tau_k}(y_t - \widetilde{\zeta}_{t,\tau_k}(\widehat{\boldsymbol{\upsilon}}_{n}^{p,q})) - \rho_{\tau_k}(y_t - \widetilde{\zeta}_{t,\tau_k}(\boldsymbol{\upsilon}_{0}^{p,q}))\right] \\
&\leqslant \frac{1}{n - m_{\max}} \sum_{t=m_{\max}+1}^{n} \sum_{k=1}^{K} |\widetilde{\zeta}_{t,\tau_k}(\widehat{\boldsymbol{\upsilon}}_{n}^{p,q}) - \widetilde{\zeta}_{t,\tau_k}(\boldsymbol{\upsilon}_{0}^{p,q})| \\
&\leqslant \sum_{k=1}^{K} \|\widehat{\boldsymbol{\upsilon}}_{n}^{p,q} - \boldsymbol{\upsilon}_{0}^{p,q}\| \left(\frac{1}{n - m_{\max}} \sum_{t=m_{\max}+1}^{n} \|\dot{\tilde{\zeta}}_{t,\tau_k}(\boldsymbol{\upsilon}_{\star}^{p,q})\|\right) \\
&\leqslant o_p(1) \sum_{k=1}^{K} \left[E\left(\sup_{\upsilon^{p,q}\in\Upsilon^{p,q}} \|\dot{\tilde{\zeta}}_{t,\tau_k}(\boldsymbol{\upsilon}^{p,q})\|\right) + o_p(1)\right] = o_p(1).
\end{aligned} \tag{S7.9}$$

where $\boldsymbol{\upsilon}_{\star}^{p,q}$ is between $\widehat{\boldsymbol{\upsilon}}_n^{p,q}$ and $\boldsymbol{\upsilon}_0^{p,q}$. Combining (S7.3)-(S7.9), we have

$$\log \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p,q}) - \log \widetilde{L}_n(\widehat{\boldsymbol{\upsilon}}_n^{p_0,q_0}) > O_p(n^{-1}) + o_p(1) + \delta.$$

Then as $n \to \infty$, it holds that

$$\text{BIC}(p,q) - \text{BIC}(p_0,q_0) = 2(n - m_{\max})\delta + O_p(1) + o_p(n) + (d - d_0)\log(n - m_{\max}) \to \infty.$$

As a result, (S7.2) holds for Case II and the proof is accomplished. □

S8 Additional simulation results

Due to space limitation, we only reported the results for ARMA-GARCH models for three simulation experiments in Section 4, with the results for estimating and predicting high conditional quantiles relegated to Table S.1. This section also provides additional results for ALDAR and ESTAR-GARCH models in three experiments.

For the first experiment, Tables S.2-S.5 list the biases, empirical standard deviations (ESDs), and asymptotic standard deviations (ASDs) of both CQRs for DGP2-DGP4. The following findings in Section 4 remain unchanged: (i) as the sample size increases, biases, ESDs, and ASDs generally decrease, and ESDs approach ASDs; (ii) most of the ASDs and ESDs increase as the distribution of η_t gets more heavy-tailed; (iii) the ASDs of the semi-parametric CQR using h_{HS} in (2.7) are slightly smaller compared to those using h_B, and closer to the corresponding ESDs; (iv) for the mis-specified situation of $Q_\tau(\boldsymbol{\lambda})$ that η_t follows F_N or F_{t_5} but the Tukey-lambda distribution is employed for $Q_\tau(\boldsymbol{\lambda})$, the biases of the parametric CQR estimator are still small, and the ESDs/ASDs are close to those of the semi-parametric CQR estimator.

For the second experiment, Tables S.6-S.9 report the biases and RMSEs of the in-sample estimation and out-of-sample prediction using the semi-parametric and parametric CQRs, GQMLE and EQMLE for DGP2 and DGP4. For the ALDAR and ESTAR-GARCH models, it can be found that (i) as the distribution of η_t becomes more heavy-tailed or the target quantile level τ gets more extreme, the biases and RMSEs of all the estimation methods generally increase, indicating that the accuracy of estimation and prediction decreases; (ii) the semi-parametric CQR and EQMLE perform similarly and they have better performance than the GQMLE; (iii) the parametric CQR outperforms the semi-parametric CQR when the quantile function $Q_\tau(\boldsymbol{\lambda})$ is correctly specified for the innovation η_t.

