# Supplementary material for “A new generalized exponentially weighted moving average quantile model and its statistical inference”

Ke Zhu

*University of Hong Kong*

This supplementary material contains two parts. In the first part, it provides some additional simulation results. In the second part, it presents the Appendix, which contains all of the proofs for the paper.

# S Additional simulations

In this section, we give some additional simulation results, when $\varepsilon_t$ follows the standardized Pearson Type IV distribution (denoted by $P(\nu, m)$) such that its $\tau$th quantile is $-1$. Note that the Pearson Type IV distribution has the density given by

$$f(x;\nu,m) = K(1+x^2)^{-m}\exp\big(-\nu\cdot\tan^{-1}(x)\big) \ \text{ with } \ K = \frac{2^{2m-1}|\Gamma(m+i\nu/2)|^2}{\pi\Gamma(2m-1)}$$

for $m > 1/2$, where it is negatively skewed when $\nu > 0$ and it has a heavier tail when the value of $m$ becomes smaller (see Zhu and Li (2015) for more details on this distribution).

## S.1 Simulation studies for $\widehat{\theta}_n$

In this subsection, we examine the finite-sample performance of the weighted quantile estimator $\widehat{\theta}_n$ in (3.4). We generate 1000 replications of sample size $n = 500$ and 1000 from the following model:

$$y_t = q_t\varepsilon_t \text{ and } q_t = \psi_0|y_{t-1}| + 0.9q_{t-1}, \tag{S.1}$$

where $\varepsilon_t$ follows the standardized $P(0.5, 5)$ or $P(0.5, 3)$ such that its $\tau$th quantile is $-1$, the values of $\psi_0$ are taken as the cases of $\gamma_s = 0$ for $\varepsilon_t \sim P(0.5, 5)$ and $P(0.5, 3)$ That is, when $\tau = 0.01$, we take $\psi_0 = 0.3603$ or $0.4144$; and when $\tau = 0.05$, we take $\psi_0 = 0.2392$ or $0.2534$.

Tables S1 and S2 report the sample bias, sample ESD, and ASD of $\widehat{\theta}_n$ based on 1000 replications for $\tau = 0.01$ and 0.05, respectively, where the ASD is calculated based on $\widehat{\Omega}_n$ in (3.5). From these two tables, we can have the similar findings as those in Tables 2 and 3.

Table S1: The results for $\widehat{\theta}_n$, when $\tau = 0.01$.

| | | | $r=2$ | | $r=1$ | | $r=0$ | |
|---|---|---|---|---|---|---|---|---|
| $\varepsilon_t$ | $n$ | | $\widehat{\psi}_n$ | $\widehat{\lambda}_n$ | $\widehat{\psi}_n$ | $\widehat{\lambda}_n$ | $\widehat{\psi}_n$ | $\widehat{\lambda}_n$ |
| | | | Panel A: $(\psi_0, \lambda_0) = (0.3603, 0.9)$ | | | | | |
| $P(0.5, 5)$ | 500 | Bias | 0.0045 | -0.0013 | 0.0064 | -0.0017 | 0.0078 | -0.0020 |
| | | ESD | 0.1310 | 0.0363 | 0.1335 | 0.0367 | 0.1370 | 0.0375 |
| | | ASD | 0.1426 | 0.0372 | 0.1434 | 0.0373 | 0.1436 | 0.0373 |
| | 1000 | Bias | 0.0043 | -0.0012 | 0.0054 | -0.0015 | 0.0058 | -0.0015 |
| | | ESD | 0.1114 | 0.0299 | 0.1130 | 0.0304 | 0.1152 | 0.0309 |
| | | ASD | 0.1092 | 0.0287 | 0.1090 | 0.0286 | 0.1093 | 0.0286 |
| $P(0.5, 3)$ | 500 | Bias | 0.0001 | -0.0001 | 0.0009 | -0.0001 | 0.0025 | -0.0004 |
| | | ESD | 0.1429 | 0.0348 | 0.1465 | 0.0354 | 0.1524 | 0.0365 |
| | | ASD | 0.1596 | 0.0363 | 0.1593 | 0.0362 | 0.1634 | 0.0368 |
| | 1000 | Bias | 0.0072 | -0.0013 | 0.0077 | -0.0014 | 0.0082 | -0.0014 |
| | | ESD | 0.1273 | 0.0303 | 0.1301 | 0.0309 | 0.1335 | 0.0315 |
| | | ASD | 0.1293 | 0.0295 | 0.1291 | 0.0294 | 0.1286 | 0.0293 |
| | | | Panel B: $(\psi_0, \lambda_0) = (0.4144, 0.9)$ | | | | | |
| $P(0.5, 5)$ | 500 | Bias | -0.0094 | 0.0017 | -0.0088 | 0.0017 | -0.0070 | 0.0014 |
| | | ESD | 0.1296 | 0.0367 | 0.1331 | 0.0375 | 0.1396 | 0.0390 |
| | | ASD | 0.1492 | 0.0388 | 0.1494 | 0.0388 | 0.1506 | 0.0390 |
| | 1000 | Bias | -0.0011 | -0.0000 | 0.0003 | -0.0004 | 0.0020 | -0.0008 |
| | | ESD | 0.1113 | 0.0298 | 0.1137 | 0.0304 | 0.1168 | 0.0311 |
| | | ASD | 0.1179 | 0.0307 | 0.1184 | 0.0308 | 0.1188 | 0.0309 |
| $P(0.5, 3)$ | 500 | Bias | 0.0163 | -0.0043 | 0.0200 | -0.0049 | 0.0214 | -0.0050 |
| | | ESD | 0.1848 | 0.0439 | 0.1888 | 0.0444 | 0.1949 | 0.0457 |
| | | ASD | 0.1801 | 0.0405 | 0.1861 | 0.0415 | 0.2821 | 0.0582 |
| | 1000 | Bias | -0.0100 | 0.0021 | -0.0065 | 0.0013 | -0.0058 | 0.0013 |
| | | ESD | 0.1315 | 0.0310 | 0.1344 | 0.0315 | 0.1389 | 0.0324 |
| | | ASD | 0.1388 | 0.0315 | 0.1394 | 0.0316 | 0.1392 | 0.0315 |

*Note*: The distribution of $\varepsilon_t$ is standardized such that its $\tau$th quantile is $-1$.

Table S2: The results for $\widehat{\theta}_n$, when $\tau = 0.05$.

