Supplementary material for

“Quantile Estimation of Regression Models with GARCH-X Errors”

Qianqian Zhu, Guodong Li and Zhijie Xiao

*Shanghai University of Finance and Economics,*

*University of Hong Kong and Boston College*

**Abstract**

This supplementary material provides Lemmas S.1-S.5 with proofs, and includes technical details for Theorems 1-3 and Corollaries 1-3, as well as additional simulation results for Section 3.1.

1 **Technical details**

This section provides the detailed proofs of Theorems 1-3 and Corollaries 1-3. To show Theorems 1-3, we introduce Lemmas S.1-S.5 with proofs. Specifically, Lemma S.1 contains some preliminary results. Lemma S.2 will be used to handle initial values in GARCH-X models. Lemma S.3 verifies the stochastic differentiability condition defined on Page 298 of Pollard (1985), and its proof mainly uses the bracketing method; see also Lee and Noh.
(2013) and Zhu and Ling (2011). Lemmas S.4 and S.5 will be used to verify the root-
$n$ consistency and the asymptotic normality of $\hat{\theta}_{\tau n}$ in Theorem 2, and their proofs are based
on Lemma S.3 and some approximation arguments.

Throughout this section $C$ is a generic positive constant which may take different values
at its different occurrences, $\rho \in (0, 1)$ is a generic constant which may take different values
at its different occurrences, $o_p(1)$ denotes a sequence of random variables converging to
zero in probability, and the notation $o^*_p(1)$ corresponds to the bootstrap probability space.
We denote by $\| \cdot \|$ the norm of a matrix or column vector, defined as $\|A\| = \sqrt{\text{tr}(AA^T)} =
\sqrt{\sum_{i,j} |a_{ij}|^2}$. For simplicity, denote $\psi_r(x) = \tau - I(x < 0)$ and $\varepsilon_{t,\tau} = \varepsilon_{t} - b_r$. In addition, let
$\sigma_t = \sigma_t(\lambda_0)$, $\tilde{\sigma}_t = \sigma_t(\hat{\lambda}_n^{\text{int}})$, $\hat{\theta}_t(\theta) = \rho_r[Y_t - \tilde{\theta}_t(\theta)]$ and $\ell_t(\theta) = \rho_r[Y_t - q_t(\theta)]$,
where $\lambda_0$ and $\hat{\lambda}_n^{\text{int}}$ are the true value and an appropriate estimator of $\lambda$, respectively,
$q_t(\theta) = \phi'X_{t-1} + b\sigma_t(\lambda)$ and $\tilde{q}_t(\theta) = \phi'X_{t-1} + b\tilde{\sigma}_t(\lambda)$ are the conditional quantile functions
of $Y_t$ without and with initial values, respectively.

**Lemma S.1.** Let $\xi_{\rho,t} = \sum_{j=0}^{\infty} \rho^j(1 + \|X_{t-j-1}\| + \|V_{t-j-1}\|^{1/2} + |u_{t-j}^\perp|)$ and $\zeta_{\rho,t} = \sum_{j=0}^{\infty} \rho^j(1 +
\|X_{t-j-1}\| + |u_{t-j}^\perp|)$ be positive random variables depending on a constant $\rho \in (0, 1)$, where
$\iota$ is a constant satisfying $\iota \in (0, 2/(4 + \delta))$ for some $\delta > 0$. If Assumption 1 holds, then

1. $\sup_{\theta} \sigma_t^2(\lambda) \leq C \xi_{\rho,t-1}^2$;
2. $\sup_{\theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial q_t(\theta)}{\partial \theta} \right| \leq C \zeta_{\rho,t}$ and $\sup_{\theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \right| \leq C \zeta_{\rho,t-1}^2$;
3. $\sup_{\theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial \tilde{q}_t(\theta)}{\partial \theta} \right| \leq C \zeta_{\rho,t}$ and $\sup_{\theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial^2 \tilde{q}_t(\theta)}{\partial \theta \partial \theta'} \right| \leq C \zeta_{\rho,t-1}^2$;
4. for any $\kappa > 0$, there exists a constant $c > 0$ such that

$$E \left\{ \sup_{\theta} \left[ \frac{\sigma_t^2(\lambda_1)}{\sigma_t^2(\lambda_2)} \right]^\kappa : \|\lambda_1 - \lambda_2\| \leq c \right\} < \infty.$$
Lemma S.2. Let \( \xi_p = \sum_{j=0}^{\infty} \rho^j (1 + \|X_{j-1}\| + |u_j|) \) be a positive random variable depending on a constant \( \rho \in (0, 1) \). If Assumption 1 holds, then

(i) \( \sup_{\Theta} |\hat{\sigma}_i(\lambda) - \sigma_i(\lambda)| \leq C \rho^i \xi_p \); (ii) \( \sup_{\Theta} \frac{1}{\sigma_i(\lambda)} \left\| \frac{\partial \hat{q}_t(\theta)}{\partial \theta} - \frac{\partial q_t(\theta)}{\partial \theta} \right\| \leq C \rho^i \xi_p \).

Lemma S.3. If Assumptions 1, 3 and 4 hold and \( E(u_i^2) < \infty \), then for \( u = o_p(1) \),

\[
\zeta_n(u) = o_p(\sqrt{n}||u|| + n||u||^2),
\]

where \( \zeta_n(u) = \sum_{t=1}^{n} \sigma_t^{-1} q_{it}(u) \{ \xi_{it}(u) - E[\xi_{it}(u)|\mathcal{F}_{t-1}] \} \) with

\[
q_{it}(u) = u \frac{\partial \hat{q}_t(\theta_{t0})}{\partial \theta} \quad \text{and} \quad \xi_{it}(u) = \int_{0}^{1} \left[ I(\varepsilon_t \leq b_r + \sigma_t^{-1} q_{it}(u)s) - I(\varepsilon_t \leq b_r) \right] ds.
\]

Lemma S.4. Suppose that \( \sqrt{n}(\tilde{\lambda}_n - \lambda_0) = O_p(1) \) and \( E(u_i^2) < \infty \). Under Assumptions 1, 3 and 4, for \( \theta - \theta_{t0} = o_p(1) \), it holds that

\[
n[\widehat{L}_n(\theta) - \hat{L}_n(\theta_{t0})] - n[\tilde{L}_n(\theta) - \tilde{L}_n(\theta_{t0})] = o_p(\sqrt{n}||\theta - \theta_{t0}|| + n||\theta - \theta_{t0}||^2),
\]

where \( \tilde{L}_n(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\sigma}_i^{-1} \rho_r [Y_i - q_t(\theta)] \) and \( \hat{L}_n(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\sigma}_i^{-1} \rho_r [Y_i - \hat{q}_t(\theta)] \).

Lemma S.5. Suppose that \( \sqrt{n}(\tilde{\lambda}_n - \lambda_0) = O_p(1) \) and \( E(u_i^2) < \infty \). Under Assumptions 1, 3 and 4, for \( \theta - \theta_{t0} = o_p(1) \), it holds that

\[
n[\tilde{L}_n(\theta) - \tilde{L}_n(\theta_{t0})] = -\sqrt{n}(\theta - \theta_{t0})' T_n + \sqrt{n}(\theta - \theta_{t0})' J_n \sqrt{n}(\theta - \theta_{t0}) + o_p(\sqrt{n}||\theta - \theta_{t0}|| + n||\theta - \theta_{t0}||^2),
\]

where \( \tilde{L}_n(\theta) = n^{-1} \sum_{i=1}^{n} \hat{\sigma}_i^{-1} \rho_r [Y_i - q_t(\theta)] \),

\[
T_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \psi_r(\varepsilon_{i\tau}) \quad \text{and} \quad J_n = \frac{f(\varepsilon_{r})}{2n} \sum_{i=1}^{n} \frac{1}{\sigma_t^2} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \frac{\partial q_t(\theta_{t0})}{\partial \theta}. 
\]
Proof of Lemma S.1. Denote \( \alpha(B) = \sum_{i=1}^{q} \alpha_i B^i \) and \( \beta(B) = 1 - \sum_{i=1}^{p} \beta_i B^i \), where \( B \) is the back-shift operator. By Assumption 1, it holds that

\[
\beta^{-1}(B) \alpha(B) = \sum_{i=1}^{\infty} a_{\gamma}(i) B^i \quad \text{and} \quad \beta^{-1}(B) = \sum_{i=0}^{\infty} a_{\beta}(i) B^i,
\]

where \( a_{\beta}(i) = e^T G^i e \) and \( a_{\gamma}(i) = \sum_{j=1}^{q} \alpha_j a_{\beta}(i-j) \). \( e = (1, 0, \ldots, 0)' \) is \( p \times 1 \) vector and \( p \times p \) matrix \( G \) is defined as below

\[
G = \begin{pmatrix}
\beta_1 & \cdots & \beta_p \\
I_{p-1} & 0
\end{pmatrix},
\]

with \( I_m \) being the \( m \times m \) identity matrix and \( 0 \) being the zero vector with compatible dimensions. By Lemma 3.1 of Berkes, Horváth, and Kokoszka (2003), we have

\[
\sup_{\Theta} a_{\gamma}(i) \leq C \rho^i \quad \text{and} \quad \sup_{\Theta} a_{\beta}(i) \leq C \rho^i
\]

(S.1)

for a constant \( 0 < C < \infty \) and a constant \( \rho \in (0, 1) \). Moreover, for \( \theta \in \Theta \) and any constant vector \( c \) with all elements being nonnegative, it holds that \( c^T e e^T c \leq c^T G \theta c \), and hence

\[
a_{\beta}(i+k) = e^T G^{i+k} e \geq e^T G^i e = a_{\beta}(i) \omega^k.
\]

This implies that \( \sup_{\theta} a_{\beta}(i)/a_{\beta}(i+k) \leq \omega^{-k} \).

Note that \( a_{\gamma}(i) \geq \alpha_1 a_{\beta}(i-1) \). As a result, for any integer \( i \) and \( k \leq \max(p, q) \), it can be verified that

\[
\sup_{\Theta} \frac{a_{\beta}(i)}{a_{\gamma}(i+k)} \leq \omega^{-k} \quad \text{and} \quad \sup_{\Theta} \frac{a_{\gamma}(i)}{a_{\gamma}(i+k)} \leq \omega^{-j+k} \sum_{j=1}^{q} \omega^{-j-k}.
\]

(S.2)

We first prove (i). Since \( \sigma_t^2(\lambda) = 1 + \sum_{i=1}^{q} \alpha_i u_{t-i}(\phi) + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2(\lambda) + \pi^T V_{t-1} \), then we have

\[
\sigma_t^2(\lambda) = \beta^{-1}(B) \left( 1 + \pi^T V_{t-1} \right) + \beta^{-1}(B) \alpha(B) u_t^2(\phi)
\]

\[
= \frac{1}{1 - \sum_{i=1}^{p} \beta_i} + \sum_{i=0}^{\infty} \sum_{k=1}^{d} \pi_k a_{\beta}(i) v_{k,t-i-1}^2 + \sum_{i=1}^{\infty} a_{\gamma}(i) u_{t-i}(\phi),
\]

(S.3)
where \( u_t(\phi) = u_t - (\phi - \phi_0)'X_{t-1} \). It follows that

\[
\sup_{\Theta} \sigma_t^2(\lambda) \leq \frac{1}{1 - \sum_{i=1}^{p} \beta_i} + \sum_{i=0}^{d} \pi_k \sup_{\Theta} a_{\beta}(i) v_{k,t-i-1}^2 + \sum_{i=1}^{\infty} \sup_{\Theta} a_{\gamma}(i) u_{t-1}^2(\phi) \\
\leq C \sum_{i=0}^{\infty} \rho_i [1 + \sup_{\Theta} u_{t-1}^2(\phi) + \sum_{k=1}^{d} \pi_k v_{k,t-i-1}^2] \leq C \xi_{\phi,t-1}^2.
\]

Hence, (i) is asserted.

We then prove (ii). Since \( q_t(\theta) = \phi'X_{t-1} + b\sigma_t(\lambda) \), it holds that

\[
\frac{\partial q_t(\theta)}{\partial \theta} = \left( \sigma_t(\lambda), \frac{b}{2\sigma_t(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \gamma}, X_{t-1} + \frac{b}{2\sigma_t(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \phi} \right)', \tag{S.4}
\]

Moreover, it can be verified that

\[
\frac{\partial \sigma_t^2(\lambda)}{\partial \gamma} = \beta^{-1}(B) z_t(\lambda) \quad \text{and} \quad \frac{\partial \sigma_t^2(\lambda)}{\partial \phi} = -2\beta^{-1}(B) \alpha(B) u_t(\phi) X_{t-1}, \tag{S.5}
\]

where \( z_t(\lambda) = (u_{t-1}^2(\phi), \ldots, u_{t-q}^2(\phi), \sigma_{t-1}^2(\lambda), \ldots, \sigma_{t-p}^2(\lambda), v_{1,t-1}, \ldots, v_{d,t-1})' \). From (S.3), it holds that \( \sigma_t^2(\lambda) \geq a_{\gamma}(i) u_{t-i}(\phi) \) and \( \sigma_t^2(\lambda) \geq \pi_k a_{\beta}(i) v_{k,t-i-1}^2 \). This together with (S.1), (S.2) and \( \sigma_t^2(\lambda) \geq 1 \), implies that

\[
\sup_{\Theta} \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \gamma} \right\| = \sup_{\Theta} \left\| \sum_{i=0}^{\infty} \frac{a_{\beta}(i) z_{t-i}(\lambda)}{\sigma_t^2(\lambda)} \right\| \\
\leq \sup_{\Theta} \sum_{i=0}^{\infty} \left[ \sum_{k=1}^{d} \frac{a_{\beta}(i) v_{k,t-i-1}^2}{\sigma_t^2(\lambda)} + \sum_{k=1}^{q} \frac{a_{\beta}(i) u_{t-i-k}(\phi)}{\sigma_t^2(\lambda)} + \sum_{k=1}^{p} \frac{a_{\beta}(i) a_{\gamma}(k)}{\sigma_t^2(\lambda)} \right] \\
\leq \sum_{k=1}^{d} \frac{1}{\pi_k} + \sum_{k=1}^{q} \sum_{i=0}^{\infty} \left[ \sup_{\Theta} \frac{a_{\beta}(i)}{a_{\gamma}(i+k)} \right]^{1-i/2} \sup_{\Theta} a_{\beta}^{i/2}(i) |u_{t-i-k}(\phi)|^{1-i/2} \\
+ \sum_{k=1}^{p} \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \left[ \sup_{\Theta} \frac{a_{\beta}(i) a_{\gamma}(j)}{a_{\gamma}(i+j+k)} \right]^{1-i/2} \sup_{\Theta} [a_{\beta}(i) a_{\gamma}(j)]^{1/2} |u_{t-i-j-k}(\phi)|^{1-i/2} \leq C \sum_{i=0}^{\infty} \rho_i [1 + \|X_{t-i-1}\|^t + |u_{t-i}|^t], \tag{S.6}
\]
and
\[
\sup_\Theta \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \phi} \right\| \leq 2 \sup_\Theta \left\| \frac{\sum_{i=1}^\infty a_\gamma(i) u_{i-1}(\phi) X_{t-i-1}}{\sigma_t(\lambda)} \right\| \leq 2 \sum_{i=1}^\infty \sup_\Theta \sqrt{a_\gamma} \| X_{t-i-1} \| \leq C \sum_{i=1}^\infty \rho^i \| X_{t-i-1} \|. \quad (S.7)
\]

In view of (S.4)-(S.7), we have
\[
\sup_\Theta \left\| \frac{1}{\sigma_t(\lambda)} \frac{\partial q_t(\theta)}{\partial \theta} \right\| \leq \frac{\gamma}{2} \sup_\Theta \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \gamma} \right\| + \| X_{t-1} \| + \frac{\gamma}{2} \sup_\Theta \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \phi} \right\| \leq C \sum_{i=0}^\infty \rho^i [1 + \| X_{t-i-1} \| + |u_{t-i}|] \leq C \zeta_{\rho,t}.
\]

We next consider the second derivatives. It holds that
\[
\frac{\partial^2 q_t(\theta)}{\partial b^2} = 0, \quad \frac{\partial^2 q_t(\theta)}{\partial \gamma \partial \gamma'} = 0, \quad \frac{\partial^2 q_t(\theta)}{\partial \gamma \partial \phi} = 0, \quad \frac{\partial^2 q_t(\theta)}{\partial \phi \partial \phi'} = 0.
\]

Moreover, it can be verified that
\[
\frac{\partial^2 \sigma_t^2(\lambda)}{\partial \gamma \partial \gamma'} = \beta^{-1}(\lambda) \frac{\partial z_t(\lambda)}{\partial \gamma'}, \quad \frac{\partial^2 \sigma_t^2(\lambda)}{\partial \gamma \partial \phi} = \beta^{-1}(\lambda) \frac{\partial z_t(\lambda)}{\partial \phi'}, \quad \frac{\partial^2 \sigma_t^2(\lambda)}{\partial \phi \partial \phi'} = 2 \beta^{-1}(\lambda) \alpha(\lambda) X_{t-1} X_{t-1}',
\]

where
\[
\frac{\partial z_t(\lambda)}{\partial \gamma} = \left( \frac{\partial^2 \sigma_{t-1}(\lambda)}{\partial \gamma}, \ldots, \frac{\partial^2 \sigma_{t-p}(\lambda)}{\partial \gamma}, 0_{(p+q+d) \times d} \right) \text{ and }
\frac{\partial z_t(\lambda)}{\partial \phi} = \left( -2u_{t-1}(\phi) X_{t-2}, \ldots, -2u_{t-q}(\phi) X_{t-q-1}, \frac{\partial \sigma_{t-1}(\lambda)}{\partial \phi}, \ldots, \frac{\partial \sigma_{t-p}(\lambda)}{\partial \phi}, 0_{d \times d} \right).
\]

Similar to the proof in (S.6) and (S.7), it can be verified that
\[
\sup_\Theta \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial^2 \sigma_t^2(\lambda)}{\partial \gamma \partial \gamma'} \right\| \leq \sum_{i=0}^\infty \sup_\Theta a_\beta(i) \sum_{k=1}^p \sup_\Theta \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_{t-i-k}(\lambda)}{\partial \gamma} \right\| \leq C \zeta_{\rho,t-1},
\]
In this paper, we use the setting for initial value of \( u_t(\phi) \) for \( t > 0 \) and \( \tilde{u}_t(\phi) = 0 \) for \( t \leq 0 \). The proof of this lemma.