For the third experiment, Figures S.1 and S.2 plot the ARE$(\widehat{\boldsymbol{\vartheta}}_n, \widetilde{\boldsymbol{\vartheta}}_n)$, ARE$(\widehat{\boldsymbol{\vartheta}}_n, \check{\boldsymbol{\vartheta}}_n)$ and ARE$(\widehat{\boldsymbol{\vartheta}}_n, \breve{\boldsymbol{\vartheta}}_n)$ defined in Remark S3 for the ALDAR and ESTAR-GARCH models. The following findings in Section 4 remain unchanged: (i) as η_t becomes more heavy-tailed, ARE$(\widehat{\boldsymbol{\vartheta}}_n, \widetilde{\boldsymbol{\vartheta}}_n)$ gets smaller than one; (ii) the semi-parametric CQR is less efficient than GQMLE (or EQMLE) when η_t approximately follows the normal (or Laplace) distribution, but it tends to be more efficient than GQMLE and EQMLE when η_t becomes more heavy-tailed; (iii) when $\delta = 0$ such that $\eta_t \sim N(0,1)$, then ARE$(\widehat{\boldsymbol{\vartheta}}_n, \check{\boldsymbol{\vartheta}}_n) < 1$ and the GQMLE is the most efficient. (iv) when $\delta = 1$ and $m(x)$ is the pdf of a standard Laplace distribution such that η_t follows a standard Laplace distribution, then ARE$(\widehat{\boldsymbol{\vartheta}}_n, \breve{\boldsymbol{\vartheta}}_n) < 1$ and the EQMLE is the most efficient.

S9 Additional results for the empirical analysis

To save space, the ACF and PACF plots of $\{y_t\}$ are provided in Figure S.3, which imply that $\{y_t\}$ is serial correlated.

Bibliography

Chan, F., M. McAleer, and M. C. Medeiros (2015). Structure and asymptotic theory for nonlinear models with GARCH errors. *EconomiA 16*, 1–21.

Fan, J. and Q. Yao (2003). *Nonlinear time series: nonparametric and parametric methods.* New York: Springer.

Francq, C. and J.-M. Zakoïan (2004). Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli 10*, 605–637.

Hall, P. and C. C. Heyde (2014). *Martingale limit theory and its application.* Academic press.

Knight, K. (1998). Limiting distributions for L1 regression estimators under general conditions. *Annals of statistics 26*, 755–770.

Ling, S. and M. McAleer (2003). Asymptotic theory for a vector ARMA-GARCH model. *Econometric Theory 19*, 280–310.

Pollard, D. (1985). New ways to prove central limit theorems. *Econometric Theory 1*, 295–313.

Serfling, R. J. (2009). *Approximation theorems of mathematical statistics*, Volume 162. John Wiley & Sons.

Tan, S. and Q. Zhu (2022). Asymmetric linear double autoregression. *Journal of Time Series Analysis 43*, 371–388.

White, H. and I. Domowitz (1984). Nonlinear regression with dependent observations. *Econometrica 52*, 143–161.

Zou, H. and M. Yuan (2008). Composite quantile regression and the oracle model selection theory. *The Annals of Statistics 36*, 1108–1126.

Table S.1: Biases and RMSEs for estimating and predicting conditional quantiles at $\tau = 0.1\%, 0.5\%, 99.5\%$ and 99.9% for DGP1, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. M1, M2, M3 and M4 represent the semi-parametric CQR, parametric CQR, GQMLE and EQMLE, respectively.