| $\varepsilon_t$ | $n$ | | $r=2$ $\widehat{\psi}_n$ | $r=2$ $\widehat{\lambda}_n$ | $r=1$ $\widehat{\psi}_n$ | $r=1$ $\widehat{\lambda}_n$ | $r=0$ $\widehat{\psi}_n$ | $r=0$ $\widehat{\lambda}_n$ |
|---|---|---|---|---|---|---|---|---|
| | | | | Panel A: $(\psi_0, \lambda_0) = (0.2392, 0.9)$ | | | | |
| $P(0.5, 5)$ | 500 | Bias | 0.0008 | -0.0003 | 0.0003 | -0.0001 | -0.0001 | 0.0001 |
| | | ESD | 0.0781 | 0.0314 | 0.0779 | 0.0314 | 0.0773 | 0.0311 |
| | | ASD | 0.0752 | 0.0298 | 0.0750 | 0.0297 | 0.0750 | 0.0297 |
| | 1000 | Bias | -0.0001 | -0.0001 | -0.0005 | 0.0001 | -0.0012 | 0.0004 |
| | | ESD | 0.0565 | 0.0225 | 0.0567 | 0.0225 | 0.0566 | 0.0225 |
| | | ASD | 0.0553 | 0.0220 | 0.0552 | 0.0220 | 0.0551 | 0.0219 |
| $P(0.5, 3)$ | 500 | Bias | 0.0046 | -0.0018 | 0.0038 | -0.0015 | 0.0034 | -0.0014 |
| | | ESD | 0.0878 | 0.0336 | 0.0881 | 0.0338 | 0.0893 | 0.0343 |
| | | ASD | 0.0807 | 0.0301 | 0.0805 | 0.0300 | 0.0805 | 0.0300 |
| | 1000 | Bias | -0.0006 | 0.0003 | -0.0014 | 0.0006 | -0.0021 | 0.0008 |
| | | ESD | 0.0614 | 0.0232 | 0.0612 | 0.0231 | 0.0618 | 0.0233 |
| | | ASD | 0.0591 | 0.0221 | 0.0590 | 0.0221 | 0.0589 | 0.0220 |
| | | | | Panel B: $(\psi_0, \lambda_0) = (0.2534, 0.9)$ | | | | |
| $P(0.5, 5)$ | 500 | Bias | -0.0011 | 0.0006 | -0.0004 | 0.0003 | -0.0003 | 0.0003 |
| | | ESD | 0.0761 | 0.0310 | 0.0776 | 0.0316 | 0.0788 | 0.0321 |
| | | ASD | 0.0775 | 0.0307 | 0.0776 | 0.0307 | 0.0775 | 0.0307 |
| | 1000 | Bias | 0.0020 | -0.0006 | 0.0020 | -0.0007 | 0.0016 | -0.0005 |
| | | ESD | 0.0572 | 0.0228 | 0.0574 | 0.0229 | 0.0582 | 0.0232 |
| | | ASD | 0.0567 | 0.0224 | 0.0567 | 0.0224 | 0.0567 | 0.0224 |
| $P(0.5, 3)$ | 500 | Bias | -0.0002 | 0.0004 | -0.0004 | 0.0005 | -0.0003 | 0.0005 |
| | | ESD | 0.0909 | 0.0340 | 0.0914 | 0.0341 | 0.0923 | 0.0344 |
| | | ASD | 0.0829 | 0.0307 | 0.0829 | 0.0307 | 0.0828 | 0.0307 |
| | 1000 | Bias | 0.0005 | -0.0001 | 0.0001 | 0.0001 | -0.0000 | 0.0001 |
| | | ESD | 0.0649 | 0.0244 | 0.0655 | 0.0246 | 0.0654 | 0.0245 |
| | | ASD | 0.0623 | 0.0232 | 0.0622 | 0.0231 | 0.0622 | 0.0231 |

*Note*: As in Table S1.

## S.2 Simulation studies for $S_n$ and $M_n$

In this subsection, we examine the finite-sample performance of the stability test $S_n$ in (4.2) and the mean invariance test $M_n$ in (4.4). We generate 1000 replications of sample size $n = 1000$ and 2000 from the following model:

$$y_t = q_t \varepsilon_t \text{ and } q_t = (\psi_0 + \zeta)|y_{t-1}| + 0.9 q_{t-1}, \tag{S.2}$$

where $\varepsilon_t$ is chosen as in (S.1), $\zeta \in \{-0.05, ..., -0.01, 0, 0.01, ..., 0.05\}$, and the values of $\psi_0$ are taken with respect to $\gamma_s = 0$ (or $\gamma_m = 1$) for $S_n$ (or $M_n$) so that $q_t$ in model (S.2) is stable for $S_n$ or mean-invariant for $M_n$ when $\zeta = 0$. Specifically, when $\varepsilon_t \sim P(0.5, 5)$ and $\tau = 0.01$, we take $\psi_0 = 0.3603$ for $S_n$ and $\psi_0 = 0.3484$ for $M_n$; when $\varepsilon_t \sim P(0.5, 5)$ and $\tau = 0.05$, we take $\psi_0 = 0.2392$ for $S_n$ and $\psi_0 = 0.2312$ for $M_n$; when $\varepsilon_t \sim P(0.5, 3)$ and $\tau = 0.01$, we take $\psi_0 = 0.4144$ for $S_n$ and $\psi_0 = 0.3980$ for $M_n$; and when $\varepsilon_t \sim P(0.5, 3)$ and $\tau = 0.05$, we take $\psi_0 = 0.2534$ for $S_n$ and $\psi_0 = 2434$ for $M_n$.

Since the power of $S_n$ and $M_n$ is invariant to the choice of $r$ due to the adaptiveness property, we only plot the power of $S_n$ and $M_n$ for $r = 2$ in Figures S1 and S2, respectively, where the sizes of $S_n$ and $M_n$ are corresponding to the cases of $\zeta = 0$. Clearly, we can obtain the similar findings from Figures S1 and S2 as those from Figures 2 and 3.

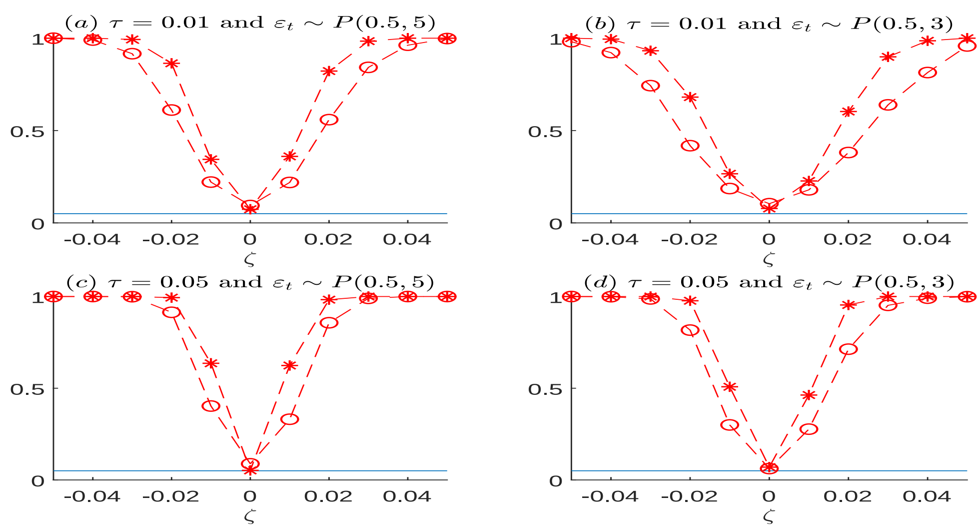


Figure S1: The power of $S_n$ across $\zeta$ in model (S.2), where $n$ is 1000 (dashed circle line) or 2000 (dashed star line). Here, the solid line stands for the significance level $\alpha = 5\%$.

## S.3 Simulation studies for $U_n$

In this subsection, we examine the finite-sample performance of the unit root test $U_n$ in (4.6). We generate 1000 replications of sample size $n = 1000$ and 2000 from the following model:

$$y_t = q_t \varepsilon_t \text{ and } q_t = \omega_0 + \psi_0 |y_{t-1}| + 0.9 q_{t-1}, \tag{S.3}$$

where $\varepsilon_t$ is chosen as in (S.1), $\omega_0 \in \{0, 10^{-3}, 10^{-2}\}$, and three different values of $\psi_0$ are taken for the cases of $\gamma_s < 0$, $\gamma_s = 0$, and $\gamma_s > 0$, respectively. For each replication, we apply $U_n$ to