Proof of Lemma S.2. In this paper, we use the setting for initial value of \( u_t(\phi) \) as follows

\[
\tilde{u}_t(\phi) = u_t(\phi) \quad \text{for} \quad t > 0; \quad \tilde{u}_t(\phi) = 0 \quad \text{for} \quad t \leq 0.
\]
We first prove (i). Note that $\tilde{\sigma}_t(\lambda) - \sigma_t(\lambda) = [\tilde{\sigma}_t^2(\lambda) - \sigma_t^2(\lambda)]/[\tilde{\sigma}_t(\lambda) + \sigma_t(\lambda)]$. Moreover, by (S.3), it follows that
\[
\tilde{\sigma}_t^2(\lambda) - \sigma_t^2(\lambda) = \sum_{i=1}^{\infty} a_\gamma(i)[\tilde{u}_{t-i}(\phi) - u_{t-i}(\phi)] = - \sum_{i=1}^{\infty} a_\gamma(i)u_{t-i}^2(\phi).
\]
Then by (S.1), (S.8) and $\sigma_t^2(\lambda) \geq a_\gamma(i)u_{t-i}^2(\phi)$ implied by (S.3), together with the fact that $u_t(\phi) = u_t - (\phi - \phi_0)'X_{t-1}$, we have
\[
\sup_{\theta} |\tilde{\sigma}_t(\lambda) - \sigma_t(\lambda)| \leq \sup_{\theta} \sum_{i=1}^{\infty} \frac{a_\gamma(i)u_{t-i}^2(\phi)}{\sigma_t(\lambda)} \leq \sup_{\theta} \sqrt{\sum_{i=1}^{\infty} a_\gamma(i)|u_{t-i}(\phi)|} \leq C\rho^j\xi_\rho.
\]
Hence, (i) holds.

We next verify (ii). By (S.4) and the fact that $\sigma_t^2(\lambda) \geq 1$, it can be verified that
\[
\frac{1}{\sigma_t(\lambda)} \left\| \frac{\partial q_i(\theta)}{\partial \theta} - \frac{\partial q_i(\theta)}{\partial \theta} \right\| \leq |\tilde{\sigma}_t(\lambda) - \sigma_t(\lambda)| + \frac{b}{\sigma_t(\lambda)} \left\| \frac{\partial \tilde{\sigma}_t^2(\lambda)}{\partial \gamma} - \frac{\partial \sigma_t^2(\lambda)}{\partial \gamma} \right\| + \frac{b}{\sigma_t(\lambda)} \left\| \frac{\partial \tilde{\sigma}_t^2(\lambda)}{\partial \phi} - \frac{\partial \sigma_t^2(\lambda)}{\partial \phi} \right\|.
\]
By (S.3), it holds that
\[
\left\| \frac{\tilde{z}_{t-i}(\lambda)}{\sigma_t(\lambda)} - \frac{z_{t-i}(\lambda)}{\sigma_t(\lambda)} \right\| \leq \sum_{k=1}^{q} \left| \frac{\tilde{u}_{t-i-k}(\phi)}{\tilde{\sigma}_t(\lambda)} - \frac{u_{t-i-k}(\phi)}{\sigma_t(\lambda)} \right| + \sum_{k=1}^{p} \left| \frac{\tilde{\sigma}_{t-i-k}(\lambda)}{\tilde{\sigma}_t(\lambda)} - \frac{\sigma_{t-i-k}(\lambda)}{\sigma_t(\lambda)} \right| + \sum_{k=1}^{d} \frac{u_{k,t-i}^2(\phi)}{\tilde{\sigma}_t(\lambda)} \leq \sum_{k=1}^{q} \left| \frac{\tilde{u}_{t-i-k}(\phi)}{\tilde{\sigma}_t(\lambda)} - \frac{u_{t-i-k}(\phi)}{\sigma_t(\lambda)} \right| + \sum_{k=1}^{p} \sum_{j=1}^{\infty} a_\gamma(j) \left| \frac{\tilde{u}_{t-i-j-k}(\phi)}{\tilde{\sigma}_t(\lambda)} - \frac{u_{t-i-j-k}(\phi)}{\sigma_t(\lambda)} \right| + \left| \tilde{\sigma}_t(\lambda) - \sigma_t(\lambda) \right| \sum_{k=1}^{d} \frac{u_{k,t-i}^2(\phi)}{\tilde{\sigma}_t(\lambda)} \sigma_t(\lambda).
\]
This together with (S.2), (S.5), (S.8) and the fact that $\tilde{\sigma}_t^2(\lambda) \geq 1$, then similar to the proof
In view of (S.9)-(S.11), together with (i) and $\theta \leq \bar{b}$ for $\theta \in \Theta$, it follows that

$$\sup_{\Theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial \hat{q}_t(\lambda)}{\partial \theta} - \frac{1}{\sigma_t(\lambda)} \frac{\partial q_t(\theta)}{\partial \theta} \right| \leq C \rho^t \xi_p \xi_{\rho,t}. \quad \square$$

Therefore, (ii) is asserted. The proof of this lemma is complete.

**Proof of Lemma S.3.** It can be verified that

$$|\zeta_n(u)| \leq \sqrt{n} \|u\| \sum_{j=1}^{m+p+q+d+1} \left| \frac{1}{\sqrt{n}} \sum_{l=1}^{n} m_{t,j} \{ \xi_{l,t}(u) - E[\xi_{l,t}(u)|F_{l-1}] \} \right|,$$

where $m_{t,j} = \sigma_t^{-1} \hat{q}_t(\theta_{t0})/\partial \theta_{(j)}$ with $\theta_{(j)}$ being the $j$th element of $\theta$. For $1 \leq j \leq m + p + \ldots$
Furthermore, by the Taylor expansion, it holds that
\[
D_n(u) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{f_t(u) - E[f_t(u)|F_{t-1}]\}.
\]

To establish Lemma S.3, it suffices to show that, for any \( \delta > 0 \),
\[
\sup_{|u| \leq \delta} \frac{|D_n(u)|}{1 + \sqrt{n}\|u\|} = o_p(1).
\] (S.12)

We follow the method in Lemma 4 of Pollard (1985) to verify (S.12). Let \( \mathfrak{F} = \{f_t(u) : \|u\| \leq \delta\} \) be a collection of functions indexed by \( u \). First, we verify that \( \mathfrak{F} \) satisfies the bracketing condition defined in Pollard (1985), page 304. Let \( B_r(\xi) \) be an open neighborhood of \( \xi \) with radius \( r > 0 \), and define a constant \( C_0 \) to be selected later. For any \( \epsilon > 0 \) and \( 0 < r \leq \delta \), there exists a sequence of small cubes \( \{B_{\epsilon r/C_0}(u_i)\}_{i=1}^{K(\epsilon)} \) to cover \( B_r(0) \), where \( K(\epsilon) \) is an integer less than \( C \epsilon^{-(m+p+q+d+1)} \), and the constant \( C \) is not depending on \( \epsilon \) and \( r \); see Huber (1967), page 227. Denote \( V_i(r) = B_{\epsilon r/C_0}(u_i) \cap B_r(0) \), and let \( U_i(r) = V_i(r) \) and \( U_i(r) = V_i(r) - \bigcup_{j=1}^{i-1} V_j(r) \) for \( i \geq 2 \). Note that \( \{U_i(r)\}_{i=1}^{K(\epsilon)} \) is a partition of \( B_r(0) \). For each \( u_i \in U_i(r) \) with \( 1 \leq i \leq K(\epsilon) \), define the following bracketing functions
\[
\begin{align*}
    f^L_t(u_i) &= g_t \int_0^1 \left[ I \left( \varepsilon_t \leq b_r + \frac{u_i \cdot \hat{q}_t(\theta_{t0})}{\sigma_t} s - \frac{er}{C_0} \left\| \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \right\| \right) - I(\varepsilon_t \leq b_r) \right] ds, \\
    f^U_t(u_i) &= g_t \int_0^1 \left[ I \left( \varepsilon_t \leq b_r + \frac{u_i \cdot \hat{q}_t(\theta_{t0})}{\sigma_t} s + \frac{er}{C_0} \left\| \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \right\| \right) - I(\varepsilon_t \leq b_r) \right] ds.
\end{align*}
\]

Since the indicator function \( I(\cdot) \) is non-decreasing and \( g_t \geq 0 \), for any \( u \in U_i(r) \), we have
\[
f^L_t(u_i) \leq f_t(u) \leq f^U_t(u_i). \tag{S.13}
\]

Furthermore, by the Taylor expansion, it holds that
\[
E[f^U_t(u_i) - f^L_t(u_i)|F_{t-1}] \leq \frac{er}{C_0} \cdot 2 \sup_{x \in \mathbb{R}} f_\varepsilon(x) \left\| \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \right\|^2. \tag{S.14}
\]

Denote \( \Delta_t = 2 \sup_{x \in \mathbb{R}} f_\varepsilon(x) \left\| \sigma_t^{-1} \hat{q}_t(\theta_{t0})/\hat{\theta} \right\|^2. \) By Assumption 4, we have \( \sup_{x \in \mathbb{R}} f_\varepsilon(x) < \)
This together with Lemma S.1, implies that $E(\Delta_t)$ exists. Let $C_0 = E(\Delta_t)$. Then by the iterated-expectation, it follows that

$$E \left[ f_t^U(u) - f_t^L(u) \right] = E \{ E \left[ f_t^U(\theta) - f_t^L(\theta) \big| \mathcal{F}_{t-1} \right] \} \leq \epsilon r.$$

This together with (S.13), implies that the family $\mathcal{F}$ satisfies the bracketing condition.

Put $r_k = 2^{-k} \delta$. Let $B(k) = B_{r_k}(0)$ and $A(k)$ be the annulus $B(k) \setminus B(k+1)$. From the bracketing condition, for fixed $\epsilon > 0$, there is a partition $U_1(r_k), U_2(r_k), \ldots, U_{K(\epsilon)}(r_k)$ of $B(k)$. First, consider the upper tail case. For $u \in U_i(r_k)$, by (S.14), it holds that

$$D_n(u) \leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{ f_t^U(u) - E \left[ f_t^U(u) \big| \mathcal{F}_{t-1} \right] \} + \frac{1}{\sqrt{n}} \sum_{t=1}^{n} E \left[ f_t^U(u) - f_t^L(u) \big| \mathcal{F}_{t-1} \right]$$

$$\leq D_n^U(u) + \sqrt{\epsilon r_k} \frac{1}{nC_0} \sum_{t=1}^{n} \Delta_t, \tag{S.15}$$

where

$$D_n^U(u) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{ f_t^U(u) - E \left[ f_t^U(u) \big| \mathcal{F}_{t-1} \right] \}.$$ 

Define the event

$$E_n = \left\{ \omega : \frac{1}{nC_0} \sum_{t=1}^{n} \Delta_t(\omega) < 2 \right\}.$$ 

For $u \in A(k)$, $1 + \sqrt{n}||u|| > \sqrt{n}r_{k+1} = \sqrt{n}r_k/2$. Then by (S.15) and the Chebyshev’s inequality, we have

$$P \left( \sup_{u \in A(k)} \frac{D_n(u)}{1 + \sqrt{n}||u||} > 6\epsilon, E_n \right)$$

$$\leq P \left( \max_{1 \leq i \leq K(\epsilon)} \sup_{u \in U_i(r_k) \cap A(k)} D_n(u) > 3\sqrt{\epsilon r_k}, E_n \right)$$

$$\leq K(\epsilon) \max_{1 \leq i \leq K(\epsilon)} P \left( D_n^U(u_i) > \sqrt{\epsilon r_k} \right)$$

$$\leq K(\epsilon) \max_{1 \leq i \leq K(\epsilon)} \frac{E[|D_n^U(u_i)|^2]}{n\epsilon^2 r_k^2}. \tag{S.16}$$
Moreover, by the iterated-expectation, the Taylor expansion, Assumption 4 and $\|u_i\| \leq r_k$ for $u_i \in U_i(r_k)$, we have

$$E \{[f_i^U(u_i)]^2 \} = E \{ E \{[f_i^U(u_i)]^2 | F_{t-1} \} \}$$
$$\leq 2E \left\{ g_t^2 \left[ \int_0^t \left[ F_x \left( b_r + \frac{u_i^t \frac{\partial q_t(\theta_{r,0})}{\partial \theta}}{\sigma_t} s + \frac{c_p}{C_0} \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{r,0})}{\partial \theta} \right) - F_x(b_r) \right] ds \right] \right\}$$
$$\leq C \sup_{x \in \mathbb{R}} f_x(x) r_k E \left[ \left\| \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{r,0})}{\partial \theta} \right\|^3 \right].$$

This, together with Lemma S.1, $E(\|X_i\|^{4+\delta}) < \infty$ for some $\delta > 0$ by Assumption 3, $\sup_{x \in \mathbb{R}} f_x(x) < \infty$ by Assumption 4 and the fact that $f_i^U(u_i) - E[f_i^U(u_i)|F_{t-1}]$ is a martingale difference sequence, implies that

$$E[[D_n^U(u_i)]^2] = \frac{1}{n} \sum_{i=1}^n E\{[f_i^U(u_i) - E[f_i^U(u_i)|F_{t-1}]]^2\}$$
$$\leq \frac{1}{n} \sum_{i=1}^n E[[f_i^U(u_i)]^2]$$
$$\leq \frac{C r_k}{n} \sum_{i=1}^n E \left[ \left\| \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{r,0})}{\partial \theta} \right\|^3 \right] := \Delta(r_k). \quad (S.17)$$

Combining (S.16) and (S.17), we have

$$\mathbb{P} \left( \sup_{u \in A(k)} \frac{D_n(u)}{1 + \sqrt{n} \|u\|} > 6\epsilon, E_n \right) \leq \frac{K(\epsilon) \Delta(r_k)}{n \epsilon^2 r_k^2}. \quad \text{for } \epsilon > 0, \quad \text{for } k \geq k_\epsilon.$$ 

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

$$\mathbb{P} \left( \sup_{u \in A(k)} \frac{|D_n(u)|}{1 + \sqrt{n} \|u\|} > 6\epsilon, E_n \right) \leq \frac{2K(\epsilon) \Delta(r_k)}{n \epsilon^2 r_k^2}. \quad (S.18)$$