		F_N				F_{t_5}				F_λ			
		Bias		RMSE		Bias		RMSE		Bias		RMSE	
τ	F	in	out	in	out	in	out	in	out	in	out	in	out
0.1%	M1	-0.124	-0.139	1.119	1.109	0.725	0.949	7.317	7.604	6.499	9.354	11.919	15.761
	M2	0.313	0.355	1.161	1.186	3.591	3.608	5.613	5.744	3.955	5.444	7.324	9.186
	M3	0.108	0.111	0.374	0.376	-5.671	-5.935	15.373	16.391	-0.503	3.125	17.961	21.450
	M4	0.159	0.159	0.946	0.921	-3.746	-3.957	12.243	14.061	-0.396	3.620	19.053	24.214
0.5%	M1	-0.024	-0.039	0.635	0.651	0.073	0.066	3.260	3.024	3.396	5.144	7.129	9.830
	M2	0.137	0.172	0.713	0.728	1.256	1.241	2.503	2.589	3.185	4.414	5.205	6.639
	M3	0.022	0.024	0.265	0.263	-2.937	-3.106	7.578	7.597	0.260	3.322	13.953	17.519
	M4	0.076	0.075	0.773	0.749	-1.718	-1.783	6.847	6.863	0.539	3.940	15.142	19.739
99.5%	M1	0.034	0.040	0.635	0.640	0.026	0.056	3.003	4.195	-3.618	-5.027	7.254	9.448
	M2	-0.137	-0.157	0.713	0.731	-1.256	-1.206	2.505	2.563	-3.183	-4.383	5.205	6.638
	M3	-0.928	-0.038	0.260	0.255	-4.914	2.622	7.458	7.204	-16.277	-3.509	14.076	18.273
	M4	-0.970	-0.093	0.774	0.750	-5.538	1.333	7.075	7.006	-16.728	-4.273	15.262	20.376
99.9%	M1	0.148	0.152	1.174	1.159	0.779	0.843	8.274	13.191	-6.475	-9.451	12.075	15.998
	M2	-0.312	-0.340	1.162	1.185	-3.592	-3.573	5.614	5.702	-3.954	-5.413	7.324	9.197
	M3	-1.791	-0.118	0.379	0.379	-11.258	4.758	13.832	13.803	-29.562	-3.197	17.575	22.278
	M4	-1.826	-0.171	0.940	0.914	-11.733	2.742	11.805	11.591	-30.057	-3.910	18.783	24.285

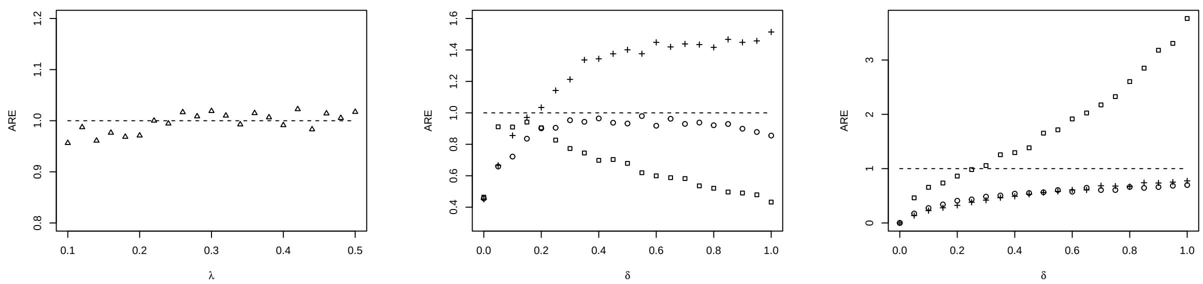

Figure S.1: The ARE($\widehat{\boldsymbol{\vartheta}}_n, \widetilde{\boldsymbol{\vartheta}}_n$) (left), ARE($\widehat{\boldsymbol{\vartheta}}_n, \check{\boldsymbol{\vartheta}}_n$) (middle) and ARE($\widehat{\boldsymbol{\vartheta}}_n, \breve{\boldsymbol{\vartheta}}_n$) (right) for the ALDAR model, where $\lambda = 0.1 + 0.02k$ and $\delta = k/20$ with $k = 0, 1, \ldots, 20$, for Tukey-lambda ($\triangle$), $N(0, 6)$ ($\square$), t_5 (+) or standard Laplace ($\circ$) distribution.