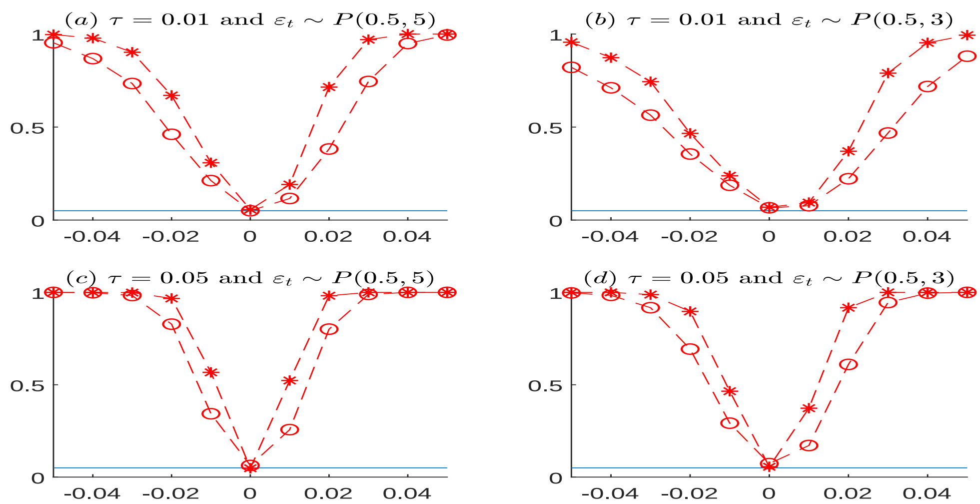


Figure S2: The power of $M_n$ across $\zeta$ in model (S.2), where $n$ is 1000 (dashed circle line) or 2000 (dashed star line). Here, the solid line stands for the significance level $\alpha = 5\%$.

detect whether $\omega_0 = 0$ and $\gamma_s = 0$. Table S3 reports the power of $U_n$ at the significance level of 5%, where the sizes of $U_n$ are corresponding to the cases that $\omega_0 = 0$ and $\gamma_s = 0$. From this table, we can get the similar findings as those in Table 4.

## S.4 Simulation studies for $D_{n,k}^{(p)}$

In this subsection, we examine the finite-sample performance of the dynamic quantile test $D_{n,k}^{(p)}$ in (5.1). We generate 1000 replications of sample size $n = 1000$ and 2000 from the following model:

$$y_t = q_t \varepsilon_t \text{ and } q_t = \psi_0 |y_{t-1}| + \zeta |y_{t-2}| + 0.9 q_{t-1}, \tag{S.4}$$

where the settings of $\varepsilon_t$ and $\psi_0$ are the same as those for $S_n$ above, and $\zeta \in \{0, 0.2, 0.4, 0.6, 0.8\}$. For each replication, we fit it by using the GEWMA quantile model, and then apply $D_{n,k}^{(p)}$ to check whether the fitted model is adequate. Below, we consider the cases that $k = 2$, 4, and 6 with $p = 1.6$ as in the paper, and only report the testing results for $r = 2$ based on 1000 replications, since the performance of $D_{n,k}^{(p)}$ is invariant to the choice of $r$. Table S4 reports the power of three tests $D_{n,2}^{(1.6)}$, $D_{n,4}^{(1.6)}$, and $D_{n,6}^{(1.6)}$, where the sizes correspond to the results for $\zeta = 0$. From this table, we can have the similar findings as those in Table 5.

Overall, our additional simulation studies in this supplement show that our proposed estimator and tests still perform well even when the model innovation is skewed.

Table S3: The power of $U_n$ across $\omega_0$, based on model (S.3).

| | | | | $r=2$ | | | $r=1$ | | | $r=0$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | $\omega_0$ | | | $\omega_0$ | | | $\omega_0$ | | |
| $\tau$ | $\varepsilon_t$ | $n$ | $\psi_0$ | 0 | $10^{-3}$ | $10^{-2}$ | 0 | $10^{-3}$ | $10^{-2}$ | 0 | $10^{-3}$ | $10^{-2}$ |
| 0.01 | $P(0.5,5)$ | 1000 | 0.3400 | 0.5080 | 0.2350 | 0.8140 | 0.5040 | 0.2310 | 0.8110 | 0.5020 | 0.2290 | 0.8130 |
| | | | 0.3603 | **0.0960** | 0.5860 | 0.1490 | **0.0950** | 0.5790 | 0.1520 | **0.0940** | 0.5820 | 0.1500 |
| | | | 0.3800 | 0.1490 | 0.1780 | 0.1180 | 0.1530 | 0.1770 | 0.1170 | 0.1450 | 0.1770 | 0.1100 |
| | | 2000 | 0.3400 | 0.7610 | 0.1980 | 0.8330 | 0.7580 | 0.1980 | 0.8330 | 0.7560 | 0.2000 | 0.8320 |
| | | | 0.3603 | **0.0590** | 0.5040 | 0.0560 | **0.0580** | 0.5040 | 0.0570 | **0.0580** | 0.5020 | 0.0570 |
| | | | 0.3800 | 0.2460 | 0.3820 | 0.4070 | 0.2450 | 0.3830 | 0.3980 | 0.2390 | 0.3740 | 0.3870 |
| | $P(0.5,3)$ | 1000 | 0.3900 | 0.4600 | 0.1900 | 0.7820 | 0.4490 | 0.2000 | 0.7810 | 0.4400 | 0.1900 | 0.7780 |
| | | | 0.4144 | **0.1090** | 0.4980 | 0.1720 | **0.1060** | 0.4950 | 0.1720 | **0.1100** | 0.4890 | 0.1680 |
| | | | 0.4300 | 0.1510 | 0.3020 | 0.0850 | 0.1510 | 0.2970 | 0.0810 | 0.1510 | 0.2910 | 0.0800 |
| | | 2000 | 0.3900 | 0.7420 | 0.1640 | 0.8010 | 0.7390 | 0.1600 | 0.8010 | 0.7290 | 0.1590 | 0.8010 |
| | | | 0.4144 | **0.0590** | 0.4070 | 0.0530 | **0.0610** | 0.4080 | 0.0530 | **0.0610** | 0.4080 | 0.0530 |
| | | | 0.4300 | 0.1060 | 0.1890 | 0.1720 | 0.1080 | 0.1880 | 0.1710 | 0.1040 | 0.1870 | 0.1630 |
| 0.05 | $P(0.5,5)$ | 1000 | 0.2200 | 0.8080 | 0.1440 | 0.9770 | 0.8040 | 0.1450 | 0.9770 | 0.8020 | 0.1410 | 0.9770 |
| | | | 0.2392 | **0.0750** | 0.6690 | 0.2110 | **0.0750** | 0.6690 | 0.2130 | **0.0750** | 0.6680 | 0.2120 |
| | | | 0.2600 | 0.1670 | 0.2510 | 0.2590 | 0.1680 | 0.2500 | 0.2570 | 0.1680 | 0.2480 | 0.2520 |
| | | 2000 | 0.2200 | 0.9780 | 0.0480 | 0.9930 | 0.9780 | 0.0470 | 0.9930 | 0.9790 | 0.0510 | 0.9930 |
| | | | 0.2392 | **0.0480** | 0.5600 | 0.0900 | **0.0480** | 0.5580 | 0.0900 | **0.0470** | 0.5570 | 0.0900 |
| | | | 0.2600 | 0.6810 | 0.7950 | 0.8130 | 0.6800 | 0.7930 | 0.8140 | 0.6780 | 0.7930 | 0.8120 |
| | $P(0.5,3)$ | 1000 | 0.2300 | 0.8450 | 0.0740 | 0.9780 | 0.8450 | 0.0750 | 0.9780 | 0.8460 | 0.0780 | 0.9780 |
| | | | 0.2534 | **0.0760** | 0.5950 | 0.2370 | **0.0740** | 0.5960 | 0.2360 | **0.0740** | 0.5970 | 0.2350 |
| | | | 0.2700 | 0.1140 | 0.1790 | 0.0990 | 0.1130 | 0.1790 | 0.0990 | 0.1110 | 0.1790 | 0.0990 |
| | | 2000 | 0.2300 | 0.9900 | 0.0180 | 0.9940 | 0.9900 | 0.0180 | 0.9940 | 0.9900 | 0.0180 | 0.9940 |
| | | | 0.2534 | **0.0580** | 0.5320 | 0.0740 | **0.0580** | 0.5340 | 0.0740 | **0.0570** | 0.5330 | 0.0740 |
| | | | 0.2700 | 0.2680 | 0.4190 | 0.4370 | 0.2690 | 0.4190 | 0.4360 | 0.2680 | 0.4210 | 0.4350 |

*Note*: The size of $U_n$ is in boldface.