Note that $\Delta(r_k) \to 0$ as $k \to \infty$, we can choose $k_\epsilon$ such that $2K(\epsilon) \Delta(r_k)/(\epsilon^2 \delta^2) < \epsilon$ for $k \geq k_\epsilon$. Let $k_n$ be the integer such that $n^{-1/2} \delta \leq r_{k_n} \leq 2n^{-1/2} \delta$, and split $B_\delta(0)$ into two events $B := B(k_n + 1)$ and $B^c := B(0) - B(k_n + 1)$. Note that $B^c = \bigcup_{k=0}^{k_n} A(k)$. Moreover,
by Lemma S.1, it follows that $\Delta(r_k)$ is bounded. This together with (S.18), implies that

$$
P\left( \sup_{u \in B} \frac{|D_n(u)|}{1 + \sqrt{n}||u||} > 6\epsilon \right) \leq \sum_{k=0}^{k_n} P\left( \sup_{u \in B} \frac{|D_n(u)|}{1 + \sqrt{n}||u||} > 6\epsilon, E_n \right) + P(E_n^c)
$$

$$
\leq \frac{1}{n} \sum_{k=0}^{k_n-1} C K(\epsilon) 2^{2k} + \frac{1}{n} \sum_{k=k_n}^{k_n-1} 2^{2k} + P(E_n^c)
$$

$$
\leq O\left( \frac{1}{n} \right) + 4\epsilon + P(E_n^c).
$$

Furthermore, for $u \in B$, we have $1 + \sqrt{n}||u|| \geq 1$ and $r_{k_n+1} \leq n^{-1/2} \delta < n^{-1/2}$. Similar to the proof of (S.16) and (S.17), we can show that

$$
P\left( \sup_{u \in B} \frac{D_n(u)}{1 + \sqrt{n}||u||} > 3\epsilon, E_n \right) \leq P\left( \max_{1 \leq t \leq K(\epsilon)} D_n^t(u_t) > \epsilon, E_n \right) \leq \frac{K(\epsilon) \Delta(r_{k_n+1})}{\epsilon^2}.
$$

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

$$
P\left( \sup_{u \in B} \frac{|D_n(u)|}{1 + \sqrt{n}||u||} > 3\epsilon \right) = P\left( \sup_{u \in B} \frac{|D_n(u)|}{1 + \sqrt{n}||u||} > 3\epsilon, E_n \right) + P(E_n^c)
$$

$$
\leq \frac{2K(\epsilon) \Delta(r_{k_n+1})}{\epsilon^2} + P(E_n^c).
$$

Note that $\Delta(r_{k_n+1}) \to 0$ as $n \to \infty$. Moreover, by the ergodic theorem, $P(E_n) \to 1$ as $n \to \infty$ and thus $P(E_n^c) \to 0$ as $n \to \infty$. (S.20) together with (S.19) asserts (S.12). The proof of this lemma is complete. \(\square\)

**Proof of Lemma S.4.** Denote $u = \theta - \theta_{\tau_0}$. Recall that $\hat{L}_n(\theta) = n^{-1} \sum_{t=1}^n \hat{\sigma}_t^{-1} \hat{\ell}_t(\theta)$ and $\tilde{L}_n(\theta) = n^{-1} \sum_{t=1}^n \sigma_t^{-1} \ell_t(\theta)$, where $\hat{\sigma}_t = \hat{\sigma}_t(\hat{\lambda}_n^{int})$, $\tilde{\sigma}_t = \sigma_t(\lambda_n^{int})$, $\hat{\ell}_t(\theta) = \rho_t[Y_t - \hat{q}_t(\theta)]$ and $\ell_t(\theta) = \rho_t[Y_t - q_t(\theta)]$. Then it can be verified that

$$
n[\hat{L}_n(\theta) - \hat{L}_n(\theta_{\tau_0})] - n[\tilde{L}_n(\theta) - \tilde{L}_n(\theta_{\tau_0})]
$$

$$
= \sum_{t=1}^n \hat{\sigma}_t^{-1} [\hat{\ell}_t(\theta) - \tilde{\ell}_t(\theta_{\tau_0})] - \sum_{t=1}^n \sigma_t^{-1} [\ell_t(\theta) - \ell_t(\theta_{\tau_0})]
$$

$$
= R_{3n}(\theta) + R_{4n}(\theta),
$$

(S.21)
Thus, (S.22) holds. Next, we verify that
\[
\tilde{\mathcal{R}}_{3n}(\theta) = \sum_{t=1}^{n} (\hat{\sigma}_t^{-1} - \tilde{\sigma}_t^{-1})(\tilde{\ell}_t(\theta) - \tilde{\ell}_t(\theta_0)) \quad \text{and} \\
\tilde{\mathcal{R}}_{4n}(\theta) = \sum_{t=1}^{n} \hat{\sigma}_t^{-1}[(\tilde{\ell}_t(\theta) - \tilde{\ell}_t(\theta_0)] - [\ell_t(\theta) - \ell_t(\theta_0)].
\]

First, we show that
\[
\tilde{\mathcal{R}}_{3n}(\theta) = o_p(\sqrt{n}\|u\|).
\]  \hspace{1cm} \text{(S.22)}

By the Taylor expansion and the Lipschitz continuity of $\rho_r(x)$, we have
\[
\left|\tilde{\ell}_t(\theta) - \tilde{\ell}_t(\theta_0)\right| \leq C|\tilde{q}_t(\theta) - \tilde{q}_t(\theta_0)| \leq C\|u\| \left|\frac{\partial \tilde{q}_t(\theta^*)}{\partial \theta} \right|,
\] \hspace{1cm} \text{(S.23)}

where $\theta^*$ is between $\theta_0$ and $\theta$. Recall that $\zeta_{\rho,t} = \sum_{j=0}^{\infty} \rho^j(1 + \|X_{t-j-1}\| + |u_{t-j}|)$. By Assumptions 1 and 3, it holds that $E(\zeta_{\rho,t}^{2+\delta}) < \infty$. Combing (S.63) and (S.23), by Lemma S.1, $\tilde{\sigma}_t^2(\lambda) \leq \sigma_t^2(\lambda)$ and $\tilde{\sigma}_t^2(\lambda) \geq 1$, we can show that
\[
\sup_{\theta} \frac{\left|\tilde{\mathcal{R}}_{3n}(\theta)\right|}{\sqrt{n}\|u\|} \leq C \frac{\zeta_{\rho,t}}{\sqrt{n}} \sum_{t=1}^{n} \rho^j \sup_{\theta} \left| \frac{1}{\tilde{\sigma}_t(\lambda)} \frac{\partial \tilde{q}_t(\theta^*)}{\partial \theta} \right| \frac{\tilde{\sigma}_t(\lambda)}{\tilde{\sigma}_t} \leq C \frac{\zeta_{\rho,t}}{\sqrt{n}} \sum_{t=1}^{n} \rho^j \frac{\zeta_{\rho,t}}{\tilde{\sigma}_t} \frac{\tilde{\sigma}_t(\lambda)}{\tilde{\sigma}_t} = o_p(1).
\]

Thus, (S.22) holds. Next, we verify that
\[
\tilde{\mathcal{R}}_{4n}(\theta) = o_p(\sqrt{n}\|u\| + n\|u\|^2).
\] \hspace{1cm} \text{(S.24)}

Denote $\nu_t(u) = q_t(\theta) - q_t(\theta_0)$ and $\tilde{\nu}_t(u) = \tilde{q}_t(\theta) - \tilde{q}_t(\theta_0)$. Define
\[
\xi_t(u) = \int_{0}^{1} \left[ I(\varepsilon_t \leq b_r + \sigma_t^{-1} \nu_t(u)s) - I(\varepsilon_t \leq b_r) \right] ds \quad \text{and} \\
\tilde{\xi}_t(u) = \int_{0}^{1} \left[ I(\varepsilon_t \leq b_r \tilde{\sigma}_t + \sigma_t^{-1} \tilde{\nu}_t(u)s) - I(\varepsilon_t \leq b_r) \right] ds.
\]

By the Knight equation (S.66), it can be verified that
\[
\tilde{\mathcal{R}}_{4n}(\theta) = \sum_{t=1}^{n} \hat{\sigma}_t^{-1} \left\{ \tilde{\nu}_t(u)[-\psi_r(\varepsilon_t - b_r \tilde{\sigma}_t^{-1}) + \tilde{\xi}_t(u)] - \nu_t(u)[-\psi_r(\varepsilon_t) + \xi_t(u)] \right\} \\
= \tilde{\Pi}_1(u) + \tilde{\Pi}_2(u) + \tilde{\Pi}_3(u) + \tilde{\Pi}_4(u),
\] \hspace{1cm} \text{(S.25)}
where \( \varepsilon_{t,\tau} = \varepsilon_t - b_{\tau} \),

\[
\hat{\Pi}_1(u) = -\sum_{t=1}^{n} \hat{\sigma}_t^{-1} [\hat{\nu}_t(u) - \nu_t(u)] \psi_t(\varepsilon_t - b_{\tau} \hat{\sigma}_t \sigma_t^{-1}),
\]

\[
\hat{\Pi}_2(u) = -\sum_{t=1}^{n} \hat{\sigma}_t^{-1} \nu_t(u) [\psi_t(\varepsilon_t - b_{\tau} \hat{\sigma}_t \sigma_t^{-1}) - \psi_t(\varepsilon_{t,\tau})],
\]

\[
\hat{\Pi}_3(u) = \sum_{t=1}^{n} \hat{\sigma}_t^{-1} [\hat{\nu}_t(u) - \nu_t(u)] \tilde{\xi}_t(u) \quad \text{and} \quad \hat{\Pi}_4(u) = \sum_{t=1}^{n} \hat{\sigma}_t^{-1} \nu_t(u) [\tilde{\xi}_t(u) - \xi_t(u)].
\]

Moreover, by the Taylor expansion, we have

\[
\nu_t(u) = u \frac{\partial \hat{q}_t(\theta^*)}{\partial \theta} \quad \text{and} \quad \hat{\nu}_t(u) = u \frac{\partial \hat{q}_t(\theta^*)}{\partial \theta}, \tag{S.26}
\]

where \( \theta^* \) is between \( \theta_{r_0} \) and \( \theta \). Then it follows that

\[
\hat{\Pi}_1(u) = -u' \sum_{t=1}^{n} \frac{1}{\hat{\sigma}_t} \left[ \frac{\partial \hat{q}_t(\theta^*)}{\partial \theta} - \frac{\partial \hat{q}_t(\theta^*)}{\partial \theta} \right] \psi_t(\varepsilon_t - b_{\tau} \hat{\sigma}_t \sigma_t^{-1}).
\]

By Lemma S.1(iv), Lemma S.2(ii) and Assumption 1, together with the fact that \( |\psi_t(x)| < 1 \), it follows that

\[
\sup_{\theta} \frac{|\hat{\Pi}_1(u)|}{\sqrt{n} \|u\|} \leq \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sup_{\theta} \frac{1}{\sigma_t(\lambda)} \left\| \frac{\partial \hat{q}_t(\theta)}{\partial \theta} - \frac{\partial \hat{q}_t(\theta^*)}{\partial \theta} \right\| \sup_{\theta} \frac{\sigma_t(\lambda)}{\sigma_t(\hat{\lambda}_n)} \leq \frac{C \xi_p}{\sqrt{n}} \sum_{t=1}^{n} \rho^t \xi_{p,t} \sup_{\theta} \frac{\sigma_t(\lambda)}{\sigma_t(\hat{\lambda}_n)} = o_p(1).
\]

Hence, we have

\[
\hat{\Pi}_1(u) = o_p(\sqrt{n} \|u\|). \tag{S.27}
\]

We next consider \( \hat{\Pi}_2(u) \). Since \( \psi_t(x) = \tau - I(x < 0) \), by the Taylor expansion, we have

\[
E \left[ \psi_t(\varepsilon_t - b_{\tau} \hat{\sigma}_t \sigma_t^{-1}) - \psi_t(\varepsilon_{t,\tau}) | \mathcal{F}_{t-1} \right] = f_\varepsilon(b_{\tau_1}) b_{\tau} \sigma_t^{-1} (\sigma_t - \hat{\sigma}_t),
\]

where \( b_{\tau_1} \) is between \( b_{\tau} \) and \( b_{\tau} \hat{\sigma}_t \sigma_t^{-1} \). As a result, by the iterated-expectation, the Cauchy-Schwarz inequality and the Taylor expansion in (S.26), together with \( \hat{\sigma}_t^2 \geq 1 \), Lemmas S.1-S.2, \( E(\xi_{p,r,t}^2) < \infty \) by Assumption 3, \( E(\xi_p^2) < \infty \) by Assumption 1 and \( E(u_t^2) < \infty \), and
sup_{x \in \mathbb{R}} f_{\varepsilon}(x) < \infty$ by Assumption 4, it follows that
\[
E \left[ \sup_{\Theta} \frac{\hat{\Pi}_2(u)}{\sqrt{n}\|u\|} \right] \leq \frac{1}{\sqrt{n}} \sum_{i=1}^{n} E \left\{ \frac{1}{\sigma_i} \sup_{\Theta} \left\| \frac{\partial q_i(\Theta)}{\partial \Theta} \right\| E \left[ \psi_{\tau}(\varepsilon \tau - b_{\tau} \tilde{\sigma}_1 \sigma_i^{-1}) - \psi_{\tau}(\varepsilon \tau, |\mathcal{F}_{t-1}|) \right] \right\}
\[
\leq \theta \sup_{x \in \mathbb{R}} f_{\varepsilon}(x) \frac{1}{\sqrt{n}} \sum_{i=1}^{n} E \left\{ |\sigma_i - \tilde{\sigma}_1| \sup_{\Theta} \sigma_i(\lambda) \sup_{\Theta} \left\| \frac{1}{\sigma_i(\lambda)} \frac{\partial q_i(\Theta)}{\partial \Theta} \right\| \right\}
\[
\leq C \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \rho_i \left[ E(\xi_i^2) \right]^{1/2} \left[ E \sup_{\Theta} \frac{\sigma_i^4(\lambda)}{\sigma_i^2} \right]^{1/4} \left[ E(\xi_i^4) \right]^{1/4} = o(1).
\]

Therefore,
\[
\hat{\Pi}_2(u) = o_p(\sqrt{n}\|u\|) \tag{S.28}
\]

Note that $|\hat{\xi}_i(u)| < 2$. For $\hat{\Pi}_3(u)$, similar to the proof of $\hat{\Pi}_1(u)$, we can show that
\[
\hat{\Pi}_3(u) = o_p(\sqrt{n}\|u\|) \tag{S.29}
\]

Finally, we consider $\hat{\Pi}_4(u)$. By the Taylor expansion, we have
\[
E[\hat{\xi}_i(u) - \xi_i(u)|\mathcal{F}_{t-1}]
\[
= \int_0^1 \left[ F_{\varepsilon}(b_{\tau} \tilde{\sigma}_1 \sigma_i^{-1} \nu_i(u)) - F_{\varepsilon}(b_{\tau} \tilde{\sigma}_1 \sigma_i)^s \right] ds - \int_0^1 \left[ F_{\varepsilon}(b_{\tau} \tilde{\sigma}_1 \sigma_i^{-1} \nu_i(u)) - F_{\varepsilon}(b_{\tau}) \right] ds
\[
= \int_0^1 f_{\varepsilon}(b_{\tau} \sigma_i^{-1} \nu_i(u)) ds - \int_0^1 f_{\varepsilon}(b_{\tau} \sigma_i^{-1} \nu_i(u)) ds
\[
= \frac{1}{2} f_{\varepsilon}(b_{\tau} \sigma_i^{-1} \nu_i(u)) + \sigma_i^{-1} \nu_i(u) \int_0^1 \left[ f_{\varepsilon}(b_{\tau} \sigma_i^{-1} \nu_i(u)) - f_{\varepsilon}(b_{\tau}) \right] ds
\]
\[
\leq 0 \quad \text{with } 0 \leq s \leq 1.
\]