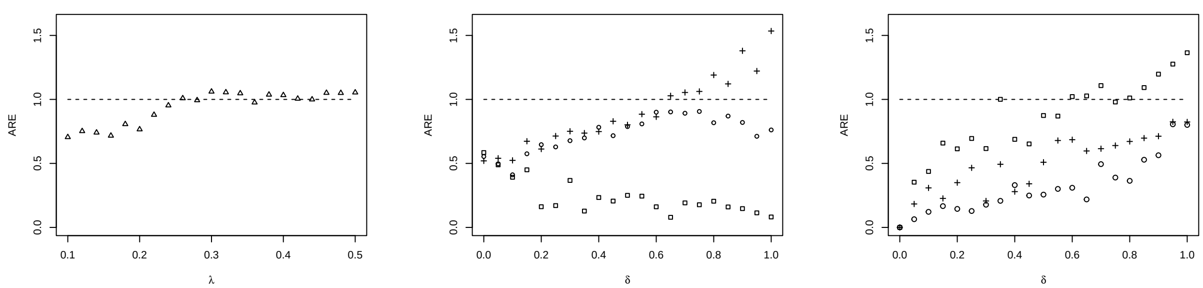

Figure S.2: The ARE($\widehat{\boldsymbol{\vartheta}}_n, \widetilde{\boldsymbol{\vartheta}}_n$) (left), ARE($\widehat{\boldsymbol{\vartheta}}_n, \check{\boldsymbol{\vartheta}}_n$) (middle) and ARE($\widehat{\boldsymbol{\vartheta}}_n, \breve{\boldsymbol{\vartheta}}_n$) (right) for the ESTAR-GARCH model, where $\lambda = 0.1+0.02k$ and $\delta = k/20$ with $k = 0, 1, \ldots, 20$, for Tukey-lambda ($\triangle$), $N(0, 6)$ ($\square$), t_5 (+) or standard Laplace ($\circ$) distribution.

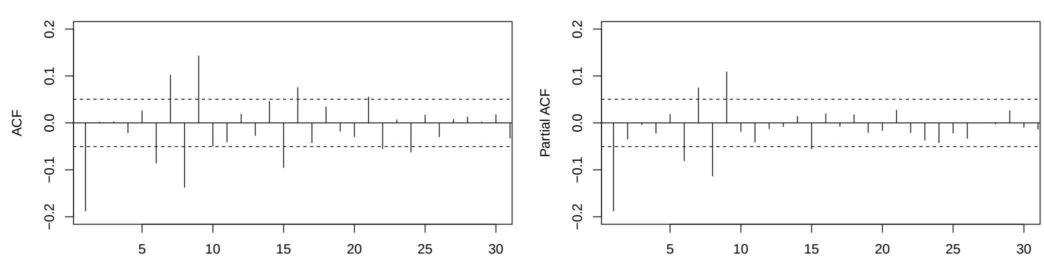

Figure S.3: The ACF and PACF plots of $\{y_t\}$, where the dashed lines are the corresponding 95% confidence bounds.

Table S.2: Biases, ASDs, and ESDs of the semi-parametric CQR estimator for DGP2 and DGP3, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. ASD_1 and ASD_2 correspond to the bandwidths h_B and h_{HS}, respectively.

	n	Bias	ASD_1	ASD_2	ESD	Bias	ASD_1	ASD_2	ESD	Bias	ASD_1	ASD_2	ESD
		$F = F_N$				$F = F_{t_5}$				$F = F_\lambda$			
DGP2													
ϕ_1	500	-0.005	0.055	0.054	0.053	-0.004	0.057	0.055	0.055	-0.006	0.079	0.077	0.067
	1000	-0.001	0.039	0.039	0.037	-0.001	0.041	0.039	0.040	-0.005	0.046	0.045	0.049
α_1^+	500	-0.007	0.085	0.084	0.090	-0.006	0.086	0.080	0.085	-0.001	0.072	0.069	0.064
	1000	-0.002	0.065	0.064	0.060	-0.004	0.060	0.056	0.058	-0.004	0.044	0.042	0.048
α_1^-	500	-0.010	0.099	0.097	0.098	-0.004	0.106	0.098	0.109	-0.005	0.116	0.111	0.093
	1000	-0.002	0.071	0.071	0.070	-0.003	0.077	0.072	0.071	-0.004	0.063	0.060	0.062
DGP3													
ϕ_1	500	-0.004	0.057	0.056	0.056	-0.008	0.059	0.057	0.062	-0.010	0.158	0.151	0.103
	1000	-0.004	0.040	0.039	0.040	-0.003	0.041	0.040	0.046	-0.014	0.141	0.135	0.096
α_1^+	500	-0.008	0.092	0.089	0.087	-0.004	0.091	0.088	0.097	0.048	0.170	0.165	0.190
	1000	-0.002	0.067	0.066	0.063	-0.006	0.066	0.062	0.067	0.036	0.188	0.181	0.169
α_2^+	500	-0.007	0.090	0.087	0.089	-0.003	0.089	0.087	0.100	0.053	0.172	0.167	0.187
	1000	-0.004	0.064	0.063	0.059	-0.003	0.065	0.061	0.068	0.042	0.175	0.167	0.177
α_1^-	500	-0.005	0.110	0.106	0.103	-0.007	0.114	0.111	0.118	0.086	0.249	0.242	0.273
	1000	-0.004	0.075	0.073	0.073	-0.005	0.076	0.072	0.089	0.068	0.216	0.207	0.265
α_1^-	500	-0.006	0.109	0.106	0.105	-0.006	0.113	0.110	0.136	0.102	0.240	0.233	0.307
	1000	-0.002	0.073	0.072	0.073	-0.005	0.076	0.072	0.089	0.063	0.214	0.205	0.252