Table S4: The power of $D_{n,k}^{(p)}$ for $p = 1.6$, based on model (S.4).

| | | | | Model (S.4) | | | | |
|---|---|---|---|---|---|---|---|---|
| $\tau$ | $\varepsilon_t$ | Tests | $n$ | $\zeta = 0$ | $\zeta = 0.2$ | $\zeta = 0.4$ | $\zeta = 0.6$ | $\zeta = 0.8$ |
| 0.01 | $P(0.5, 5)$ | $D_{n,2}^{(p)}$ | 1000 | 0.0550 | 0.0570 | 0.0510 | 0.0680 | 0.0740 |
| | | | 2000 | 0.0650 | 0.0550 | 0.0960 | 0.1280 | 0.1810 |
| | | $D_{n,4}^{(p)}$ | 1000 | 0.0870 | 0.0690 | 0.0980 | 0.1070 | 0.1240 |
| | | | 2000 | 0.0670 | 0.0720 | 0.0980 | 0.1450 | 0.1940 |
| | | $D_{n,6}^{(p)}$ | 1000 | 0.1000 | 0.1020 | 0.1080 | 0.1140 | 0.1250 |
| | | | 2000 | 0.0720 | 0.0800 | 0.1190 | 0.1510 | 0.1730 |
| | $P(0.5, 3)$ | $D_{n,2}^{(p)}$ | 1000 | 0.0440 | 0.0410 | 0.0580 | 0.0590 | 0.0680 |
| | | | 2000 | 0.0560 | 0.0510 | 0.0620 | 0.0790 | 0.1010 |
| | | $D_{n,4}^{(p)}$ | 1000 | 0.0780 | 0.0760 | 0.0780 | 0.0800 | 0.1070 |
| | | | 2000 | 0.0590 | 0.0780 | 0.0880 | 0.0960 | 0.1220 |
| | | $D_{n,6}^{(p)}$ | 1000 | 0.0870 | 0.0970 | 0.0930 | 0.1360 | 0.1410 |
| | | | 2000 | 0.0900 | 0.0880 | 0.1100 | 0.1240 | 0.1440 |
| 0.05 | $P(0.5, 5)$ | $D_{n,2}^{(p)}$ | 1000 | 0.0450 | 0.0900 | 0.2020 | 0.3540 | 0.4960 |
| | | | 2000 | 0.0530 | 0.1710 | 0.4440 | 0.7140 | 0.8540 |
| | | $D_{n,4}^{(p)}$ | 1000 | 0.0700 | 0.1040 | 0.1800 | 0.2830 | 0.4260 |
| | | | 2000 | 0.0460 | 0.1410 | 0.3520 | 0.5850 | 0.7530 |
| | | $D_{n,6}^{(p)}$ | 1000 | 0.0840 | 0.1090 | 0.2010 | 0.2650 | 0.3880 |
| | | | 2000 | 0.0630 | 0.1350 | 0.3120 | 0.5120 | 0.7140 |
| | $P(0.5, 3)$ | $D_{n,2}^{(p)}$ | 1000 | 0.0570 | 0.0810 | 0.1320 | 0.2160 | 0.3430 |
| | | | 2000 | 0.0540 | 0.1270 | 0.2630 | 0.4760 | 0.6460 |
| | | $D_{n,4}^{(p)}$ | 1000 | 0.0760 | 0.0880 | 0.1470 | 0.2020 | 0.2610 |
| | | | 2000 | 0.0690 | 0.0970 | 0.2280 | 0.4230 | 0.5140 |
| | | $D_{n,6}^{(p)}$ | 1000 | 0.0900 | 0.1110 | 0.1610 | 0.1810 | 0.2700 |
| | | | 2000 | 0.0680 | 0.1180 | 0.2230 | 0.3510 | 0.4760 |

# Appendix: Proofs

Define five $[0,\infty]$-valued processes

$$
\begin{aligned}
v_t(\theta) &= \sum_{i=1}^{\infty} \frac{\psi|\varepsilon_{t-i}|}{\lambda_0+\psi_0|\varepsilon_{t-i}|} \prod_{j=1}^{i-1} \frac{\lambda}{\lambda_0+\psi_0|\varepsilon_{t-j}|}, \\
d_t^{\psi}(\theta) &= \sum_{i=1}^{\infty} \frac{|\varepsilon_{t-i}|}{\lambda_0+\psi_0|\varepsilon_{t-i}|} \prod_{j=1}^{i-1} \frac{\lambda}{\lambda_0+\psi_0|\varepsilon_{t-j}|}, \\
d_t^{\lambda}(\theta) &= \sum_{i=2}^{\infty} \frac{(i-1)\psi|\varepsilon_{t-i}|}{\lambda(\lambda_0+\psi_0|\varepsilon_{t-i}|)} \prod_{j=1}^{i-1} \frac{\lambda}{\lambda_0+\psi_0|\varepsilon_{t-j}|}, \\
d_t^{\psi\lambda}(\theta) &= \sum_{i=2}^{\infty} \frac{(i-1)|\varepsilon_{t-i}|}{\lambda(\lambda_0+\psi_0|\varepsilon_{t-i}|)} \prod_{j=1}^{i-1} \frac{\lambda}{\lambda_0+\psi_0|\varepsilon_{t-j}|}, \\
d_t^{\lambda\lambda}(\theta) &= \sum_{i=3}^{\infty} \frac{(i-1)(i-2)\psi|\varepsilon_{t-i}|}{\lambda^2(\lambda_0+\psi_0|\varepsilon_{t-i}|)} \prod_{j=1}^{i-1} \frac{\lambda}{\lambda_0+\psi_0|\varepsilon_{t-j}|}
\end{aligned}
$$

with the convention $\prod_{k=1}^{j-1} = 1$ when $j \le 1$. Let $\Theta_0 = \{\theta \in \Theta : \lambda < e^{\gamma_s}\}$. For any $\theta \in \Theta_0$, by Cauchy root test and Assumptions 3.1 and 3.3, it is not hard to see that $v_t(\theta)$, $1/v_t(\theta)$, $d_t^{\psi}(\theta)$, $d_t^{\lambda}(\theta)$, $d_t^{\psi\lambda}(\theta)$, and $d_t^{\lambda\lambda}(\theta)$ are stationary and ergodic with moments of any order.

To facilitate the proofs, we need four technical lemmas. Lemmas A.1 and A.2 are key to our proofs, and they show that normalized by $q_t$, the nonstationary process $Q_t(\theta)$ and its first and second derivatives can be well approximated by some stationary processes. Lemmas A.3 and A.4 are used to prove the consistency and asymptotic normality of $\widehat{\theta}_n$, respectively.

**LEMMA A.1.** *Suppose Assumptions 3.1 and 3.3 hold. Then, there exists a constant $c_0 > 0$ such that, as $t \to \infty$,*

(i) $e^{c_0 t} \sup_{\theta\in\Theta_0} \left| \frac{Q_t(\theta)}{q_t} - v_t(\theta) \right| \xrightarrow{a.s.} 0$;

(ii) $e^{c_0 t} \sup_{\theta\in\Theta_0} \left| \frac{q_t}{Q_t(\theta)} - \frac{1}{v_t(\theta)} \right| \xrightarrow{a.s.} 0$;

(iii) $\frac{Q_t(\theta)}{q_t} \xrightarrow{a.s.} \infty$ *for any* $\theta \notin \Theta_0$.