As a result, by the iterated-expectation and the Taylor expansion in (S.26), it follows that
\[
E \left[ \sup_{\Theta} \frac{\hat{\Pi}_4(u)}{n\|u\|^2} \right] \leq \hat{K}_1(u) + \hat{K}_2(u) + \hat{K}_3(u), \tag{S.30}
\]
where
\[
\tilde{K}_1(u) = \frac{1}{2} \sum_{t=1}^{n} E \left[ \frac{1}{\sigma_t} \sup_{\theta} \left| \frac{\partial q_t(\theta)}{\partial \theta} \right| \left| \frac{\partial \tilde{q}_t(\theta)}{\partial \theta'} - \frac{\partial \tilde{q}_t(\theta)}{\partial \theta'} \right| \right],
\]
\[
\tilde{K}_2(u) = \frac{1}{n} \sum_{t=1}^{n} E \left[ \frac{1}{\sigma_t} \sup_{\theta} \left| \frac{\partial q_t(\theta)}{\partial \theta} \right| \left| \frac{\partial \tilde{q}_t(\theta)}{\partial \theta'} \right| \sup_{\theta} \int_{0}^{1} |f_\varepsilon(b_{r+2}) - f_\varepsilon(b_r)| ds \right],
\]
\[
\tilde{K}_3(u) = \sum_{t=1}^{n} E \left[ \frac{1}{\sigma_t} \sup_{\theta} \left| \frac{\partial q_t(\theta)}{\partial \theta} \right|^2 \sup_{\theta} \int_{0}^{1} |f_\varepsilon(b_{r+3}) - f_\varepsilon(b_r)| ds \right].
\]

First, consider \( \tilde{K}_1(\theta) \). By the Hölder inequality and Lemmas S.1-S.2, together with \( E(\xi_{\rho,t}^{4+\delta}) < \infty \) by Assumption 3, \( E(\xi_{\rho}^2) < \infty \) by Assumption 1 and \( E(u_t^2) < \infty \), and the fact that \( \tilde{\sigma}_t^2 \leq \sigma_t^2 \), we have
\[
\tilde{K}_1(u) \leq C n \sum_{t=1}^{n} \left( \sup_{\theta} \frac{\sigma_t^2(\lambda)}{\sigma_t^2} \right) \left\{ \sup_{\theta} \frac{1}{\sigma_t(\lambda)} \left| \frac{\partial q_t(\theta)}{\partial \theta} \right| \left| \frac{\partial \tilde{q}_t(\theta)}{\partial \theta'} - \frac{\partial \tilde{q}_t(\theta)}{\partial \theta'} \right| \right\} \]
\[
\leq C n \sum_{t=1}^{n} \rho'E \left( \sup_{\theta} \frac{\sigma_t^2(\lambda)}{\sigma_t^2} \right) \left[ E(\xi_{\rho,t}^{4+\delta}) \right]^{\frac{1}{2}} \left[ E(\xi_{\rho}^{4}) \right]^{\frac{1}{2}} \left( E \sup_{\theta} \left| \frac{\sigma_t(\lambda)}{\sigma_t} \right|^{\frac{4+\delta}{3+\delta}} \right) \]
\[
= o(1).
\]
Hence,
\[
\tilde{K}_1(u) = o_p(1). \tag{S.31}
\]

Next, we consider \( \tilde{K}_2(u) \). By the Taylor expansion, it follows that
\[
\int_{0}^{1} |f_\varepsilon(b_{r+2}) - f_\varepsilon(b_r)| ds \leq \frac{1}{2\sigma_t} \sup_{x \in \mathbb{R}} |f_\varepsilon(x)||b_r(\tilde{\sigma}_t - \sigma_t) + \tilde{\nu}_t(u)s_2|.
\]

Then by the Taylor expansion in (S.26) and the Hölder inequality, together with Lemmas
S.1-S.2, $\sup_{x \in \mathbb{R}} |\tilde{f}_z(x)| < \infty$ by Assumption 4, $|b_r| < \overline{b}$ and the fact that $\sigma^2_t \geq 1$, we have

$$
\sup_{|u| \leq \eta} \tilde{K}_2(u) \leq \frac{C}{n} \sum_{t=1}^{n} E \left\{ \sup_{\theta} \left[ \frac{1}{\sigma_t} \frac{\partial_q(\theta)}{\theta} \left| \frac{1}{\sigma_t} \frac{\partial_q(\theta)}{\theta} \right| \right] \left[ \frac{\overline{b}}{\sigma_t} |\tilde{\sigma}_t - \sigma_t| + \sup_{|u| \leq \eta} \left| \tilde{\nu}_t(u) \right| \right] \right\}
$$

$$
\leq \frac{C}{n} \sum_{t=1}^{n} \left\{ E \left[ \sup_{\theta} \frac{\sigma_t(\lambda)}{\sigma_t} \right]^{\frac{2}{3+\delta}} \left[ E(\zeta^4_{p,t}) \right]^{\frac{2}{3+\delta}} \right\}
$$

$$
+ \frac{C\eta}{n} \sum_{t=1}^{n} \left\{ E \left[ \sup_{\theta} \frac{\sigma_t(\lambda)}{\sigma_t} \right]^{\frac{4}{3+\delta}} \left[ E(\zeta^2_{p,t}) \right]^{\frac{2}{3+\delta}} \right\}
$$

tends to 0 as $\eta \to 0$ and $n$ is large enough. Similar to (S.50) and (S.51), we can show that

$$
\tilde{K}_2(u) = o_p(1).
$$

(S.32)

We finally consider $\tilde{K}_3(u)$. By the Taylor expansion, we have

$$
\int_0^1 |f_z(b,t) - f_z(b,r)|sd\sigma \leq \sup_{x \in \mathbb{R}} |\tilde{f}_z(x)| \sigma_t^{-1} |\nu_t(u)|.
$$

Then by the Hölder inequality and Lemma S.1, it follows that

$$
\sup_{|u| \leq \eta} \tilde{K}_3(u) \leq \frac{C}{n} \sum_{t=1}^{n} E \left\{ \sup_{\theta} \left[ \frac{1}{\sigma_t \sigma_t} \left| \frac{\partial_q(\theta)}{\theta} \right| \right]^{\frac{2}{3+\delta}} \sup_{|u| \leq \eta} \left| \nu_t(u) \right| \right\}
$$

$$
\leq \frac{C\eta}{n} \sum_{t=1}^{n} \left\{ E \left[ \sup_{\theta} \frac{\sigma_t(\lambda)}{\sigma_t} \right]^{\frac{3}{3+\delta}} \left[ E(\zeta^3_{p,t}) \right] \right\}
$$

tends to 0 as $\eta \to 0$ and $n$ is large enough. Similar to (S.50) and (S.51), we can show that

$$
\tilde{K}_3(u) = o_p(1).
$$

(S.33)

In view of (S.30)-(S.33), we have

$$
\tilde{P}_4(u) = o_p(n\|u\|^2).
$$

(S.34)

Combing (S.25), (S.27)-(S.29) and (S.34), we assert (S.24). By (S.21), (S.22) and (S.24),
it follows that
\[ n[\hat{L}_n(\theta) - \tilde{L}_n(\theta_0)] - n[\tilde{L}_n(\theta) - \tilde{L}_n(\theta_0)] = o_p(\sqrt{n}\|u\| + n\|u\|^2). \]  
\[ (S.35) \]

Hence, the proof of this lemma is complete.

Proof of Lemma S.5. Denote \( u = \theta - \theta_0 \). Recall that 
\( \tilde{L}_n(\theta) = n^{-1} \sum_{t=1}^{n} \tilde{\sigma}_t^{-1} \ell_t(\theta) \) and 
\( L_n(\theta) = n^{-1} \sum_{t=1}^{n} \sigma_t^{-1} \ell_t(\theta) \), where \( \ell_t(\theta) = \rho_t[Y_t - q_t(\theta)] \). To show this lemma, we decompose the proof into two steps. In the first step, we will show that
\[ n[\tilde{L}_n(\theta) - \tilde{L}_n(\theta_0)] - n[L_n(\theta) - L_n(\theta_0)] = o_p(\sqrt{n}\|u\| + n\|u\|^2). \]  
\[ (S.36) \]

Denote \( \varepsilon_{t,\tau} = \varepsilon_t - b_{\tau} \) and \( \nu_t(u) = q_t(\theta) - q_t(\theta_0) \). Define
\[ \xi_t(u) = \int_0^1 \left[I(\varepsilon_t \leq b_{\tau} + \sigma_t^{-1}\nu_t(u)s) - I(\varepsilon_t \leq b_{\tau})\right]ds. \]  
\[ (S.37) \]

By the Knight equation \( (S.66) \), it can be verified that
\[ n[\tilde{L}_n(\theta) - \tilde{L}_n(\theta_0)] - n[L_n(\theta) - L_n(\theta_0)] \]
\[ = \sum_{t=1}^{n} (\tilde{\sigma}_t^{-1} - \sigma_t^{-1}) \left[ \rho_{\tau} (\varepsilon_{t,\tau}\sigma_t - \nu_t(u)) - \rho_{\tau} (\varepsilon_{t,\tau}\sigma_t) \right] \]
\[ = K_{1n}(u) + K_{2n}(u), \]  
\[ (S.38) \]

where
\[ K_{1n}(u) = -\sum_{t=1}^{n} \left( \frac{1}{\tilde{\sigma}_t} - \frac{1}{\sigma_t} \right) \nu_t(u) \psi_t(\varepsilon_{t,\tau}) \] and 
\( K_{2n}(u) = \sum_{t=1}^{n} \left( \frac{1}{\tilde{\sigma}_t} - \frac{1}{\sigma_t} \right) \nu_t(u) \xi_t(u) \).

By the Taylor expansion, it holds that
\[ \frac{1}{\tilde{\sigma}_t} - \frac{1}{\sigma_t} = -(\tilde{\lambda}_n - \lambda_0)' \frac{1}{2\sigma^2_t(\lambda_0^*)} \frac{\partial \sigma^2_t(\lambda^*)}{\partial \lambda} \] and 
\( \nu_t(u) = u \frac{\partial q_t(\theta^*)}{\partial \theta} \),  
\[ (S.39) \]
where $\lambda^*$ is between $\lambda_{n}^{\text{int}}$ and $\lambda_0$, and $\theta^*$ is between $\theta$ and $\theta_{\tau_0}$. By (S.39), it follows that

$$K_{1n}(u) = \sqrt{n}(\hat{\lambda}_n^{\text{int}} - \lambda_0)' \cdot \frac{1}{n} \sum_{t=1}^{n} Z_{1t}(\theta^*) \cdot \sqrt{n}u,$$

(S.40)

where

$$Z_{1t}(\theta^*) = \frac{1}{2\sigma_t^2(\lambda^*)} \frac{\partial \sigma_t^2(\lambda^*)}{\partial \lambda} \frac{\partial q_t(\theta^*)}{\partial \theta'} \psi_{t,\tau}(\varepsilon_{t,\tau}).$$

By the iterated-expectation and Lemma S.1, together with the fact that $E[\psi_{t,\tau}(\varepsilon_{t,\tau})] = 0$, we can show that $E[Z_{1t}(\theta^*)] = 0$. Furthermore, by the Cauchy-Schwarz inequality, Lemma S.1 and the fact that $|\psi_{t}(x)| \leq 1$, we have

$$E \left( \sup_{\theta} \|Z_{1t}(\theta^*)\| \right) \leq \frac{1}{2} \left[ E \sup_{\theta} \left\| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \lambda} \right\|^2 \right]^{1/2} \left[ E \sup_{\theta} \left\| \frac{1}{\sigma_t(\lambda)} \frac{\partial q_t(\theta)}{\partial \theta} \right\|^2 \right]^{1/2} < \infty.$$

Then by the Theorem 3.1 in Ling and McAleer (2003), it follows that

$$\sup_{\theta} \left\| \frac{1}{n} \sum_{t=1}^{n} Z_{1t}(\theta^*) \right\| = o_p(1).$$

This together with (S.40) and the fact that $\sqrt{n}(\hat{\lambda}_n^{\text{int}} - \lambda_0) = O_p(1)$, implies that

$$K_{1n}(u) = o_p(\sqrt{n}\|u\|).$$

(S.41)

For $K_{2n}(u)$, by (S.39), it can be verified that

$$K_{2n}(u) = -(\hat{\lambda}_n - \lambda_0)'Z_{2t}(u),$$

(S.42)

where

$$Z_{2t}(u) = \sum_{t=1}^{n} \frac{1}{2\sigma_t^2(\lambda^*)} \frac{\partial \sigma_t^2(\lambda^*)}{\partial \lambda} \nu_t(u)\xi_t(u).$$
By the Taylor expansion and Assumption 4, it follows that

\[
E[\xi_t(u)|\mathcal{F}_{t-1}] = \int_0^1 \left[ F_{\varepsilon} \left( b_r + \frac{\nu_t(u) s}{\sigma_t} \right) - F_{\varepsilon}(b_r) \right] ds = \int_0^1 f_{\varepsilon} \left( b_r + \frac{\nu_t(u) s^*}{\sigma_t} \right) \frac{\nu_t(u)}{\sigma_t} ds,
\]

where \(s^*\) is between 0 and \(s\). Then by the iterated-expectation and the Cauchy-Schwarz inequality, together with (S.39), Lemma S.1 and \(\sup_{x \in \mathbb{R}} f_{\varepsilon}(x) < \infty\) by Assumption 4, we have

\[
E \left[ \sup_{\Theta} \| Z_2(u) \| \right] \\
\leq \frac{1}{4} \sup_{x \in \mathbb{R}} f_{\varepsilon}(x) \frac{1}{n} \sum_{t=1}^n E \left[ \sup_{\Theta} \left| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \lambda} \right| \sup_{\Theta} \frac{\nu_t^2(\Theta)}{\sigma_t \sigma_t(\lambda^*) \| u \|^2} \right] \\
\leq C E \left[ \sup_{\Theta} \left| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \lambda} \right| \sup_{\Theta} \frac{\nu_t^2(\Theta)}{\sigma_t \sigma_t(\lambda^*) \| u \|^2} \right] \\
\leq C \left[ E \sup_{\Theta} \left| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \lambda} \right| \right]^{1/4} \left[ E \sup_{\Theta} \frac{\sigma_t^4(\lambda)}{\sigma_t^2} \right]^{1/4} \left[ E \sup_{\Theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial \nu_t(\Theta)}{\partial \lambda} \right| \right]^{1/4} \leq C \left[ E \sup_{\Theta} \left| \frac{1}{\sigma_t^2(\lambda)} \frac{\partial \sigma_t^2(\lambda)}{\partial \lambda} \right| \right]^{1/4} \left[ E \sup_{\Theta} \frac{\sigma_t^4(\lambda)}{\sigma_t^2} \right]^{1/4} \left[ E \sup_{\Theta} \left| \frac{1}{\sigma_t(\lambda)} \frac{\partial \nu_t(\Theta)}{\partial \lambda} \right| \right]^{1/4} \leq C < \infty.
\]

This together with (S.42), the ergodic theorem and the fact that \(\sqrt{\mu(\hat{\lambda}_n^\text{int} - \lambda_0)} = O_p(1)\), implies that

\[
K_{2n}(u) = o_p(n\|u\|^2). \tag{S.43}
\]

By (S.38), (S.41) and (S.43), it follows that (S.36) holds.