Table S.3: Biases, ASDs, and ESDs of the parametric CQR estimator for DGP2 and DGP3, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively.

	n	Bias	ASD	ESD	Bias	ASD	ESD	Bias	ASD	ESD
		$F = F_N$			$F = F_{t_5}$			$F = F_\lambda$		
DGP2										
ϕ_1	500	-0.002	0.079	0.053	-0.001	0.090	0.056	-0.002	0.069	0.067
	1000	0.000	0.056	0.037	0.000	0.049	0.040	-0.003	0.047	0.049
α_1^+	500	-0.038	0.080	0.084	-0.041	0.075	0.075	-0.003	0.054	0.054
	1000	-0.034	0.059	0.056	-0.039	0.046	0.050	-0.004	0.038	0.042
α_1^-	500	-0.046	0.093	0.092	-0.048	0.092	0.094	-0.010	0.089	0.072
	1000	-0.037	0.065	0.066	-0.046	0.059	0.062	-0.005	0.056	0.055
DGP3										
ϕ_1	500	-0.001	0.113	0.056	-0.004	0.070	0.063	0.001	0.149	0.110
	1000	-0.003	0.057	0.040	-0.001	0.050	0.047	-0.004	0.131	0.099
α_1^+	500	-0.041	0.097	0.080	-0.047	0.066	0.076	-0.007	0.118	0.096
	1000	-0.037	0.060	0.057	-0.048	0.050	0.053	-0.002	0.118	0.110
α_2^+	500	-0.039	0.095	0.078	-0.045	0.065	0.076	0.003	0.117	0.124
	1000	-0.038	0.058	0.053	-0.046	0.049	0.053	-0.002	0.106	0.090
α_1^-	500	-0.045	0.115	0.092	-0.063	0.082	0.091	0.006	0.164	0.177
	1000	-0.046	0.066	0.066	-0.062	0.057	0.069	0.004	0.121	0.166
α_2^-	500	-0.044	0.116	0.094	-0.062	0.082	0.109	0.013	0.162	0.201
	1000	-0.042	0.066	0.065	-0.060	0.057	0.071	-0.003	0.125	0.137

Table S.4: Biases, ASDs, and ESDs of the semi-parametric CQR estimator for DGP4, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. ASD_1 and ASD_2 correspond to the bandwidths h_B and h_{HS}, respectively.

	n	Bias	ASD_1	ASD_2	ESD	Bias	ASD_1	ASD_2	ESD	Bias	ASD_1	ASD_2	ESD
		$F = F_N$				$F = F_{t_5}$				$F = F_\lambda$			
α_{00}	500	-0.038	0.062	0.061	0.102	-0.031	0.094	0.091	0.123	-0.053	0.126	0.124	0.173
	1000	-0.023	0.053	0.052	0.083	-0.012	0.052	0.050	0.070	-0.036	0.086	0.084	0.111
α_{01}	500	0.140	0.253	0.246	0.292	0.104	0.359	0.349	0.349	0.111	0.357	0.352	0.403
	1000	0.105	0.206	0.203	0.231	0.041	0.235	0.230	0.232	0.069	0.265	0.262	0.304
α_{10}	500	-0.095	0.337	0.326	0.358	-0.024	0.303	0.292	0.288	-0.024	0.341	0.334	0.332
	1000	-0.054	0.304	0.294	0.307	-0.002	0.204	0.196	0.192	-0.012	0.204	0.199	0.208
α_{11}	500	0.008	0.221	0.215	0.212	-0.013	0.088	0.086	0.115	-0.009	0.093	0.091	0.089
	1000	-0.010	0.180	0.174	0.183	-0.011	0.053	0.050	0.059	-0.007	0.076	0.075	0.069
γ	500	0.234	0.599	0.578	0.626	0.223	0.457	0.439	0.404	0.236	0.655	0.646	0.790
	1000	0.222	0.558	0.544	0.497	0.171	0.307	0.298	0.318	0.162	0.507	0.500	0.557
c	500	-0.017	0.271	0.263	0.309	-0.003	0.354	0.344	0.332	0.015	0.271	0.265	0.246
	1000	-0.009	0.246	0.239	0.246	0.016	0.164	0.158	0.191	0.022	0.165	0.161	0.165
a	500	0.080	0.109	0.104	0.155	0.425	0.191	0.187	0.212	0.313	0.152	0.151	0.190
	1000	0.071	0.082	0.078	0.106	0.405	0.160	0.154	0.168	0.316	0.107	0.105	0.126
b	500	0.037	0.057	0.055	0.083	0.027	0.059	0.057	0.052	0.020	0.053	0.053	0.055
	1000	0.020	0.047	0.045	0.060	0.024	0.036	0.033	0.035	0.010	0.037	0.036	0.039