*Proof.* By (1.1)–(1.2), we have $q_{t-1}/q_t = 1/(\lambda_0+\psi_0|\varepsilon_t|)$. Hence, it follows that

$$
\begin{aligned}
\frac{Q_t(\theta)}{q_t} &= \psi \frac{q_{t-1}}{q_t} \frac{|y_{t-1}|}{q_{t-1}} + \lambda \frac{q_{t-1}}{q_t} \frac{Q_{t-1}(\theta)}{q_{t-1}} \\
&= \frac{\psi|\varepsilon_{t-1}|}{\lambda_0+\psi_0|\varepsilon_{t-1}|} + \frac{\lambda}{\lambda_0+\psi_0|\varepsilon_{t-1}|} \frac{Q_{t-1}(\theta)}{q_{t-1}}
\end{aligned}
$$

$$= \sum_{i=1}^{t} \frac{\psi|\varepsilon_{t-i}|}{\lambda_0 + \psi_0|\varepsilon_{t-i}|} \prod_{j=1}^{i-1} \frac{\lambda}{\lambda_0 + \psi_0|\varepsilon_{t-j}|}.$$

By using the preceding equation, the conclusions follow from the similar arguments as for Lemma A.1 in Li et al. (2018), and hence the details are omitted. □

**LEMMA A.2.** *Suppose Assumptions 3.1 and 3.3 hold. Then, there exists a constant $c_0 > 0$ such that, as $t \to \infty$,*

(i) $e^{c_0 t} \sup_{\theta \in \Theta_0} \left\| \frac{1}{q_t} \frac{\partial Q_t(\theta)}{\partial \theta} - d_t(\theta) \right\| \xrightarrow{a.s.} 0$*;*

(ii) $e^{c_0 t} \sup_{\theta \in \Theta_0} \left\| \frac{1}{q_t} \frac{\partial^2 Q_t(\theta)}{\partial \theta \partial \theta'} - \Sigma_t(\theta) \right\| \xrightarrow{a.s.} 0,$

*where*

$$d_t(\theta) = \begin{pmatrix} d_t^{\psi}(\theta) \\ d_t^{\lambda}(\theta) \end{pmatrix} \quad \text{and} \quad \Sigma_t(\theta) = \begin{pmatrix} 0 & d_t^{\psi\lambda}(\theta) \\ d_t^{\psi\lambda}(\theta) & d_t^{\lambda\lambda}(\theta) \end{pmatrix}.$$

*Proof.* Since $Q_t(\theta) = \psi \sum_{i=1}^{t} \lambda^{i-1}|y_{t-i}|$, we have

$$\frac{\partial Q_t(\theta)}{\partial \psi} = \sum_{i=1}^{t} \lambda^{i-1}|y_{t-i}|, \qquad \frac{\partial Q_t(\theta)}{\partial \lambda} = \psi \sum_{i=2}^{t}(i-1)\lambda^{i-2}|y_{t-i}|,$$

$$\frac{\partial^2 Q_t(\theta)}{\partial \psi^2} = 0, \qquad \frac{\partial^2 Q_t(\theta)}{\partial \psi \partial \lambda} = \sum_{i=2}^{t}(i-1)\lambda^{i-2}|y_{t-i}|,$$

$$\text{and} \quad \frac{\partial^2 Q_t(\theta)}{\partial \lambda^2} = \psi \sum_{i=3}^{t}(i-1)(i-2)\lambda^{i-3}|y_{t-i}|.$$

Then, the conclusion follows from the similar argument as for Lemma A.1. □

**LEMMA A.3.** *Suppose Assumptions 3.1 and 3.3 hold. Then, $\widehat{\theta}_{n,r}$ in (3.3) satisfies*

(i) $\widehat{\theta}_{n,r} \in \Theta_0$ *a.s.;*

(ii) $\sqrt{n}(\widehat{\theta}_{n,r} - \theta_0^*) \xrightarrow{d} N(0, b_r W_r^{-1})$ *as* $n \to \infty$,

*where* $\theta_0^* = (\psi_0^*, \lambda_0)'$ *with* $\psi_0^* = \kappa_r \psi_0$*, and*

$$b_r = \begin{cases} \frac{Var(|\varepsilon_t^*|^r)}{r^2}, & \text{if } r > 0, \\ E[(\log|\varepsilon_t^*|)^2], & \text{if } r = 0, \end{cases} \qquad W_r = \begin{pmatrix} \frac{1}{(\psi_0^*)^2} & \frac{\nu_1}{\psi_0^* \lambda_0 (1-\nu_1)} \\ \frac{\nu_1}{\psi_0^* \lambda_0 (1-\nu_1)} & \frac{(1+\nu_1)\nu_2}{\lambda_0^2 (1-\nu_1)(1-\nu_2)} \end{pmatrix}.$$

*Proof.* The conclusions follow from the similar arguments as for Theorems 3.1–3.2 in Li et al. (2018). Below, we give the proof for the case of $r > 0$, and the case of $r = 0$ can be handled in a similar way.

(i) Since $y_t = q_t^* \varepsilon_t^*$, we can rewrite $\widehat{\theta}_{n,r} = \arg\min_{\theta \in \Theta_r} P_{n,r}(\theta)$, where

$$P_{n,r}(\theta) = \frac{1}{n} \sum_{t=1}^{n} \left\{ |\varepsilon_t^*|^r \left[ \left( \frac{1}{Q_t(\theta)/q_t^*} \right)^r - 1 \right] + r \log\left( \frac{Q_t(\theta)}{q_t^*} \right) \right\}.$$

By Lemma A.1(iii) and the fact that $q_t^* = \kappa_r q_t$, we have $Q_t(\theta)/q_t^* \xrightarrow{a.s.} \infty$ for $\theta \notin \Theta_0$ and large $t$. Therefore, the result (i) follows directly.

(ii) Due to the result (i), it is sufficient to consider the case that $\widehat{\theta}_{n,r} = \arg\min_{\theta\in\Theta_0^*} P_{n,r}(\theta)$, where $\Theta_0^*$ is an arbitrary compact subset of $\Theta_0$. Rewrite $P_{n,r}(\theta) = O_{n,r}(\theta) + R_{n,r}(\theta)$, where

$$O_{n,r}(\theta) = \frac{1}{n}\sum_{t=1}^{n}\left\{|\varepsilon_t^*|^r\left[\Big(\frac{1}{v_t(\theta)/\kappa_r}\Big)^r - 1\right] + r\log\Big(\frac{v_t(\theta)}{\kappa_r}\Big)\right\} \text{ and } R_{n,r}(\theta) = P_{n,r}(\theta) - O_{n,r}(\theta).$$

We first prove that $\widehat{\theta}_{n,r} \xrightarrow{a.s.} \theta_0^*$. Since $E|\varepsilon_t^*|^r = 1$, by the strong law of large numbers for stationary and ergodic sequences, we have

$$O_{n,r}(\theta) \xrightarrow{a.s.} E\Big\{\Big(\frac{1}{v_t(\theta)/\kappa_r}\Big)^r - 1 + r\log\Big(\frac{v_t(\theta)}{\kappa_r}\Big)\Big\} \geq 0,$$

where the equality holds if and only if $v_t(\theta) = \kappa_r$ a.s. or equivalently, $\theta = \theta_0^*$ by using a similar argument as for Lemma A.2 in Francq and Zakoïan (2012). Meanwhile, by Lemma A.1(i)–(ii), it is not hard to see that $\sup_{\theta\in\Theta_0^*}|R_{n,r}(\theta)| \xrightarrow{a.s.} 0$. Hence, by the standard arguments, it follows that $\widehat{\theta}_{n,r} \xrightarrow{a.s.} \theta_0^*$.