In the second step, we will establish that

\[
n[L_n(\theta) - L_n(\theta_{r0})] = -\sqrt{n} u' T_n + \sqrt{n} u' J_n \sqrt{n} u + o_p(\sqrt{n}\|u\| + n\|u\|^2). \tag{S.44}
\]

By the Knight equation (S.66), we have

\[
n[L_n(\theta) - L_n(\theta_{r0})] = R_{1n}(u) + R_{2n}(u), \tag{S.45}
\]

where

\[
R_{1n}(u) = -\sum_{t=1}^n \frac{1}{\sigma_t} \nu_t(u) \psi(\varepsilon_{t,\tau}) \quad \text{and} \quad R_{2n}(u) = \sum_{t=1}^n \frac{1}{\sigma_t} \nu_t(u) \xi_t(u).
\]
By the Taylor expansion, we have $\nu_t(u) = q_{1t}(u) + q_{2t}(u)$, where

$$q_{1t}(u) = u' \frac{\partial q_t(\theta_{t0})}{\partial \theta}$$

and

$$q_{2t}(u) = \frac{u' \frac{\partial^2 q_t(\theta^*)}{\partial \theta \partial \theta'}}{2} u$$

with $\theta^*$ between $\theta$ and $\theta_{t0}$. Then it can be verified that

$$R_{1n}(u) = -\sqrt{n} u'T_n - \sqrt{n} u'K_{3n}(\theta^*)\sqrt{n}u,$$  \hspace{1cm} (S.46)

where

$$T_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \psi_t(\varepsilon_{t,\tau})$$

and

$$K_{3n}(\theta^*) = \frac{1}{2n} \sum_{t=1}^{n} \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta^*)}{\partial \theta \partial \theta'} \psi_t(\varepsilon_{t,\tau}).$$

By the iterated-expectation and the fact that $E[\psi_t(\varepsilon_{t,\tau})] = 0$, it follows that

$$E \left[ \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta^*)}{\partial \theta \partial \theta'} \psi_t(\varepsilon_{t,\tau}) \right] = 0.$$  

Moreover, by Lemma S.1 and the fact that $|\psi_t(\varepsilon_{t,\tau})| \leq 1$, we have

$$E \left[ \sup_{\theta} \left\| \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \psi_t(\varepsilon_{t,\tau}) \right\| \right] \leq \left[ E \sup_{\theta} \frac{\sigma_t^2(\lambda)}{\sigma_t^2} \right]^{1/2} \left[ E \sup_{\theta} \left\| \frac{1}{\sigma_t(\lambda)} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \right\|^2 \right]^{1/2} < \infty.$$  

Then by the Theorem 3.1 in Ling and McAleer (2003), it follows that

$$\sup_{\theta} \|K_{3n}(\theta^*)\| = o_p(1).$$

This together with (S.46), implies that

$$R_{1n}(u) = -\sqrt{n} u'T_n + o_p(n\|u\|^2).$$  \hspace{1cm} (S.47)
By simple calculation, we have $\xi_t(u) = \xi_{1t}(u) + \xi_{2t}(u)$, where

$$
\xi_{1t}(u) = \int_0^1 [I(\varepsilon_t \leq b_r + \sigma^{-1}_t q_{1t}(u)s) - I(\varepsilon_t \leq b_r)] ds
$$

and

$$
\xi_{2t}(u) = \int_0^1 [I(\varepsilon_t \leq b_r + \sigma^{-1}_t \nu_t(u)s) - I(\varepsilon_t \leq b_r + \sigma^{-1}_t q_{1t}(u)s)] ds.
$$

Then for $R_{2n}(u)$, it can be verified that

$$
R_{2n}(u) = K_{4n}(u) + K_{5n}(u) + K_{6n}(u) + K_{7n}(u),
$$

where

$$
K_{4n}(u) = u' \sum_{t=1}^n \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} E[\xi_{1t}(u) | F_{t-1}],
$$

$$
K_{5n}(u) = u' \sum_{t=1}^n \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \{\xi_{1t}(u) - E[\xi_{1t}(u) | F_{t-1}]\},
$$

$$
K_{6n}(u) = \frac{u'}{2} \sum_{t=1}^n \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta_{t0})}{\partial \theta \partial \theta'} \xi_{2t}(u) \quad \text{and} \quad K_{7n}(u) = \frac{u'}{2} \sum_{t=1}^n \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta_{t0})}{\partial \theta \partial \theta'} \xi_{1t}(u) u.
$$

First, consider $K_{4n}(u)$. By the Taylor expansion, it follows that

$$
E[\xi_{1t}(u) | F_{t-1}] = \int_0^1 [F_x(b_r + \sigma^{-1}_t q_{1t}(u)s) - F_x(b_r)] ds
$$

$$
= \frac{1}{2} f_x(b_r) \sigma^{-1}_t q_{1t}(u) + \sigma^{-1}_t q_{1t}(u) \int_0^1 [f_x(b_r + \sigma^{-1}_t q_{1t}(u)s^*) - f_x(b_r)] s ds,
$$

where $s^*$ is between 0 and $s$. Therefore, we have

$$
K_{4n}(u) = \sqrt{n} u' J_n \sqrt{n} u + \sqrt{n} u' \Pi_{1n}(u) \sqrt{n} u,
$$

where

$$
J_n = \frac{f_x(b_r)}{2n} \sum_{t=1}^n \frac{1}{\sigma^2_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \frac{\partial q_t(\theta_{t0})}{\partial \theta'} \quad \text{and}
$$

$$
\Pi_{1n}(u) = \frac{1}{n} \sum_{t=1}^n \frac{1}{\sigma^2_t} \frac{\partial q_t(\theta_{t0})}{\partial \theta} \frac{\partial q_t(\theta_{t0})}{\partial \theta'} \int_0^1 [f_x(b_r + \sigma^{-1}_t q_{1t}(u)s^*) - f_x(b_r)] s ds.
$$

By the iterated-expectation, the Taylor expansion, Lemma S.1 and $\sup_{x \in \mathbb{R}} |\tilde{f}(x)| < \infty$ by
Assumption 4, for any $\eta > 0$, it holds that

$$
E \left( \sup_{||u|| \leq \eta} \| \Pi_{1n}(u) \| \right) \leq \frac{1}{n} \sum_{t=1}^{n} E \left( \sup_{||u|| \leq \eta} \left\| \frac{1}{\sigma_{t}^{2}} \frac{\partial q_{t}^{*}(\theta_{0}) \partial q_{t}^{*}(\theta_{0})}{\partial \theta} \sup_{x \in \mathbb{R}} |\hat{f}(x)| \frac{u^t \cdot \partial q_{t}^{*}(\theta_{0})}{\partial \theta} \right\| \right)
$$

$$
\leq \eta \sup_{x \in \mathbb{R}} |\hat{f}(x)| E \left( \left\| \frac{1}{\sigma_{t}} \frac{\partial q_{t}^{*}(\theta_{0})}{\partial \theta} \right\|^{3} \right)
$$

tends to 0 as $\eta \to 0$. Therefore, for any $\epsilon, \delta > 0$, there exists $\eta_{0} = \eta_{0}(\epsilon) > 0$ such that

$$
P \left( \sup_{||u|| \leq \eta_{0}} \| \Pi_{1n}(u) \| > \delta \right) < \frac{\epsilon}{2}
$$

(S.50)

for all $n \geq 1$. Since $u = o_{p}(1)$, it follows that

$$
P (\|u\| > \eta_{0}) < \frac{\epsilon}{2}
$$

(S.51)

as $n$ is large enough. From (S.50) and (S.51), we have

$$
P (\|\Pi_{1n}(u)\| > \delta) \leq P (\|\Pi_{1n}(u)\| > \delta, \|u\| \leq \eta_{0}) + P (\|u\| > \eta_{0})
$$

$$
\leq P \left( \sup_{||u|| \leq \eta_{0}} \| \Pi_{1n}(u) \| > \delta \right) + \frac{\epsilon}{2} < \epsilon
$$

as $n$ is large enough. Therefore, $\Pi_{1n}(u) = o_{p}(1)$. This together with (S.49), implies that

$$
K_{4n}(u) = \sqrt{n}u^t J_n \sqrt{n}u + o_{p}(n\|u\|^2).
$$

(S.52)

For $K_{5n}(u)$, by Lemma S.3, it holds that

$$
K_{5n}(u) = o_{p}(\sqrt{n}\|u\| + n\|u\|^2).
$$

(S.53)

Next, we consider $K_{6n}(u)$. By the Taylor expansion, it follows that

$$
E[\xi_{2t}(u)|\mathcal{F}_{t-1}] = \int_{0}^{1} \left[ F_{\epsilon}(b_{r} + \sigma_{t}^{-1}\nu(u)s) - F_{\epsilon}(b_{r} + \sigma_{t}^{-1}q_{1t}(u)s) \right] ds
$$

$$
= \sigma_{t}^{-1}q_{2t}(u) \int_{0}^{1} f_{\epsilon}(b^{*}) ds,
$$

where $b^{*}$ is between $b_{r} + \sigma_{t}^{-1}q_{1t}(u)s$ and $b_{r} + \sigma_{t}^{-1}\nu(u)$. Then by the iterated-expectation
and the Cauchy-Schwarz inequality, together with Lemma S.1 and sup<sub>x∈R</sub> f<sub>ε</sub>(x) < ∞ by Assumption 4, for any η > 0, it holds that

\[
E \left( \sup_{|u| \leq \eta} \left| K_{\theta n}(u) \right| \right) 
\leq \frac{\eta}{n} \sum_{t=1}^{n} E \left\{ \left\| \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{\tau 0})}{\partial \theta} \right\| \frac{1}{2} \sup_{x \in \mathbb{R}} f_{\varepsilon}(x) \sup_{\theta} \left\| \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \right\| \right\} 
\leq C \eta \left[ \frac{1}{\sigma_t} \frac{\partial q_t(\theta_{\tau 0})}{\partial \theta} \right]^{4/4} \left[ E \sup_{\theta} \frac{\sigma_t^4(\lambda)}{\sigma_t^4} \right]^{1/4} \left[ E \sup_{\theta} \left\| \frac{1}{\sigma_t(\lambda)} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \right\|^{2} \right]^{1/2} 
\text{(S.54)}
\]
tends to 0 as η → 0. Similar to (S.50) and (S.51), we can show that

\[K_{\theta n}(u) = o_p(n\|u\|^2).\]  
\text{(S.55)}

Finally, for \(K_{\theta n}(u)\), it follows that

\[K_{\theta n}(u) = \sqrt{n} u \Pi_{2n}(u) \sqrt{n} u + \sqrt{n} u \Pi_{3n}(u) \sqrt{n} u, \]
\text{(S.56)}

where

\[\Pi_{2n}(u) = \frac{1}{2n} \sum_{t=1}^{n} \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta^*)}{\partial \theta \partial \theta'} \xi_{1t}(u) \text{ and } \Pi_{3n}(u) = \frac{1}{2n} \sum_{t=1}^{n} \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta^*)}{\partial \theta \partial \theta'} \xi_{2t}(u).\]

As for the proof in (S.54), we can show that

\[
E \left( \sup_{|u| \leq \eta} \left\| \Pi_{2n}(u) \right\| \right) 
\leq C \eta \left[ E \sup_{\theta} \left\| \frac{1}{\sigma_t} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \right\|^{2} \right]^{1/2} \left[ E \sup_{\theta} \frac{\sigma_t^4(\lambda)}{\sigma_t^4} \right]^{1/4} \left[ E \left\| \frac{1}{\sigma_t(\lambda)} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'} \right\|^2 \right]^{1/4}
\]
tends to 0 as η → 0. Then similar to (S.50) and (S.51), we can prove that \(\Pi_{2n}(u) = o_p(1)\).
Similarly, for \( \Pi_{3n}(u) \), by the Hölder inequality, for some \( \delta > 0 \), we have

\[
E\left( \sup_{u \in \eta} \| \Pi_{3n}(u) \| \right) \leq C \eta^2 \left\{ E \sup_{\theta} \left[ \frac{1}{\sigma_t(\lambda)} \frac{\partial^2 q_t(\theta)}{\partial \theta \partial \theta'}^2 \right]^{\frac{2}{2+\delta}} \right\}^{\frac{2}{2+\delta}}
\]

tends to 0 as \( \eta \to 0 \). This together with the proof similar to that of (S.50) and (S.51), implies that \( \Pi_{3n}(u) = o_p(1) \). Therefore, by (S.56), it follows that

\[
K_{\gamma_n}(u) = o_p(n\|u\|^2).
\]

(S.57)

Combing (S.48), (S.52), (S.53), (S.55) and (S.57), we have

\[
R_{2n}(u) = \sqrt{n}u'J_n\sqrt{n}u + o_p(\sqrt{n}\|u\| + n\|u\|^2).
\]

(S.58)

By (S.45), (S.47) and (S.58), it follows that (S.44) holds. In view of (S.36) and (S.44), the proof of this lemma is complete.

Proof of Theorem 1. Denote \( \ell_t(\theta) = \rho_\tau[Y_t - q_t(\theta)] \) and \( \tilde{\ell}_t(\theta) = \rho_\tau[Y_t - \tilde{q}_t(\theta)] \), where \( q_t(\theta) = \phi' X_{t-1} + b \sigma_t(\lambda) \) and \( \tilde{q}_t(\theta) = \phi' X_{t-1} + b \tilde{\sigma}_t(\lambda) \). Define the following functions

\[
\hat{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n \hat{\sigma}_t^{-1} \tilde{\ell}_t(\theta), \quad \tilde{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n \tilde{\sigma}_t^{-1} \ell_t(\theta) \quad \text{and} \quad \tilde{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^n \tilde{\sigma}_t^{-1} \tilde{\ell}_t(\theta),
\]

where \( \hat{\sigma}_t = \tilde{\sigma}_t(\hat{\lambda}_n^{n}) \), \( \tilde{\sigma}_t = \tilde{\sigma}_t(\tilde{\lambda}_n^{n}) \) and \( \sigma_t = \sigma_t(\lambda_0) \). To show the consistency of \( \hat{\theta}_m \), we first verify the following claims:

(i) \( \sup_{\theta} [\hat{L}_n(\theta) - L_n(\theta)] = o_p(1) \);

(ii) \( E[\sup_{\theta} \sigma_t^{-1} \ell_t(\theta)] < \infty \);

(iii) \( E[\sigma_t^{-1} \ell_t(\theta)] \) has a unique minimum at \( \theta^* \);

(iv) For any \( \theta^* \in \Theta \), \( E[\sup_{\theta \in B_\eta(\theta^*)} \sigma_t^{-1}[\ell_t(\theta) - \ell_t(\theta^*)]] \to 0 \) as \( \eta \to 0 \), where \( B_\eta(\theta^*) = \{ \theta \in \Theta : \| \theta - \theta^* \| < \eta \} \) is an open neighborhood of \( \theta^* \) with radius \( \eta > 0 \).
To prove Claim (i), we need to verify that

$$\sup_{\theta \in \Theta} |\hat{L}_n(\theta) - \tilde{L}_n(\theta)| = o_p(1)$$  \hspace{1cm} (S.59)

and

$$\sup_{\theta \in \Theta} |L_n(\theta) - \tilde{L}_n(\theta)| = o_p(1).$$  \hspace{1cm} (S.60)

We first show that (S.60) holds. By Taylor expansion and the fact that $\sigma_t^2(\lambda) \geq 1$,

$$|\hat{\sigma}_t^{-1} - \sigma_t^{-1}| = |\sigma_t^{-1}(\hat{\lambda}_n^{int}) - \sigma_t^{-1}(\lambda_0)| \leq \frac{1}{2} \sup_{\theta} \left\| \frac{\partial \sigma_t^2(\lambda)}{\partial \lambda} \right\| \| \hat{\lambda}_n^{int} - \lambda_0 \|.$$

Moreover, by the fact that $|\rho_r(x)| \leq |x|$, together with $E(Y_t^2) < \infty$ and $E \sup_{\Theta} q_t^2(\theta) < \infty$ implied by Assumption 1 and $E(u_t^2) < \infty$, we have

$$E \sup_{\Theta} \{ \rho_r[Y_t - q_t(\theta)] \}^2 \leq 2E(Y_t^2) + 2E \sup_{\Theta} q_t^2(\theta) < \infty.$$  \hspace{1cm} (S.61)

This, together with Lemma S.1, the ergodic theorem and $\hat{\lambda}_n^{int} - \lambda_0 = o_p(1)$, leads to

$$\sup_{\Theta} |L_n(\theta) - \tilde{L}_n(\theta)| \leq \frac{1}{n} \sum_{t=1}^{n} |\hat{\sigma}_t^{-1} - \sigma_t^{-1}| \sup_{\Theta} \rho_r[Y_t - q_t(\theta)] = o_p(1).$$