Table S.5: Biases, ASDs, and ESDs of the parametric CQR for DGP4, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively.

	n	Bias	ASD	ESD	Bias	ASD	ESD	Bias	ASD	ESD
		$F = F_N$			$F = F_{t_5}$			$F = F_\lambda$		
α_{00}	500	-0.017	0.057	0.075	-0.021	0.080	0.093	-0.038	0.123	0.162
	1000	-0.014	0.047	0.068	-0.014	0.037	0.055	-0.029	0.087	0.105
α_{01}	500	0.083	0.303	0.251	0.050	0.313	0.275	0.135	0.440	0.409
	1000	0.062	0.240	0.211	0.006	0.173	0.191	0.076	0.284	0.303
α_{10}	500	0.010	0.244	0.235	-0.039	0.242	0.223	-0.006	0.238	0.290
	1000	-0.016	0.218	0.210	-0.038	0.158	0.141	-0.008	0.165	0.184
α_{11}	500	0.027	0.191	0.168	-0.006	0.083	0.092	-0.002	0.090	0.098
	1000	0.003	0.164	0.144	-0.013	0.079	0.073	-0.001	0.057	0.059
γ	500	0.087	0.602	0.575	0.136	0.571	0.623	0.159	0.684	0.644
	1000	0.127	0.550	0.521	0.100	0.404	0.437	0.114	0.508	0.545
c	500	0.025	0.223	0.195	-0.020	0.314	0.289	0.013	0.451	0.497
	1000	0.005	0.209	0.174	-0.012	0.216	0.198	0.004	0.287	0.301
a	500	-0.100	0.074	0.047	-0.121	0.057	0.034	-0.018	0.078	0.078
	1000	-0.098	0.053	0.039	-0.126	0.030	0.022	-0.016	0.055	0.056
b	500	-0.026	0.086	0.106	0.005	0.049	0.052	0.006	0.055	0.056
	1000	-0.006	0.046	0.052	0.011	0.034	0.040	0.002	0.039	0.039

Table S.6: Biases and RMSEs for estimating and predicting conditional quantiles at $\tau = 5\%, 10\%, 90\%$ and 95% for DGP2, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. M1, M2, M3 and M4 represent the semi-parametric CQR, parametric CQR, GQMLE and EQMLE, respectively.

		F_N				F_{t_5}				F_λ			
		Bias		RMSE		Bias		RMSE		Bias		RMSE	
τ	F	in	out	in	out	in	out	in	out	in	out	in	out
5%	M1	0.002	0.002	0.137	0.136	-0.004	-0.005	0.407	0.341	-0.024	-0.003	1.204	0.924
	M2	-0.026	-0.043	1.406	0.772	0.519	0.520	0.695	0.685	-0.026	-0.043	1.406	0.772
	M3	-0.582	-0.617	1.216	1.277	-1.341	-1.401	3.582	3.960	-5.217	-5.364	15.278	16.392
	M4	0.010	0.008	0.134	0.137	-0.006	-0.003	0.463	0.378	0.009	0.057	1.085	1.121
10%	M1	0.003	0.005	0.116	0.113	0.000	-0.003	0.283	0.255	-0.008	0.008	0.848	0.686
	M2	0.035	0.036	0.108	0.107	0.079	0.081	0.305	0.283	-0.008	-0.016	1.093	0.580
	M3	-0.457	-0.484	0.950	1.000	-0.933	-0.978	2.482	2.797	-3.611	-3.712	10.539	11.317
	M4	0.006	0.005	0.115	0.118	-0.011	-0.007	0.330	0.254	0.010	0.054	0.748	0.765
90%	M1	0.000	-0.002	0.115	0.108	0.004	-0.002	0.282	0.231	0.022	0.029	0.858	0.721
	M2	-0.035	-0.037	0.107	0.105	-0.079	-0.070	0.297	0.222	0.006	0.020	0.986	0.575
	M3	0.459	0.483	0.962	0.990	0.922	0.992	2.518	3.096	3.611	3.756	10.602	11.729
	M4	-0.005	-0.005	0.116	0.124	-0.004	0.002	0.303	0.241	-0.024	0.006	0.796	1.242
95%	M1	-0.001	-0.002	0.137	0.132	0.013	0.005	0.403	0.338	0.011	0.037	1.208	1.027
	M2	-0.188	-0.190	0.229	0.230	-0.519	-0.508	0.691	0.627	0.024	0.047	1.288	0.782
	M3	0.588	0.620	1.231	1.274	1.334	1.439	3.637	4.489	5.176	5.376	15.232	16.719
	M4	-0.007	-0.006	0.134	0.139	-0.007	0.001	0.441	0.344	-0.049	-0.025	1.085	1.565