Next, by Taylor's expansion we have

$$\sqrt{n}(\widehat{\theta}_{n,r} - \theta_0^*) = -\Big(\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 \ell_{t,r}(\xi)}{\partial\theta\partial\theta'}\Big)^{-1}\Big(\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \ell_{t,r}(\theta_0^*)}{\partial\theta}\Big), \tag{A.1}$$

where $\xi$ lies between $\widehat{\theta}_{n,r}$ and $\theta_0^*$. By direct calculations and Lemmas A.1(i)–(ii) and A.2(i), we can show

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \ell_{t,r}(\theta_0^*)}{\partial\theta} &= \frac{r}{\sqrt{n}}\sum_{t=1}^{n}\frac{(\partial Q_t(\theta_0^*)/\partial\theta)/q_t}{Q_t(\theta_0^*)/q_t}\left[1 - \frac{|\varepsilon_t^*|^r}{(Q_t(\theta_0^*)/q_t^*)^r}\right] \\
&= \frac{r}{\sqrt{n}}\sum_{t=1}^{n}\frac{d_t(\theta_0^*)}{v_t(\theta_0^*)}\left[1 - \frac{|\varepsilon_t^*|^r}{(v_t(\theta_0^*)/\kappa_r)^r}\right] + o_p(1) \\
&= \frac{r}{\sqrt{n}}\sum_{t=1}^{n}\begin{pmatrix} d_t^{\psi}(\theta_0)/\kappa_r \\ d_t^{\lambda}(\theta_0)\end{pmatrix}[1 - |\varepsilon_t^*|^r] + o_p(1),
\end{aligned}$$

in view of the fact that $d_t^{\psi}(\theta_0^*) = d_t^{\psi}(\theta_0)$, $d_t^{\lambda}(\theta_0^*) = \kappa_r d_t^{\lambda}(\theta_0)$, and $v_t(\theta_0^*) = \kappa_r$. Since $E|\varepsilon_t^*|^r = 1$, it follows that $\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\frac{\partial \ell_{t,r}(\theta_0^*)}{\partial\theta} \xrightarrow{d} N(0, r^4 b_r W_r)$ by the martingale central limit theorem. Moreover, since $\widehat{\theta}_{n,r} \xrightarrow{a.s.} \theta_0^*$, by the dominated convergence theorem and similar arguments as before, we have $\frac{1}{n}\sum_{t=1}^{n}\frac{\partial^2 \ell_{t,r}(\xi)}{\partial\theta\partial\theta'} \xrightarrow{p} r^2 W_r$. Hence, the result (ii) follows from (A.1). □

Proof of Theorem 3.1. First, we denote that $\widehat{\theta}_n = \arg\min_{\theta\in\Theta} H_n(\theta)$, where

$$H_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\rho_\tau\Big(\frac{y_t}{Q_t(\widehat{\theta}_{n,r})} + \frac{Q_t(\theta)}{Q_t(\widehat{\theta}_{n,r})}\Big) = \frac{1}{n}\sum_{t=1}^{n}\rho_\tau\Big(\frac{\varepsilon_t}{Q_t(\widehat{\theta}_{n,r})/q_t} + \frac{Q_t(\theta)/q_t}{Q_t(\widehat{\theta}_{n,r})/q_t}\Big).$$

By Lemmas A.1(i) and A.3(i), $Q_t(\widehat{\theta}_{n,r})/q_t$ is bounded (a.s.) for large t. However, by Lemma A.1(iii), $Q_t(\theta)/q_t \xrightarrow{a.s.} \infty$ for $\theta \notin \Theta_0$ and large $t$. Therefore, $H_n(\theta) \xrightarrow{a.s.} \infty$ for $\theta \notin \Theta_0$, implying that the minimum value of $H_n(\theta)$ can not be reached outside $\Theta_0$, so $\widehat{\theta}_n = \arg\min_{\theta\in\Theta_0} H_n(\theta)$.

Next, we further rewrite $\widehat{\theta}_n = \arg\min_{\theta\in\Theta_0} L_n(\theta)$, where

$$L_n(\theta) = \frac{1}{n}\sum_{t=1}^{n}\Big[\rho_\tau\Big(\frac{y_t}{Q_t(\widehat{\theta}_{n,r})} + \frac{Q_t(\theta)}{Q_t(\widehat{\theta}_{n,r})}\Big) - \rho_\tau\Big(\frac{y_t}{Q_t(\widehat{\theta}_{n,r})} + \frac{1}{\kappa_r}\Big)\Big].$$

Then, by using the identity

$$\rho_\tau(x-y) - \rho_\tau(x) = -y\pi_\tau(x) + \int_0^y [\mathrm{I}(x \le s) - \mathrm{I}(x \le 0)]\, ds \tag{A.2}$$

with $\pi_\tau(x) = \tau - \mathrm{I}(x < 0)$, it follows that

$$\begin{aligned} L_n(\theta) &= \frac{1}{n}\sum_{t=1}^{n}\Big[\frac{Q_t(\theta)}{Q_t(\widehat{\theta}_{n,r})} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{y_t}{Q_t(\widehat{\theta}_{n,r})} + \frac{1}{\kappa_r}\Big) \\ &\quad + \frac{1}{n}\sum_{t=1}^{n}\int_0^{\frac{1}{\kappa_r} - \frac{Q_t(\theta)}{Q_t(\widehat{\theta}_{n,r})}} \mathrm{I}\Big(\frac{y_t}{Q_t(\widehat{\theta}_{n,r})} \le s - \frac{1}{\kappa_r}\Big) - \mathrm{I}\Big(\frac{y_t}{Q_t(\widehat{\theta}_{n,r})} \le -\frac{1}{\kappa_r}\Big) ds \\ &\triangleq I_{1n}(\theta) + I_{2n}(\theta). \end{aligned} \tag{A.3}$$

For $I_{1n}(\theta)$, by Lemma A.1(i), Lemma A.3(ii), the boundedness of $\pi_\tau(\cdot)$ and $\sup_x f(x)$, and Taylor's expansion, we have

$$\begin{aligned} I_{1n}(\theta) &= \frac{1}{n}\sum_{t=1}^{n}\Big[\frac{Q_t(\theta)/q_t}{Q_t(\widehat{\theta}_{n,r})/q_t} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{Q_t(\widehat{\theta}_{n,r})/q_t} + \frac{1}{\kappa_r}\Big) \\ &= \frac{1}{n}\sum_{t=1}^{n}\Big[\frac{v_t(\theta)}{v_t(\widehat{\theta}_{n,r})} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{Q_t(\widehat{\theta}_{n,r})/q_t} + \frac{1}{\kappa_r}\Big) + o_p(1) \\ &= \frac{1}{n}\sum_{t=1}^{n}\Big[\frac{v_t(\theta)}{v_t(\theta_0^*)} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{Q_t(\widehat{\theta}_{n,r})/q_t} + \frac{1}{\kappa_r}\Big) + o_p(1) \\ &= \frac{1}{n}\sum_{t=1}^{n}\Big[\frac{v_t(\theta)}{v_t(\theta_0^*)} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{v_t(\widehat{\theta}_{n,r})} + \frac{1}{\kappa_r}\Big) + o_p(1), \end{aligned} \tag{A.4}$$

where $o_p(1)$ holds uniformly in $\theta \in \Theta_0$. Furthermore, it follows that