Hence (S.60) holds. We next verify (S.59). It can be verified that

$$\hat{L}_n(\theta) - \tilde{L}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \hat{\sigma}_t^{-1} \rho_r[Y_t - \tilde{q}_t(\theta)] - \frac{1}{n} \sum_{t=1}^{n} \tilde{\sigma}_t^{-1} \rho_r[Y_t - q_t(\theta)]$$

$$= R_{1n}(\theta) + R_{2n}(\theta),$$  \hspace{1cm} (S.62)

where

$$R_{1n}(\theta) = \frac{1}{n} \sum_{t=1}^{n} (\hat{\sigma}_t^{-1} - \tilde{\sigma}_t^{-1}) \rho_r[Y_t - q_t(\theta)]$$

and

$$R_{2n}(\theta) = \frac{1}{n} \sum_{t=1}^{n} \tilde{\sigma}_t^{-1} \{ \rho_r[Y_t - \tilde{q}_t(\theta)] - \rho_r[Y_t - q_t(\theta)] \}.$$
Recall that \( \xi_{\rho,t} = \sum_{j=0}^{\infty} \rho^j(1 + \|X_{t-j-1}\| + \|V_{t-j-1}\|^{1/2} + |u_{t-j}|) \) and \( \xi_{\rho} = \sum_{j=0}^{\infty} \rho^j(1 + \|X_{j-1}\| + |u_{j-1}|) \), where \( \rho \in (0, 1) \) is a constant. By Assumption 1, it is clear that \( E(\xi_{\rho,t}^2) < \infty \) and \( E(\xi_{\rho}^2) < \infty \). Then similar to the proof in (S.61), by Lemma S.1(i), we can show that \( E(\sup_{\Theta}[Y_t - q_t(\theta)]^2) < \infty \). This together with (S.63), implies that

\[
\sup_{\Theta} |R_{1n}(\theta)| \leq \frac{1}{n} \sum_{i=1}^{n} \sup_{\Theta} |\hat{\sigma}_t^{-1} - \tilde{\sigma}_t^{-1}| \sup_{\Theta} \rho_t[Y_t - q_t(\theta)]
\]

\[
\leq \frac{C \xi_{\rho}}{n} \sum_{t=1}^{n} \rho^t \sup_{\Theta} \rho_t[Y_t - q_t(\theta)] = o_p(1). \tag{S.64}
\]

Note that \( \tilde{q}_t(\theta) - q_t(\theta) = b[\tilde{\sigma}_t(\lambda) - \sigma_t(\lambda)] \) and \( b < \bar{b} \) for \( \theta \in \Theta \). By Lemma S.2(i), the Lipschitz continuity of \( \rho_t(x) \) and the fact that \( \tilde{\sigma}_t^2 \geq 1 \), it follows that

\[
\sup_{\Theta} |R_{2n}(\theta)| \leq \frac{C}{n} \sum_{t=1}^{n} \sup_{\Theta} |	ilde{\sigma}_t(\theta) - \tilde{\sigma}_t(x)| \sup_{\Theta} \rho_t[\tilde{q}_t(\theta) - q_t(\theta)]
\]

\[
\leq \frac{C}{n} \sum_{t=1}^{n} \sup_{\Theta} |	ilde{\sigma}_t(\lambda) - \sigma_t(\lambda)| \leq \frac{C \xi_{\rho}}{n} \sum_{t=1}^{n} \rho^t = o_p(1). \tag{S.65}
\]

From (S.62), (S.64) and (S.65), we show that (S.59) holds. Combing (S.59) and (S.60), Claim (i) is verified. Moreover, Claim (ii) is implied by (S.61) and the fact that \( \sigma_t^2 \geq 1 \).

We next prove Claim (iii). For \( x \neq 0 \), it holds that

\[
\rho_t(x - y) - \rho_t(x) = -y \psi_t(x) + y \int_0^1 [I(x \leq ys) - I(x \leq 0)]ds
\]

\[
= -y \psi_t(x) + (x - y)[I(0 < x > y) - I(0 < x < y)], \tag{S.66}
\]

where \( \psi_t(x) = \tau - I(x < 0) \); see Knight (1998). Denote \( \nu_t(\theta) = q_t(\theta) - q_t(\theta_{t0}) \) and \( \varepsilon_{t,\tau} = \varepsilon_t - b_\tau \). Note that \( \psi_t(\varepsilon_{t,\tau} \sigma_t) = \psi_t(\varepsilon_{t,\tau}) \) and \( E[\psi_t(\varepsilon_{t,\tau})] = 0 \). Then by (S.66), we
have

\[
E[\sigma_t^{-1}\ell_t(\theta)] - E[\sigma_t^{-1}\ell_t(\theta_0)] = E \left\{ \sigma_t^{-1} \left[ \rho_t(\varepsilon_{t,r} - \sigma_t^{-1}\nu_t(\theta)) - \rho_t(\varepsilon_{t,r}) \right] \right\} = E \left\{ \sigma_t^{-1}[\varepsilon_{t,r} - \sigma_t^{-1}\nu_t(\theta)] [I(0 > \varepsilon_{t,r} > \sigma_t^{-1}\nu_t(\theta)) - I(0 < \varepsilon_{t,r} < \sigma_t^{-1}\nu_t(\theta))] \right\} \geq 0,
\]

and, by Assumption 2, the equality holds if and only if \( \nu_t(\theta) = q_t(\theta) - q_t(\theta_0) = 0 \) with probability one. Note that

\[

\nu_t(\theta) = b_\tau \sqrt{1 + \sum_{i=1}^{q} \alpha_i u_{t-1,i}^2 + \sum_{j=1}^{p} \beta_j \sigma_t^2 / \sigma_j^2 + \sum_{k=1}^{d} \pi_k \tau_k^2 u_{t-1,k}^2} - b_\tau \sqrt{1 + \sum_{i=1}^{q} \alpha_i u_{t-1,i}(\phi) + \sum_{j=1}^{p} \beta_j \sigma_t^2 (\lambda) + \sum_{k=1}^{d} \pi_k \tau_k^2 u_{t-1,k}^2 + (\phi - \phi_0)'X_{t-1},}
\]

where \( u_{t-1} = \sigma_{t-1}\varepsilon_{t-1} \), \( u_{t-1}(\phi) = u_{t-1} - (\phi - \phi_0)'X_{t-2} \), \( \sigma_t^2 \) and \( \sigma_t^2 (\lambda) \) are \( \mathcal{F}_{t-1} \)-measurable, and \( \{X_{t-1}\} \) are independent of \( \{\varepsilon_t\} \). As a result, the random variable \( \varepsilon_{t-1} \) in \( \nu_t(\theta) \) is independent of all the others, and then it holds that \( \phi = \phi_0 \) and \( b_\tau \alpha_{i0} = b_\tau \alpha_i \) for \( i \geq 2 \), \( b_\tau = b \) and \( \pi_k \tau_0 = \pi_k \) for \( k = 1, \ldots, d \). Finally, we verify that \( \beta_j = \beta_j \) for \( j = 1, \ldots, p \). Thus, the proof of Claim (iii) is accomplished.

Finally, we assert Claim (iv). By the Taylor expansion, it holds that

\[
|q_t(\theta) - q_t(\theta^\dagger)| \leq \|\theta - \theta^\dagger\| \left\| \frac{\partial q_t(\theta)}{\partial \theta} \right\|,
\]

where \( \theta \) is between \( \theta \) and \( \theta^\dagger \). This together with the Lipschitz continuity of \( \rho_t(x) \), the Cauchy-Schwarz inequality and Lemma S.1, implies that

\[
E[\sup_{\theta \in B_\eta(\theta^\dagger)} \sigma_t^{-1}|\ell_t(\theta) - \ell_t(\theta^\dagger)|] \leq C\eta \left\{ E \sup_{\theta} \frac{\sigma_t^2 (\lambda)}{\sigma_t^2} \right\}^{1/2} \left[ E \sup_{\theta} \left\| \frac{1}{\sigma_t (\lambda)} \frac{\partial q_t(\theta)}{\partial \theta} \right\|^2 \right]^{1/2}
\]
tends to 0 as \( \eta \to 0 \). Hence, Claim (iv) holds.

Based on Claims (i)-(iv), by a method similar to that in Huber (1973), we next verify
the consistency. Let $V$ be any open neighborhood of $\theta_0 \in \Theta$. By Claim (iv), for any $\theta^* \in V^c = \Theta/V$ and $\epsilon > 0$, there exists an $\eta_0 > 0$ such that

$$E[\min_{\theta \in \mathbb{B}_\eta(\theta^*)} \sigma_t^{-1}\ell_t(\theta)] \geq E[\sigma_t^{-1}\ell_t(\theta^*)] - \epsilon. \quad (S.67)$$

From Claim (ii), by the ergodic theorem, it follows that

$$\frac{1}{n} \sum_{t=1}^n \min_{\theta \in \mathbb{B}_\eta(\theta^*)} \sigma_t^{-1}\ell_t(\theta) \geq \min_{\theta \in \mathbb{B}_\eta(\theta^*)} \sigma_t^{-1}\ell_t(\theta) - \epsilon \quad (S.68)$$

as $n$ is large enough. Since $V^c$ is compact, we can choose $\{\mathbb{B}_\eta(\theta_i) : \theta_i \in V^c, i = 1, \ldots, k\}$ to be a finite covering of $V^c$. Then by (S.67) and (S.68), as $n$ is large enough, we have

$$\inf_{\theta \in V^c} L_n(\theta) = \min_{1 \leq i \leq k} \inf_{\theta \in \mathbb{B}_\eta(\theta_i)} L_n(\theta)$$

$$\geq \min_{1 \leq i \leq k} \frac{1}{n} \sum_{t=1}^n \min_{\theta \in \mathbb{B}_\eta(\theta_i)} \sigma_t^{-1}\ell_t(\theta)$$

$$\geq \min_{1 \leq i \leq k} E[\min_{\theta \in \mathbb{B}_\eta(\theta_i)} \sigma_t^{-1}\ell_t(\theta)] - \epsilon. \quad (S.69)$$

Moreover, for each $\theta_i \in V^c$, by Claim (iii), there exists an $\epsilon_0 > 0$ such that

$$E[\min_{\theta \in \mathbb{B}_\eta(\theta_i)} \sigma_t^{-1}\ell_t(\theta)] \geq E[\sigma_t^{-1}\ell_t(\theta_0)] + 3\epsilon_0. \quad (S.70)$$

Therefore, by (S.69) and (S.70), taking $\epsilon = \epsilon_0$, it holds that

$$\inf_{\theta \in V^c} L_n(\theta) \geq E[\sigma_t^{-1}\ell_t(\theta_0)] + 2\epsilon_0. \quad (S.71)$$

Furthermore, by the ergodic theorem, it follows that

$$\inf_{\theta \in V} L_n(\theta) \leq L_n(\theta_0) = \frac{1}{n} \sum_{t=1}^n \sigma_t^{-1}\ell_t(\theta_0) \leq E[\sigma_t^{-1}\ell_t(\theta_0)] + \epsilon_0. \quad (S.72)$$

Combining (S.71) and (S.72), we have

$$\inf_{\theta \in V^c} L_n(\theta) \geq E[\sigma_t^{-1}\ell_t(\theta_0)] + 2\epsilon_0 > E[\sigma_t^{-1}\ell_t(\theta_0)] + \epsilon_0 \geq \inf_{\theta \in V} L_n(\theta), \quad (S.73)$$
which together with Claim (i), implies that

\[ \hat{\theta}_{\tau n} \in V \text{ in probability for } \forall V, \text{ as } n \text{ is large enough.} \]

By the arbitrariness of \( V \), it implies that \( \hat{\theta}_{\tau n} \rightarrow \theta_{\tau 0} \) in probability. The proof of this theorem is complete.

**Proof of Theorem 2.** Denote \( \hat{u}_n = \hat{\theta}_{\tau n} - \theta_{\tau 0} \). From Theorem 1, we have \( \hat{u}_n = o_p(1) \). Since \( \hat{\theta}_{\tau n} \) minimizes \( \hat{L}_n(\theta) \), then \( \hat{u}_n \) is the minimizer of \( \hat{H}_n(u) = n[\hat{L}_n(\theta_{\tau 0} + u) - \hat{L}_n(\theta_{\tau 0})] \).

Define \( J = f_\varepsilon(b_\tau)^{-1}\Sigma(\tau)/2 \). By Lemma S.1 and the ergodic theorem, we have \( J_n = J + o_p(1) \).

Moreover, by Lemmas S.4 and S.5, it follows that

\[
\hat{H}_n(\hat{u}_n) = -\sqrt{n}\hat{u}_n'T_n + \sqrt{n}\hat{u}_n'J\sqrt{n}\hat{u}_n + o_p(\sqrt{n}\|\hat{u}_n\| + n\|\hat{u}_n\|^2) \tag{S.74}
\]

\[
\geq -\sqrt{n}\|\hat{u}_n\|[\|T_n\| + o_p(1)] + n\|\hat{u}_n\|^2[\lambda_{\text{min}} + o_p(1)],
\]

where \( \lambda_{\text{min}} \) is the smallest eigenvalue of \( J \). Note that, as \( n \rightarrow \infty \), \( T_n \) converges in distribution to a normal random variable with mean zero and variance matrix \( \tau(1 - \tau)\Sigma(\tau) \).

Since \( \hat{H}_n(\hat{u}_n) \leq 0 \), by (S.74), it holds that

\[
\sqrt{n}\|\hat{u}_n\| \leq [\lambda_{\text{min}} + o_p(1)]^{-1}[\|T_n\| + o_p(1)] = O_p(1). \tag{S.75}
\]

This together with Theorem 1, verifies the root-\( n \) consistency of \( \hat{\theta}_{\tau n} \) in probability.

Let \( \sqrt{n}u_n = J^{-1}T_n/2 = f_\varepsilon^{-1}(b_\tau)\Sigma^{-1}(\tau)T_n \), then we have

\[
\sqrt{n}u_n \rightarrow N\left(0, \frac{\tau(1 - \tau)}{f_\varepsilon^2(b_\tau)\Sigma^{-1}(\tau)}\right)
\]

in distribution as \( n \rightarrow \infty \). Therefore, it suffices to show that \( \sqrt{n}u_n - \sqrt{n}\hat{u}_n = o_p(1) \). By
(S.74) and (S.75), we have

\[
\hat{H}_n(\hat{u}_n) = -\sqrt{n}\hat{\mu}'_n T_n + \sqrt{n}\hat{\mu}'_n J_n \sqrt{n}\hat{\mu}_n + o_p(1)
\]

\[
= -2\sqrt{n}\hat{\mu}'_n J_n \sqrt{n}\hat{u}_n + \sqrt{n}\hat{\mu}'_n J_n \sqrt{n}\hat{\mu}_n + o_p(1) \quad \text{and} \quad (S.76)
\]

\[
\hat{H}_n(u_n) = -\sqrt{n}u'_n T_n + \sqrt{n}u'_n J_n \sqrt{n}u_n + o_p(1) = -\sqrt{n}u'_n J_n \sqrt{n}u_n + o_p(1). \quad (S.77)
\]

From (S.76) and (S.77), it follows that

\[
\hat{H}_n(\hat{u}_n) - \hat{H}_n(u_n) = (\sqrt{n}\hat{\mu}_n - \sqrt{n}u_n)' J_n(\sqrt{n}\hat{\mu}_n - \sqrt{n}u_n) + o_p(1)
\]

\[
\geq \lambda_{\min} \| \sqrt{n}\hat{\mu}_n - \sqrt{n}u_n \|^2 + o_p(1). \quad (S.78)
\]

Since \( \hat{H}_n(\hat{u}_n) - \hat{H}_n(u_n) = n[\hat{L}_n(\theta_{r0} + \hat{u}_n) - \hat{L}_n(\theta_{r0} + u_n)] \leq 0 \text{ a.s.} \), then (S.78) implies that \( \| \sqrt{n}\hat{\mu}_n - \sqrt{n}u_n \| = o_p(1) \). The proof of this theorem is hence accomplished.

**Proof of Corollary 1.** First, we show the consistency of \( \hat{\theta}_{\text{rn}} \). The proof follows the same lines as that of Theorem 1, while functions \( L(\theta) \), \( \tilde{L}_n(\theta) \) and \( \hat{L}_n(\theta) \) are defined as \( L_n(\theta) = n^{-1} \sum_{t=1}^n \rho_t [ Y_t - q_t(\theta) ] \), \( \tilde{L}_n(\theta) = n^{-1} \sum_{t=1}^n \rho_t [ Y_t - \tilde{q}_t(\theta) ] \) and \( \hat{L}_n(\theta) = n^{-1} \sum_{t=1}^n \rho_t [ Y_t - \tilde{q}_t(\theta) ] \), respectively. Next, we show the root-\( n \) consistency and asymptotic normality of \( \hat{\theta}_{\text{rn}} \) as in Theorem 2, where functions \( L(\theta) \), \( \tilde{L}_n(\theta) \) and \( \hat{L}_n(\theta) \) are defined as previous, the function \( \zeta_n(u) \) in Lemma S.3 is defined as \( \zeta_n(u) = \sum_{t=1}^n q_{1t}(u) \{ \xi_{1t}(u) - E[\xi_{1t}(u)|F_{t-1}] \} \), Lemma S.4 remains unchanged while Lemma S.5 is consequently revised as below.