Table S.7: Biases and RMSEs for estimating and predicting conditional quantiles at $\tau = 0.1\%, 0.5\%, 99.5\%$ and 99.9% for DGP2, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. M1, M2, M3 and M4 represent the semi-parametric CQR, parametric CQR, GQMLE and EQMLE, respectively.

		F_N				F_{t_5}				F_λ			
		Bias		RMSE		Bias		RMSE		Bias		RMSE	
τ	F	in	out	in	out	in	out	in	out	in	out	in	out
0.1%	M1	0.033	0.039	0.438	0.441	1.854	1.784	3.208	3.240	8.695	8.462	11.040	9.790
	M2	1.417	1.423	1.458	1.462	9.737	9.729	10.536	10.580	-0.837	-1.002	6.673	5.679
	M3	-0.870	-0.934	2.132	2.241	-3.251	-3.758	15.282	24.095	-13.905	-15.075	51.190	62.931
	M4	0.177	0.174	0.394	0.385	1.667	1.648	5.105	4.863	2.905	2.744	8.858	7.249
0.5%	M1	0.010	0.011	0.241	0.249	-0.099	-0.140	1.750	1.761	0.586	0.582	3.344	2.898
	M2	0.889	0.893	0.925	0.928	4.165	4.160	4.566	4.579	-0.274	-0.352	3.323	2.455
	M3	-0.860	-0.917	1.870	1.972	-3.042	-3.256	8.918	11.218	-11.164	-11.535	34.679	38.720
	M4	0.047	0.045	0.244	0.242	0.213	0.187	1.565	1.448	0.560	0.603	3.383	3.036
99.5%	M1	-0.011	-0.015	0.243	0.257	0.067	0.093	1.695	1.649	-0.532	-0.576	3.208	2.756
	M2	-0.890	-0.895	0.925	0.930	-4.165	-4.149	4.565	4.534	0.272	0.357	3.210	2.501
	M3	-0.579	0.927	1.895	1.982	-3.524	3.325	9.124	12.402	-5.034	12.245	35.770	38.335
	M4	-0.038	-0.039	0.241	0.244	-0.223	-0.191	1.682	1.504	-0.246	-0.255	3.414	3.376
99.9%	M1	-0.053	-0.058	0.448	0.456	-1.928	-1.861	3.246	3.188	-8.701	-8.490	11.127	9.858
	M2	-1.418	-1.424	1.459	1.464	-9.737	-9.717	10.536	10.537	0.835	1.006	6.583	5.732
	M3	-1.808	0.936	2.153	2.233	-11.026	3.985	16.936	24.536	-19.367	15.509	54.370	59.331
	M4	-0.170	-0.170	0.387	0.387	-1.635	-1.519	5.131	4.767	-2.498	-2.492	8.807	8.801

Table S.8: Biases and RMSEs for estimating and predicting conditional quantiles at $\tau = 5\%, 10\%, 90\%$ and 95% for DGP4, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. M1, M2, M3 and M4 represent the semi-parametric CQR, parametric CQR, GQMLE and EQMLE, respectively.