$$\begin{aligned} I_{1n}(\theta) &= E\Big\{\Big[\frac{v_t(\theta)}{v_t(\theta_0^*)} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{v_t(\widehat{\theta}_{n,r})} + \frac{1}{\kappa_r}\Big)\Big\} + o_p(1) \\ &= E\Big\{\Big[\frac{v_t(\theta)}{v_t(\theta_0^*)} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{v_t(\theta_0^*)} + \frac{1}{\kappa_r}\Big)\Big\} + o_p(1) \\ &= E\Big\{\Big[\frac{v_t(\theta)}{v_t(\theta_0^*)} - \frac{1}{\kappa_r}\Big]\pi_\tau\Big(\frac{\varepsilon_t}{\kappa_r} + \frac{1}{\kappa_r}\Big)\Big\} + o_p(1) \end{aligned}$$

$$= o_p(1), \tag{A.5}$$

where $o_p(1)$ holds uniformly in $\theta \in \Theta_0$, the first equality holds by the uniform ergodic theorem, the second equality holds by the dominated convergence theorem and Lemma A.3(ii), the third equality holds since $v_t(\theta_0^*) = \kappa_r v_t(\theta_0) = \kappa_r$, and the last equality holds by the double expectation and the fact that the $\tau$th quantile of $\varepsilon_t$ is $-1$.

For $I_{2n}(\theta)$, by the similar arguments as for $I_{1n}(\theta)$, we have

$$I_{2n}(\theta) = E\Big[\int_0^{\frac{1}{\kappa_r}-\frac{v_t(\theta)}{\kappa_r}} \mathrm{I}\Big(\frac{\varepsilon_t}{\kappa_r} \le s - \frac{1}{\kappa_r}\Big) - \mathrm{I}\Big(\frac{\varepsilon_t}{\kappa_r} \le -\frac{1}{\kappa_r}\Big)ds\Big] + o_p(1),$$

where $o_p(1)$ holds uniformly in $\theta \in \Theta_0$. Since $\sup_{\theta\in\Theta_0}|\frac{1}{\kappa_r} - \frac{v_t(\theta)}{\kappa_r}|$ is bounded (a.s.) for some constant $c_1 > 0$, by the double expectation and Taylor's expansion, it follows that

$$\begin{aligned}
I_{2n}(\theta) &= E\Big[\int_0^{\frac{1}{\kappa_r}-\frac{v_t(\theta)}{\kappa_r}} F\big(s\kappa_r - 1\big) - F\big(-1\big)ds\Big] + o_p(1) \\
&= \kappa_r E\Big[\int_0^{\frac{1}{\kappa_r}-\frac{v_t(\theta)}{\kappa_r}} f(\xi)sds\Big] + o_p(1) \\
&\ge \frac{\inf_{|x|\le c_1} f(x)}{2\kappa_r} E[1 - v_t(\theta)]^2 + o_p(1)
\end{aligned} \tag{A.6}$$

with the equality holds if and only if $v_t(\theta) = 1$ (a.s.), or equivalently, $\theta = \theta_0$ (see Lemma A.2 in Francq and Zakoïan (2012)), where $F(\cdot)$ is the distribution function of $\varepsilon_t$, and $\xi$ lies between $s\kappa_r - 1$ and $-1$.

Finally, the conclusion holds by (A.3), (A.5)–(A.6), and standard arguments. □

**LEMMA A.4.** *Suppose Assumptions 3.1–3.3 hold. Then,*

(i) $I_{3n} = o_p\big(\sqrt{n}(\widehat{\theta}_n - \theta_0)\big) + o_p(1)$;

(ii) $I_{4n} = o_p\big(\sqrt{n}(\widehat{\theta}_n - \theta_0)\big) + \dfrac{f(-1)}{\kappa_r} E\big(d_t(\theta_0)d_t'(\theta_0)\big)\big(\sqrt{n}(\widehat{\theta}_n - \theta_0)\big)$;

(iii) $I_{5n} = o_p(1)$;

(iv) $I_{6n} = \dfrac{1}{\sqrt{n}\kappa_r}\displaystyle\sum_{t=1}^{n} \pi_\tau\big(\varepsilon_t + 1\big)d_t(\theta_0) + o_p(1)$,

*where*

$$I_{3n} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \pi_\tau\big(y_t + Q_t(\widehat{\theta}_n)\big)\Big[\frac{1}{Q_t(\widehat{\theta}_{n,r})}\frac{\partial Q_t(\widehat{\theta}_n)}{\partial\theta} - \frac{1}{Q_t(\theta_0^*)}\frac{\partial Q_t(\theta_0)}{\partial\theta}\Big],$$

$$I_{4n} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Big[\pi_\tau\big(y_t + Q_t(\widehat{\theta}_n)\big) - \pi_\tau\big(y_t + Q_t(\theta_0)\big)\Big]\frac{1}{Q_t(\theta_0^*)}\frac{\partial Q_t(\theta_0)}{\partial\theta},$$

$$I_{5n} = \frac{1}{\sqrt{n}}\sum_{t=1}^{n} \Big[\pi_\tau\big(y_t + Q_t(\theta_0)\big) - \pi_\tau\big(y_t + q_t\big)\Big]\frac{1}{Q_t(\theta_0^*)}\frac{\partial Q_t(\theta_0)}{\partial\theta},$$

$$I_{6n} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \pi_\tau\big(y_t + q_t\big) \frac{1}{Q_t(\theta_0^*)} \frac{\partial Q_t(\theta_0)}{\partial \theta}.$$

*Proof.* (i) Rewrite $I_{3n}$ as $I_{3n} = I_{3n,1} + I_{3n,2}$, where

$$I_{3n,1} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \pi_\tau\big(y_t + Q_t(\widehat{\theta}_n)\big) \frac{1}{Q_t(\widehat{\theta}_{n,r})} \Big[\frac{\partial Q_t(\widehat{\theta}_n)}{\partial \theta} - \frac{\partial Q_t(\theta_0)}{\partial \theta}\Big],$$

$$I_{3n,2} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \pi_\tau\big(y_t + Q_t(\widehat{\theta}_n)\big) \Big[\frac{1}{Q_t(\widehat{\theta}_{n,r})} - \frac{1}{Q_t(\theta_0^*)}\Big] \frac{\partial Q_t(\theta_0)}{\partial \theta}.$$

For $I_{3n,1}$, by Lemmas A.1(i) and A.2(ii), Theorem 3.1, the mean value theorem, and the similar arguments as for (A.4), we can show

$$\begin{aligned} I_{3n,1} &= \Big[\frac{1}{n} \sum_{t=1}^{n} \pi_\tau\big(y_t + Q_t(\widehat{\theta}_n)\big) \frac{\Sigma_t(\theta_0)}{v_t(\theta_0^*)}\Big] \big[\sqrt{n}(\widehat{\theta}_n - \theta_0)\big] + o_p(1) \\ &= \Big[\frac{1}{n} \sum_{t=1}^{n} \pi_\tau\big(\varepsilon_t + v_t(\theta_0)\big) \frac{\Sigma_t(\theta_0)}{v_t(\theta_0^*)}\Big] \big[\sqrt{n}(\widehat{\theta}_n - \theta_0)\big] + o_p(1) \\ &= o_p(1)\big[\sqrt{n}(\widehat{\theta}_n - \theta_0)\big] + o_p(1), \end{aligned}$$

in view of the fact that $v_t(\theta_0) = 1$ and $\varepsilon_t$ has $\tau$th quantile $-1$. Similarly, by Lemma A.3(ii) we have $I_{3n,2} = o_p(1)\big[\sqrt{n}(\widehat{\theta}_{n,r} - \theta_0^*)\big] + o_p(1) = o_p(1)$. Hence, the result (i) holds.