**Lemma 5’.** Suppose \( E(|u_t|^{2+\delta}) < \infty \) for some \( \delta > 0 \). By Assumptions 1, 3 and 4, we have

\[
n[\tilde{L}_n(\theta) - \hat{L}_n(\theta_{r0})] = -\sqrt{n}(\theta - \theta_{r0})' T_n + \sqrt{n}(\theta - \theta_{r0})' J_n \sqrt{n}(\theta - \theta_{r0})
\]

\[
+ o_p(\sqrt{n}\|\theta - \theta_{r0}\| + n\|\theta - \theta_{r0}\|^2).
\]

32
for $\theta - \theta_{r0} = o_p(1)$, where $\tilde{L}_n(\theta) = n^{-1} \sum_{t=1}^{n} \rho_t [Y_t - \hat{q}_t(\theta)]$.

$$T_n = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \frac{\hat{q}_t(\theta_{r0})}{\partial \theta} \psi_t (\varepsilon_t - b_t)$$ and $$J_n = \frac{f_z(b)}{2n} \sum_{t=1}^{n} \frac{1}{\sigma_t} \frac{\hat{q}_t(\theta_{r0})}{\partial \theta} \frac{\hat{q}_t(\theta_{r0})}{\partial \theta'}.$$

Note that without weights $\sigma_t^{-1}$, additional moment condition on $u_t$ will be needed in some intermediate steps of the proof. Therefore, instead of $E|u_t|^2 < \infty$, higher moment condition, $E|u_t|^{2+\delta} < \infty$ for some $\delta > 0$, is required for the proof of Corollary 1.

\[ \square \]

**Proof of Corollary 2.** Denote $z_t(\gamma) = (u^2_{t-1}, \ldots, u^2_{t-q}, \sigma^2_{t-1}(\gamma), \ldots, \sigma^2_{t-p}(\gamma), v^2_{1,t-1}, \ldots, v^2_{d,t-1})'$ and $\tilde{z}_t(\gamma) = (\tilde{u}^2_{t-1}, \ldots, \tilde{u}^2_{t-q}, \tilde{\sigma}^2_{t-1}(\gamma), \ldots, \tilde{\sigma}^2_{t-p}(\gamma), \tilde{v}^2_{1,t-1}, \ldots, \tilde{v}^2_{d,t-1})'$, where $u_t = u_t(\phi_0)$, $\tilde{u}_t = u_t(\phi_n)$, $\sigma^2_t(\gamma) = 1 + \sum_{i=1}^{q} \alpha_i u^2_{t-i} + \sum_{j=1}^{p} \beta_j \sigma^2_{t-j}(\gamma) + \sum_{k=1}^{d} \pi_k v^2_{k,t-1}$ and $\tilde{\sigma}^2_t(\gamma) = 1 + \sum_{i=1}^{q} \alpha_i \tilde{u}^2_{t-i} + \sum_{j=1}^{p} \beta_j \tilde{\sigma}^2_{t-j}(\gamma) + \sum_{k=1}^{d} \pi_k \tilde{v}^2_{k,t-1}$. Let $\sigma^2_t(\gamma) = 1 + \gamma' z_t(\gamma)$ and $\tilde{\sigma}^2_t(\gamma) = 1 + \gamma' \tilde{z}_t(\gamma)$. Note that $u_t = u_t(\phi_0) = Y_t - \Phi_0 X_{t-1}$ and $\tilde{u}_t = u_t(\phi_n) = Y_t - \Phi_n X_{t-1}$. Let $\gamma_t = (b_t, \gamma')'$ and denote by $\gamma_{r0} = (b_r, \gamma'_0)'$ its true value. Denote by $\Theta' \subset \mathbb{R}^{p+q+d+1}$ the parameter space of $\gamma_t$, which satisfies

$$b \leq |b| \leq \tilde{b}, \sum_{j=1}^{p} \beta_j \leq \rho_0, \quad w \leq \min(\alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p, \pi_1, \ldots, \pi_d) \leq \max(\alpha_1, \ldots, \alpha_q, \beta_1, \ldots, \beta_p, \pi_1, \ldots, \pi_d) \leq \bar{w},$$

where $0 < b < \tilde{b}$, $0 < w < \bar{w}$, $0 < \rho_0 < 1$ and $pw < \rho_0$. Moreover, we assume $\Theta'$ is compact and $\gamma_{r0}$ is an interior of $\Theta'$. Recall that $\tilde{\theta}_{r0} = (\gamma_{r0}, \phi'_n)'$, where $\gamma_{r0} = (\tilde{b}_{r0}, \gamma'_0)'$. For the least square estimator $\tilde{\phi}_n$, by the model assumption in (2.1) and (2.2), we have

$$\sqrt{n}(\tilde{\phi}_n - \phi_0) = \left( \frac{1}{n} \sum_{t=1}^{n} X_{t-1} X'_{t-1} \right)^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_{t-1} \varepsilon_t.$$  \hspace{1cm} (S.79)

Then $\sqrt{n}(\tilde{\phi}_n - \phi_0) = O_p(1)$ and $\sqrt{n}(\tilde{\phi}_n - \phi_0) \rightarrow N(0, \Sigma_{22})$ in distribution as $n \rightarrow \infty$, where $\Sigma_{22} = \omega^* D_{0}^{-1} D_{2} D_{0}^{-1}$ with $\omega^* = \text{var}(\varepsilon_t)$ and $D_i = E(\sigma^2_{i,t-1} X_{t-1} X'_{t-1})$.
In the following proof, we focus on $\gamma_{r\eta}$. First, we verify its consistency. Define

$$L_n(\gamma) = \frac{1}{n} \sum_{t=1}^{n} \ell_t(\gamma) \quad \text{and} \quad \bar{L}_n(\gamma) = \frac{1}{n} \sum_{t=1}^{n} \bar{\ell}_t(\gamma),$$

where $\bar{\ell}_t(\gamma) = \rho_t[Y_t - \phi'_n X_{t-1} - b\bar{\sigma}_t(\gamma)]$ and $\ell_t(\gamma) = \rho_t[Y_t - \phi'_0 X_{t-1} - b\sigma_t(\gamma)]$. To show the consistency, we first verify the following claims:

(i) $\sup_{\Theta'} |\bar{L}_n(\gamma) - L_n(\gamma)| = o_p(1)$;

(ii) $E[\sup_{\Theta'} \ell_t(\gamma)] < \infty$;

(iii) $E[\ell_t(\gamma)]$ has a unique minimum at $\gamma_{r0}$;

(iv) For any $\gamma^\dagger \in \Theta'$, $E[\sup_{\gamma \in B(\gamma^\dagger)} |\ell_t(\gamma) - \ell_t(\gamma^\dagger)|] \to 0$ as $\eta \to 0$, where $B_\eta(\gamma^\dagger) = \{\gamma \in \Theta' : \|\gamma - \gamma^\dagger\| < \eta\}$ is an open neighborhood of $\gamma^\dagger$ with radius $\eta > 0$.

We first prove Claim (i). Note that $\hat{\sigma}_t(\gamma) - \sigma_t(\gamma) = [\hat{\sigma}_t^2(\gamma) - \sigma_t^2(\gamma)]/[\hat{\sigma}_t(\gamma) + \sigma_t(\gamma)]$.

Similar to the proof of Lemma S.2(i), by the Taylor expansion, it can be verified that

$$\hat{\sigma}_t^2(\gamma) - \sigma_t^2(\gamma) = -2(\hat{\phi}_n - \phi'_0)\sum_{i=1}^\infty a_{\gamma}(i)u_{t-i}(\phi^*)X_{t-i-1},$$

where $\phi^*$ is between $\phi_0$ and $\hat{\phi}_n$. By Assumption 1, together with the facts that $\hat{\sigma}_t^2(\gamma) \geq a_{\gamma}(i)\bar{u}^2_{t-i}$ and $\sigma_t^2(\gamma) \geq a_{\gamma}(i)u^2_{t-i}$, we can show that

$$\sup_{\Theta'} \|\hat{\sigma}_t(\gamma) - \sigma_t(\gamma)\| \leq 2\|\hat{\phi}_n - \phi_0\| \sum_{i=1}^\infty a_{\gamma}(i)\|u_{t-i}(\phi^*)\|\|X_{t-i-1}\| \leq 2\|\hat{\phi}_n - \phi_0\| \rho^\ell \xi_p.$$

Then by the Lipschitz continuity of $\rho_t(x)$ and Assumption 1, we have

$$\sup_{\Theta'} |\bar{L}_n(\gamma) - L_n(\gamma)| \leq 2\sum_{t=1}^n \sup_{\Theta'} \|[(\hat{\phi}_n - \phi'_0)X_{t-1} + b[\hat{\sigma}_t(\gamma) - \sigma_t(\gamma)]]\|
\leq 2\|\hat{\phi}_n - \phi_0\| \frac{1}{n} \sum_{t=1}^n \|X_{t-1}\| + C\bar{b} \rho^\ell \xi_p.$$
This together with \( \sqrt{n}(\hat{\phi}_n - \phi_0) = O_p(1) \), implies that Claim (i) holds. Similar to proofs of Claims (ii)-(iv) in the proof of Theorem 1, we can verify Claims (ii)-(iv). Finally, by Claims (i)-(iv) and the similar arguments as in the proof of Theorem 1, we can show that
\[
\tilde{\gamma}_{tn} - \gamma_{t0} = o_p(1).
\]

We next prove the asymptotic normality of \( \tilde{\gamma}_{tn} \). We use the same technical tools as in the proof of Theorem 2. We only sketch the key steps as below. Denote \( u = \gamma_r - \gamma_{0r} \). Define \( q_t(\gamma_{t0}) = \phi_0'X_{t-1} + b_r \sigma_t(\gamma_0) \) and \( \hat{q}_t(\gamma_r) = \hat{\phi}_0'X_{t-1} + b\hat{\sigma}_t(\gamma) \). Denote \( \nu_t(u) = q_t(\gamma_r) - \hat{q}_t(\gamma_{t0}) \). This together with \( \tilde{u}_t = Y_t - \hat{\phi}_n'X_{t-1} \) and \( u_t = Y_t - \phi_0'X_{t-1} \), implies that \( \nu_t(u) = -(\tilde{u}_t - u_t) + [b\hat{\sigma}_t(\gamma) - b_r \sigma_t(\gamma_0)] \). Note that \( \sigma_t = \sigma_t(\gamma_0) \) and \( u_t = \sigma_t \varepsilon_t \), then \( u_t - b_r \sigma_t(\gamma_0) = \sigma_t(\varepsilon_t - b_r) \) and hence \( \psi_t(u_t - b_r \sigma_t(\gamma_0)) = \psi_t(\varepsilon_t - b_r) \). This together with the Knight equation (S.66), we can show that
\[
n[u(L_n(\gamma_r) - L_n(\gamma_{t0}))] = R_{1n}(u) + R_{2n}(u), \tag{S.80}
\]
where \( R_{1n}(u) = - \sum_{t=1}^n \nu_t(u) \psi_t(\varepsilon_t - b_r) \) and
\[
R_{2n}(u) = \sum_{t=1}^n \nu_t(u) \int_0^1 [I(\varepsilon_t \leq b_r + \sigma_t^{-1} \nu_t(u)s) - I(\varepsilon_t \leq b_r)] ds.
\]

Since \( \tilde{u}_t - u_t = (\hat{\phi}_n - \phi_0)'X_{t-1} \), it can be verified that \( \nu_t(u) = q_{1t}(u) + q_{2t}(u) \), where
\[
q_{1t}(u) = (\hat{\phi}_n - \phi_0)'X_{t-1} + b_r[\hat{\sigma}_t(\alpha_0) - \sigma_t(\gamma_0)] + (b - b_r)\sigma_t + b_r[\sigma_t(\gamma) - \sigma_t(\gamma_0)],
\]
\[
q_{2t}(u) = b_r[\hat{\sigma}_t(\gamma) - \sigma_t(\gamma)] - [\hat{\sigma}_t(\gamma_0) - \sigma_t(\gamma_0)] + (b - b_r)[\sigma_t(\gamma) - \sigma_t(\gamma_0)] + (b - b_r)[\hat{\sigma}_t(\gamma_0) - \sigma_t(\gamma_0)].
\]

Recall that \( M_t = X_{t-1} + 0.5\sigma_t^{-1}b_r \hat{\sigma}_t^2(\lambda_0)/\hat{\phi}' \) and \( W_t = (\sigma_t, 0.5\sigma_t^{-1}b_r \hat{\sigma}_t^2(\lambda_0)/\hat{\gamma}''). \)

Following the lines in the proof of Theorem 2, if Assumptions 1, 3 and 4 hold and \( E|u_t|^{2+\delta} < \infty \)
\( \infty \) for some \( \delta > 0 \), we can show that

\[
R_{1n}(u) = -\sqrt{n}u' \frac{1}{\sqrt{n}} \sum_{t=1}^{n} W_t \psi_{\tau}(\varepsilon_t - b_t) - T_{2n} + o_p(n\|u\|^2), \quad \text{with} \quad (S.81)
\]

\[
T_{2n} = \sqrt{n}(\hat{\phi}_n - \phi_0)' \frac{1}{\sqrt{n}} \sum_{t=1}^{n} M_t \psi_{\tau}(\varepsilon_t - b_t).
\]

For \( R_{2n}(u) \), we have

\[
R_{2n}(u) = \sqrt{n}u' \frac{f_\varepsilon(b_r)}{2} \frac{1}{n} \sum_{t=1}^{n} \sigma_t^{-1} W_t W'_t \sqrt{n}u
\]

\[
+ \sqrt{n}u' f_\varepsilon(b_r) \frac{1}{n} \sum_{t=1}^{n} \sigma_t^{-1} W_t M'_t \sqrt{n}(\hat{\phi}_n - \phi_0)
\]

\[
+ T_{3n} + o_p(\sqrt{n}\|u\| + n\|u\|^2), \quad \text{with} \quad (S.82)
\]

\[
T_{3n} = \sqrt{n}(\hat{\phi}_n - \phi_0)' \frac{f_\varepsilon(b_r)}{2} \frac{1}{n} \sum_{t=1}^{n} \sigma_t^{-1} M_t M'_t \sqrt{n}(\hat{\phi}_n - \phi_0).
\]

Hence, by (S.80)-(S.82), we have

\[
n[\hat{L}_n(\gamma_{\tau}) - L_n(\gamma_{\tau_0})]
\]

\[
= -\sqrt{n}u' \left[ \frac{1}{\sqrt{n}} \sum_{t=1}^{n} W_t \psi_{\tau}(\varepsilon_t, \tau) - f_\varepsilon(b_r) \frac{1}{n} \sum_{t=1}^{n} \sigma_t^{-1} W_t M'_t \sqrt{n}(\hat{\phi}_n - \phi_0) \right]
\]

\[
+ \sqrt{n}u' f_\varepsilon(b_r) \frac{1}{2n} \sum_{t=1}^{n} \sigma_t^{-1} W_t W'_t \sqrt{n}u - T_{2n} + T_{3n} + o_p(\sqrt{n}\|u\| + n\|u\|^2).
\]

By the consistency, we have \( \gamma_{\tau n} - \gamma_{\tau_0} = o_p(1) \). Moreover, \( \gamma_{\tau n} \) is the minimizer of \( \hat{L}_n(\gamma_{\tau}) \).