		F_N				F_{t_5}				F_λ			
		Bias		RMSE		Bias		RMSE		Bias		RMSE	
τ	F	in	out	in	out	in	out	in	out	in	out	in	out
5%	M1	-0.001	-0.004	0.116	0.129	0.142	0.142	0.285	0.308	0.083	-0.021	0.835	1.163
	M2	0.008	0.012	0.076	0.107	0.170	0.224	0.319	0.382	0.079	0.019	0.764	0.758
	M3	0.004	0.007	0.061	0.071	0.116	0.146	0.381	0.463	-0.665	-0.996	7.775	7.839
	M4	0.002	0.006	0.068	0.073	0.127	0.169	0.283	0.361	-0.159	-0.140	2.674	2.317
10%	M1	0.005	0.003	0.096	0.113	0.092	0.092	0.212	0.262	0.074	-0.010	0.750	1.005
	M2	0.002	0.005	0.064	0.098	0.098	0.131	0.216	0.257	0.076	0.015	0.633	0.681
	M3	0.002	0.005	0.054	0.064	0.076	0.092	0.300	0.394	-0.350	-0.660	6.921	6.731
	M4	0.001	0.005	0.060	0.065	0.085	0.111	0.214	0.254	-0.081	-0.047	2.167	2.272
90%	M1	-0.006	-0.004	0.097	0.112	-0.119	-0.142	0.240	0.299	-0.086	-0.008	0.815	1.038
	M2	-0.002	-0.004	0.063	0.061	-0.105	-0.140	0.229	0.290	-0.072	-0.014	0.632	0.912
	M3	-0.002	0.002	0.054	0.067	-0.090	-0.121	0.317	0.455	0.228	-0.088	6.507	5.350
	M4	-0.002	0.001	0.059	0.064	-0.098	-0.129	0.227	0.289	0.086	0.197	2.353	3.449
95%	M1	-0.003	0.000	0.117	0.125	-0.172	-0.202	0.318	0.399	-0.083	-0.020	0.824	0.992
	M2	-0.008	-0.010	0.075	0.066	-0.178	-0.234	0.334	0.412	-0.075	-0.018	0.763	0.920
	M3	-0.003	0.002	0.061	0.074	-0.135	-0.176	0.402	0.579	0.634	0.334	7.499	5.680
	M4	-0.001	0.001	0.068	0.070	-0.145	-0.191	0.303	0.394	0.167	0.249	2.005	3.368

Table S.9: Biases and RMSEs for estimating and predicting conditional quantiles at $\tau = 0.1\%, 0.5\%, 99.5\%$ and 99.9% for DGP4, where the innovations follow the standard normal, Student's t_5 or Tukey-lambda distribution with the shape parameter $\lambda = 0.1$, denoted by F_N, F_{t_5} or F_λ, respectively. M1, M2, M3 and M4 represent the semi-parametric CQR, parametric CQR, GQMLE and EQMLE, respectively.

		F_N				F_{t_5}				F_λ			
		Bias		RMSE		Bias		RMSE		Bias		RMSE	
τ	F	in	out	in	out	in	out	in	out	in	out	in	out
0.1%	M1	0.015	0.009	0.192	0.205	0.248	0.254	0.678	0.782	0.008	-0.022	1.778	1.865
	M2	0.042	0.047	0.157	0.169	0.693	0.874	1.120	1.374	0.017	-0.079	1.243	1.470
	M3	0.022	0.024	0.096	0.103	0.402	0.485	0.893	1.086	-2.292	-2.594	13.249	12.173
	M4	0.013	0.018	0.104	0.103	0.406	0.507	0.771	0.948	-0.632	-0.584	6.622	4.861
0.5%	M1	0.082	0.073	0.246	0.263	1.455	1.714	1.941	2.333	-0.054	-0.150	2.633	2.176
	M2	0.079	0.083	0.243	0.245	1.470	1.829	2.269	2.818	-0.088	-0.095	1.792	1.799
	M3	0.064	0.065	0.147	0.149	1.031	1.257	1.863	2.277	-3.585	-3.894	18.642	16.482
	M4	0.052	0.056	0.151	0.148	1.020	1.228	1.755	2.202	-1.063	-1.067	9.717	8.260
99.5%	M1	0.004	0.009	0.233	0.230	-1.434	-1.733	1.935	2.511	0.068	0.116	2.632	2.662
	M2	-0.080	-0.082	0.244	0.221	-1.477	-1.838	2.284	2.839	0.092	0.096	2.025	1.876
	M3	-0.067	-0.060	0.146	0.142	-1.118	-1.421	1.929	2.486	3.576	3.072	16.328	11.101
	M4	-0.055	-0.051	0.150	0.137	-1.090	-1.400	1.790	2.509	1.203	1.107	11.003	10.569
99.9%	M1	0.034	0.039	0.211	0.213	-0.286	-0.340	0.677	0.882	-0.008	0.056	1.777	1.775
	M2	-0.042	-0.045	0.157	0.138	-0.701	-0.884	1.136	1.396	-0.004	0.023	1.203	1.207
	M3	-0.025	-0.019	0.094	0.104	-0.460	-0.596	0.948	1.209	2.291	1.976	12.313	8.805
	M4	-0.016	-0.014	0.101	0.101	-0.469	-0.615	0.833	1.149	0.621	0.698	7.120	9.797