Then it follows that

\[
\sqrt{n}(\gamma_{\tau n} - \gamma_{\tau_0}) = \frac{\Omega_{-1}}{f_\varepsilon(b_r)} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} W_t \psi_{\tau}(\varepsilon_t - b_t) - \Omega_{-1}^{-1} \Gamma_1 \sqrt{n}(\hat{\phi}_n - \phi_0) + o_p(1),
\]

where \( \Omega_1 = E(\sigma_t^{-1} W_t W'_t) \) and \( \Gamma_1 = E(\sigma_t^{-1} W_t M'_t) \). Similar to the proof of Theorem 2, we can verify that \( \sqrt{n}(\gamma_{\tau n} - \gamma_{\tau_0}) = O_p(1) \) and \( \sqrt{n}(\gamma_{\tau n} - \gamma_{\tau_0}) \to N(0, \Sigma_{11}(\tau)) \) in distribution as \( n \to \infty \), where \( \Sigma_{11}(\tau) \) is defined in Section 2.2. This together with (S.79), we complete the proof by the central limit theorem and the CraméR-Wold device. \( \square \)
Proof of Corollary 3. First, consider the conditional quantile \( q_{n+1}(\hat{\theta}_{\tau n}) \) for the jointly weighted estimator \( \hat{\theta}_{\tau n} = (\hat{\gamma}'_{\tau n}, \hat{\phi}_n)' \), where \( \hat{\gamma}_{\tau n} = (\hat{b}_{\tau n}, \hat{\gamma}_n)' \). By Theorem 2, we have \( \sqrt{n}(\hat{\theta}_{\tau n} - \theta_{\tau 0}) = O_p(1) \). Then by the Taylor expansion and Lemma S.1, given \( F_n \), it follows that

\[
q_{n+1}(\hat{\theta}_{\tau n}) - q_{n+1}(\theta_{\tau 0}) = \frac{\partial q_{n+1}(\theta_{\tau 0})}{\partial \theta_{\tau 0}} (\hat{\theta}_{\tau n} - \theta_{\tau 0}) + \frac{1}{2} \frac{\partial^2 q_{n+1}(\theta)}{\partial \theta \partial \theta'} (\hat{\theta}_{\tau n} - \theta_{\tau 0})^2 + o_p(n^{-1/2}).
\]

Similarly, by Corollary 1 and the Taylor expansion, we can verify the representation of \( q_{n+1}(\hat{\theta}_{\tau n}) \) for the jointly weighted estimator \( \hat{\theta}_{\tau n} = (\hat{\gamma}'_{\tau n}, \hat{\phi}_n)' \), where \( \hat{\gamma}_{\tau n} = (\hat{b}_{\tau n}, \hat{\gamma}_n)' \).

Finally, consider \( q_{n+1}(\hat{\theta}_{\tau n}) \) for the two-step estimator \( \hat{\theta}_{\tau n} = (\hat{\gamma}'_{\tau n}, \hat{\phi}_n)' \), where \( \hat{\gamma}_{\tau n} = (\hat{b}_{\tau n}, \hat{\gamma}_n)' \). Recall that \( \tilde{\alpha}_t^2(\gamma) = 1 + \gamma' \tilde{z}_t(\gamma) \) and \( \tilde{\alpha}_t^2(\gamma) = 1 + \gamma' \tilde{z}_t(\gamma) \) in the proof of Corollary 2. Note that \( \sigma_t = \tilde{\alpha}_t(\gamma_0) = \sigma_t(\lambda_0) \). By Corollary 2, we have \( \sqrt{n}(\gamma_{\tau n} - \gamma_{\tau 0}) = O_p(1) \) and \( \sqrt{n}(\phi_n - \phi_0) = O_p(1) \). Then by the Taylor expansion, Lemma S.1 and the proof of Corollary 2, given \( F_n \), it holds that

\[
q_{n+1}(\hat{\theta}_{\tau n}) - q_{n+1}(\theta_{\tau 0}) = (\hat{\phi}_n - \phi_0)' X_n + [\hat{b}_{\tau n} \tilde{\alpha}_{n+1}(\hat{\gamma}_n) - b_\tau \tilde{\alpha}_{n+1}(\gamma_0)] + (\hat{\phi}_n - \phi_0)' X_n + o_p(n^{-1/2})
\]

where \( M_t = X_{t-1} + 0.5 \sigma_t^{-1} b_\tau \tilde{\alpha}_t^2(\lambda_0)/\tilde{\alpha}_t \) and \( W_t = (\sigma_t, 0.5 \sigma_t^{-1} b_\tau \tilde{\alpha}_t^2(\lambda_0)/\tilde{\alpha}_t)' \). Hence, the proof of this corollary is accomplished.

Proof of Theorem 3. Since the proofs for (ii) and (iii) are similar to the proof of (i), in
below we only provide detailed proof for establishing the bootstrap consistency for the joint
weighed estimator \( \hat{\theta}_{\tau n} \) in (i). Recall that \( \ell_t(\theta) = \rho_t[Y_t - q_t(\theta)] \) and \( \tilde{\ell}_t(\theta) = \rho_t[Y_t - \tilde{q}_t(\theta)] \),
where \( q_t(\theta) = \phi'X_{t-1} + b\sigma_t(\lambda) \) and \( \tilde{q}_t(\theta) = \phi'X_{t-1} + b\tilde{\sigma}_t(\lambda) \). Define the following functions
\[
\hat{L}_n^*(\theta) = \frac{1}{n} \sum_{t=1}^n \omega_t \hat{\sigma}_t^{-1}\tilde{\ell}_t(\theta), \quad \tilde{L}_n^*(\theta) = \frac{1}{n} \sum_{t=1}^n \omega_t \tilde{\sigma}_t^{-1}\ell_t(\theta), \quad L_n^*(\theta) = \frac{1}{n} \sum_{t=1}^n \omega_t \sigma_t^{-1}\ell_t(\theta),
\]
where \( \hat{\sigma}_t = \hat{\sigma}_t(\hat{\lambda}_n) \), \( \tilde{\sigma}_t = \tilde{\sigma}_t(\tilde{\lambda}_n) \) and \( \sigma_t = \sigma_t(\lambda_0) \).

Similar to the consistency proof of Theorem 1 and Lemma A.3 of Zhu, Zeng, and Li (2019), by Assumptions 1, 3-5, we can show that

\[
\hat{\theta}_{\tau n}^* - \theta_{\tau 0} = o_p^*(1). \quad (S.83)
\]

Let \( \zeta_n^*(u) = \sum_{t=1}^n \omega_t \sigma_t^{-1}q_{1t}(u) \{\xi_{1t}(u) - E[\xi_{1t}(u)|F_{t-1}]\} \), where \( q_{1t}(u) \) and \( \xi_{1t}(u) \) are defined as in Lemma S.3. In line with the proof of Lemma S.3, by Assumptions 1, 3-5, for \( u = o_p(1) \), we can show that

\[
\zeta_n^*(u) - \zeta_n(u) = \sum_{t=1}^n (\omega_t - 1)\sigma_t^{-1}q_{1t}(u) \{\xi_{1t}(u) - E[\xi_{1t}(u)|F_{t-1}]\} = o_p^*(\sqrt{n}\|u\| + n\|u\|^2),
\]
which implies that

\[
\zeta_n^*(u) = o_p^*(\sqrt{n}\|u\| + n\|u\|^2). \quad (S.84)
\]

In line with the proof of Lemma S.4, by Assumption 5, for \( \theta - \theta_{\tau 0} = o_p(1) \) we have

\[
n[\hat{L}_n^*(\theta) - \hat{L}_n^*(\theta_{\tau 0})] - n[\tilde{L}_n^*(\theta) - \tilde{L}_n^*(\theta_{\tau 0})] = o_p^*(\sqrt{n}\|\theta - \theta_{\tau 0}\| + n\|\theta - \theta_{\tau 0}\|^2). \quad (S.85)
\]
Similar to the proof of Lemma S.5, together with (S.84), we can verify that

\[ n[\tilde{L}_n^*(\theta) - \tilde{L}_n^*(\theta_{r0})] = -\sqrt{n}(\theta - \theta_{r0})' T_n^* + \sqrt{n}(\theta - \theta_{r0})' J_n \sqrt{n}(\theta - \theta_{r0}) \]

\[ + o_p^*(\sqrt{n} \| \theta - \theta_{r0} \| + n \| \theta - \theta_{r0} \|^2), \tag{S.86} \]

where \( \theta - \theta_{r0} = o_p(1) \) and \( T_n^* = n^{-1/2} \sum_{t=1}^{n} \omega_t \sigma_t^{-1} \hat{c}_t \hat{q}_t(\theta_{r0}) / \hat{c}_t \psi_t(\varepsilon_{t,r}) \). By methods similar to (S.74)-(S.77), together with (S.83), (S.85) and (S.86), it can be shown that

\[ \sqrt{n}(\hat{\theta}_{\tau n}^* - \theta_{r0}) = f_\varepsilon^{-1}(b_\tau) \Sigma^{-1}(\tau) T_n^* + o_p^*(1). \]

This together with \( \sqrt{n}(\hat{\theta}_{\tau n} - \theta_{r0}) = f_\varepsilon^{-1}(b_\tau) \Sigma^{-1}(\tau) T_n + o_p(1) \), implies that

\[ \sqrt{n}(\hat{\theta}_{\tau n}^* - \hat{\theta}_{\tau n}) = f_\varepsilon^{-1}(b_\tau) \Sigma^{-1}(\tau) \frac{1}{\sqrt{n}} \sum_{t=1}^{n} (\omega_t - 1) \frac{1}{\sigma_t} \hat{c}_t (\theta_{r0})^2 / \hat{c}_t \psi_t(\varepsilon_{t,r}) + o_p^*(1). \]

This together with Assumption 5, we complete the proof by applying Lindeberg’s central limit theorem and the Cramér-Wold device.

\[ \square \]

2 Additional simulation results

Tables S.1-S.3 report additional simulation results for the first experiment in Section 3.1 with sample size \( n = 1000 \), which aims to evaluate the finite-sample performance of the three proposed estimators \( \hat{\theta}_{\tau n}, \tilde{\theta}_{\tau n} \) and \( \hat{\theta}_{\tau n} \), and their bootstrapping approximations. Simulation findings are provided in Section 3.1 of the main manuscript.
Table S.1: Biases, ESDs, ASDs and ECRs of 95% confidence intervals for $\hat{\theta}_{\tau n}$ at $\tau = 0.05$ and 0.10, for normally distributed $X_{t-1}$ and $v_{t-1}$. The innovations are normally or Student’s $t_5$ distributed.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$n$</th>
<th>Bias</th>
<th>ESD</th>
<th>ASD</th>
<th>ECR</th>
<th>$\tau$</th>
<th>$n$</th>
<th>Bias</th>
<th>ESD</th>
<th>ASD</th>
<th>ECR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1000</td>
<td>$b_\tau$</td>
<td>-0.227</td>
<td>0.418</td>
<td>0.376</td>
<td>0.942</td>
<td>-0.192</td>
<td>0.361</td>
<td>0.344</td>
<td>0.947</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>-0.012</td>
<td>0.053</td>
<td>0.052</td>
<td>0.909</td>
<td>-0.010</td>
<td>0.062</td>
<td>0.059</td>
<td>0.892</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>-0.041</td>
<td>0.074</td>
<td>0.079</td>
<td>0.952</td>
<td>-0.041</td>
<td>0.080</td>
<td>0.089</td>
<td>0.963</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi$</td>
<td>0.086</td>
<td>0.221</td>
<td>0.201</td>
<td>0.931</td>
<td>0.109</td>
<td>0.270</td>
<td>0.234</td>
<td>0.953</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>0.007</td>
<td>0.251</td>
<td>0.250</td>
<td>0.950</td>
<td>-0.004</td>
<td>0.277</td>
<td>0.287</td>
<td>0.959</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>1000</td>
<td>$b_\tau$</td>
<td>-0.185</td>
<td>0.334</td>
<td>0.347</td>
<td>0.954</td>
<td>-0.135</td>
<td>0.294</td>
<td>0.290</td>
<td>0.965</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>-0.014</td>
<td>0.056</td>
<td>0.053</td>
<td>0.898</td>
<td>-0.010</td>
<td>0.058</td>
<td>0.055</td>
<td>0.884</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>-0.040</td>
<td>0.074</td>
<td>0.089</td>
<td>0.958</td>
<td>-0.038</td>
<td>0.084</td>
<td>0.093</td>
<td>0.968</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi$</td>
<td>0.081</td>
<td>0.234</td>
<td>0.207</td>
<td>0.930</td>
<td>0.089</td>
<td>0.248</td>
<td>0.221</td>
<td>0.933</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>0.001</td>
<td>0.212</td>
<td>0.215</td>
<td>0.951</td>
<td>0.005</td>
<td>0.205</td>
<td>0.212</td>
<td>0.950</td>
<td></td>
</tr>
</tbody>
</table>

Table S.2: Biases, ESDs, ASDs and ECRs of 95% confidence intervals for $\tilde{\theta}_{\tau n}$ at $\tau = 0.05$ and 0.10, for normally distributed $X_{t-1}$ and $v_{t-1}$. The innovations are normally or Student’s $t_5$ distributed.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$n$</th>
<th>Bias</th>
<th>ESD</th>
<th>ASD</th>
<th>ECR</th>
<th>$\tau$</th>
<th>$n$</th>
<th>Bias</th>
<th>ESD</th>
<th>ASD</th>
<th>ECR</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1000</td>
<td>$b_\tau$</td>
<td>-0.134</td>
<td>0.336</td>
<td>0.307</td>
<td>0.952</td>
<td>-0.144</td>
<td>0.341</td>
<td>0.314</td>
<td>0.949</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>-0.009</td>
<td>0.053</td>
<td>0.052</td>
<td>0.927</td>
<td>-0.010</td>
<td>0.064</td>
<td>0.060</td>
<td>0.886</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>-0.024</td>
<td>0.064</td>
<td>0.069</td>
<td>0.953</td>
<td>-0.029</td>
<td>0.077</td>
<td>0.085</td>
<td>0.952</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi$</td>
<td>0.089</td>
<td>0.249</td>
<td>0.234</td>
<td>0.940</td>
<td>0.110</td>
<td>0.296</td>
<td>0.264</td>
<td>0.964</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>0.007</td>
<td>0.259</td>
<td>0.260</td>
<td>0.947</td>
<td>-0.003</td>
<td>0.284</td>
<td>0.301</td>
<td>0.958</td>
<td></td>
</tr>
<tr>
<td>0.10</td>
<td>1000</td>
<td>$b_\tau$</td>
<td>-0.132</td>
<td>0.314</td>
<td>0.294</td>
<td>0.971</td>
<td>-0.102</td>
<td>0.284</td>
<td>0.267</td>
<td>0.961</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>-0.011</td>
<td>0.055</td>
<td>0.054</td>
<td>0.908</td>
<td>-0.008</td>
<td>0.061</td>
<td>0.058</td>
<td>0.878</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>-0.028</td>
<td>0.076</td>
<td>0.079</td>
<td>0.956</td>
<td>-0.028</td>
<td>0.084</td>
<td>0.090</td>
<td>0.963</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi$</td>
<td>0.082</td>
<td>0.256</td>
<td>0.245</td>
<td>0.949</td>
<td>0.089</td>
<td>0.266</td>
<td>0.251</td>
<td>0.946</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>0.004</td>
<td>0.219</td>
<td>0.222</td>
<td>0.947</td>
<td>0.005</td>
<td>0.211</td>
<td>0.219</td>
<td>0.961</td>
<td></td>
</tr>
</tbody>
</table>

40
Table S.3: Biases, ESDs, ASDs and ECRs of 95% confidence intervals for $\hat{\theta}_{\tau n}$ at $\tau = 0.05$ and 0.10, for normally distributed $X_{t-1}$ and $\nu_{t-1}$. The innovations are normally or Student’s $t_5$ distributed.

<table>
<thead>
<tr>
<th>$\tau$</th>
<th>$n$</th>
<th>Normal</th>
<th></th>
<th>t$_5$</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Bias</td>
<td>ESD</td>
<td>ASD</td>
<td>ECR</td>
</tr>
<tr>
<td>0.05</td>
<td>1000</td>
<td>$b_\tau$</td>
<td>-0.131</td>
<td>0.332</td>
<td>0.297</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>-0.010</td>
<td>0.052</td>
<td>0.052</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>-0.023</td>
<td>0.066</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi$</td>
<td>0.086</td>
<td>0.233</td>
<td>0.238</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>0.002</td>
<td>0.151</td>
<td>0.147</td>
</tr>
<tr>
<td>0.10</td>
<td>1000</td>
<td>$b_\tau$</td>
<td>-0.129</td>
<td>0.318</td>
<td>0.288</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha$</td>
<td>-0.011</td>
<td>0.055</td>
<td>0.053</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\beta$</td>
<td>-0.027</td>
<td>0.076</td>
<td>0.078</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi$</td>
<td>0.087</td>
<td>0.256</td>
<td>0.248</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>0.002</td>
<td>0.151</td>
<td>0.147</td>
</tr>
</tbody>
</table>

References