(ii) For $I_{4n}$, we can show

$$\begin{aligned} I_{4n} &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Big[\mathrm{I}\Big(\varepsilon_t < -\frac{Q_t(\theta_0)}{q_t}\Big) - \mathrm{I}\Big(\varepsilon_t < -\frac{Q_t(\widehat{\theta}_n)}{q_t}\Big)\Big] \frac{d_t(\theta_0)}{v_t(\theta_0^*)} + o_p(1) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \Big[\mathrm{I}(\varepsilon_t < -v_t(\theta_0)) - \mathrm{I}(\varepsilon_t < -v_t(\widehat{\theta}_n))\Big] \frac{d_t(\theta_0)}{v_t(\theta_0^*)} + o_p(1), \end{aligned} \tag{A.7}$$

where the first equality holds by the boundedness of indicator function and Lemmas A.1(i) and A.2(i), and the second equality holds by Markov's inequality, Lemma A.1(i), and the fact that $\sup_x f(x) < \infty$ by Assumption 3.2. Let $\widehat{u}_n = \widehat{\theta}_n - \theta_0$, and

$$K_t(u) = \frac{d_t(\theta_0)}{\kappa_r} \Big[\mathrm{I}(\varepsilon_t < -v_t(\theta_0)) - \mathrm{I}(\varepsilon_t < -v_t(u + \theta_0))\Big].$$

Then, since $v_t(\theta_0^*) = \kappa_r$, by (A.7) we have $I_{4n} = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} K_t(\widehat{u}_n) + o_p(1)$. Rewrite

$$\frac{1}{\sqrt{n}} \sum_{t=1}^{n} K_t(u) = W_{1n}(u) + W_{2n}(u),$$

where

$$W_{1n}(u) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \{K_t(u) - E(K_t(u)|\mathcal{F}_{t-1})\} \text{ and } W_{2n}(u) = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} E(K_t(u)|\mathcal{F}_{t-1}).$$

Since $\widehat{u}_n = o_p(1)$ by Theorem 3.1, by the similar arguments as for Lemmas 2.2–2.3 in Zhu and Ling (2012), we can prove that $W_{1n}(\widehat{u}_n) = o_p(\sqrt{n}\widehat{u}_n) + o_p(1)$ and

$$W_{2n}(\widehat{u}_n) = \frac{f(-1)}{\kappa_r} E\big(d_t(\theta_0)d_t(\theta_0)'\big)(\sqrt{n}\widehat{u}_n) + o_p(\sqrt{n}\widehat{u}_n),$$

and hence the conclusion follows directly.

(iii)& (iv) The conclusions hold by using the similar arguments as for (A.7). $\square$

Proof of Theorem 3.2. Following the same argument as for Theorem 2.1 in Francq and Zakoïan (2012), the subgradient derivative with respect to $\theta$ is asymptotically equal to zero at the minimum, since $\widehat{\theta}_n \xrightarrow{p} \theta_0$ by Theorem 3.1, and $\theta_0$ belongs to the interior of $\Theta_0$. This implies

$$0 = \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \pi_\tau\big(y_t + Q_t(\widehat{\theta}_n)\big) \frac{1}{Q_t(\widehat{\theta}_{n,r})} \frac{\partial Q_t(\widehat{\theta}_n)}{\partial \theta} = I_{3n} + I_{4n} + I_{5n} + I_{6n},$$

where $I_{in}$, $i = 3, 4, 5, 6$, are defined as in Lemma A.4. From this lemma, it follows that

$$\begin{aligned} &\sqrt{n}(\widehat{\theta}_n - \theta_0) \\ &= \Big[\frac{f(-1)}{\kappa_r} E\big(d_t(\theta_0)d_t'(\theta_0)\big) + o_p(1)\Big]^{-1} \Big[-\frac{1}{\sqrt{n}\kappa_r}\sum_{t=1}^{n} \pi_\tau\big(\varepsilon_t + 1\big) d_t(\theta_0) + o_p(1)\Big]. \end{aligned} \tag{A.8}$$

Hence, the conclusion holds by the martingale central limit theorem and the fact that $E\big(d_t(\theta_0)d_t'(\theta_0)\big) = \Sigma_{\theta_0}$ via direct calculations. $\square$

Proof of Theorem 3.3. Since $\overline{Q}_t(\lambda) = (1-\lambda)\sum_{i=1}^{t} \lambda^{i-1}|y_{t-i}|$, we have

$$\begin{aligned} \frac{\partial \overline{Q}_t(\lambda)}{\partial \lambda} &= -\sum_{i=1}^{t} \lambda^{i-1}|y_{t-i}| + (1-\lambda)\sum_{i=2}^{t}(i-1)\lambda^{i-2}|y_{t-i}|, \\ \frac{\partial^2 \overline{Q}_t(\lambda)}{\partial \lambda^2} &= -2\sum_{i=2}^{t}(i-1)\lambda^{i-2}|y_{t-i}| + (1-\lambda)\sum_{i=3}^{t}(i-1)(i-2)\lambda^{i-3}|y_{t-i}|. \end{aligned}$$

By the similar arguments as for Lemma A.2, we can show that there exists a constant $c_0 > 0$ such that, as $t \to \infty$,

$$\begin{aligned} e^{c_0 t} \sup_{\lambda \in \Theta_\lambda} \left\| \frac{1}{q_t} \frac{\partial \overline{Q}_t(\lambda)}{\partial \lambda} - (\overline{d}_t^{\lambda}(\lambda_0) - \overline{d}_t^{\psi}(\lambda_0)) \right\| &\xrightarrow{a.s.} 0, \\ e^{c_0 t} \sup_{\theta \in \Theta_0} \left\| \frac{1}{q_t} \frac{\partial^2 \overline{Q}_t(\lambda)}{\partial \lambda^2} - (\overline{d}_t^{\lambda\lambda}(\lambda_0) - 2\overline{d}_t^{\psi\lambda}(\lambda_0)) \right\| &\xrightarrow{a.s.} 0, \end{aligned}$$

where $\overline{d}_t^{\lambda}(\lambda)$, $\overline{d}_t^{\psi}(\lambda)$, $\overline{d}_t^{\psi\lambda}(\lambda)$, and $\overline{d}_t^{\lambda\lambda}(\lambda)$ are defined in the same way as $d_t^{\lambda}(\theta)$, $d_t^{\psi}(\theta)$, $d_t^{\psi\lambda}(\theta)$, and $d_t^{\lambda\lambda}(\theta)$, respectively, with $\theta$ and $\theta_0$ replaced by $(1-\lambda,\lambda)'$ and $(1-\lambda_0,\lambda_0)'$. Then, the conclusions follow by the similar arguments as for Theorems 3.1–3.2 and the fact that $E[(\overline{d}_t^{\lambda}(\lambda_0) - \overline{d}_t^{\psi}(\lambda_0))^2] = \Sigma_{\lambda_0}$. $\square$

PROOF OF THEOREM 4.1. Define $\gamma_{s,n}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log(\lambda + \psi|\varepsilon_t(\theta)|)$, where $\varepsilon_t(\theta) = y_t/Q_t(\theta)$. Then, $\widehat{\gamma}_{s,n} = \gamma_{s,n}(\widehat{\theta}_n)$ and

$$\sqrt{n}(\widehat{\gamma}_{s,n} - \gamma_s) = \sqrt{n}(\gamma_{s,n}(\theta_0) - \gamma_s) + \sqrt{n}(\gamma_{s,n}(\widehat{\theta}_n) - \gamma_{s,n}(\theta_0)) \triangleq I_{7n} + I_{8n}. \tag{A.9}$$

For $I_{8n}$, since $\varepsilon_t(\widehat{\theta}_n) = \varepsilon_t/(Q_t(\widehat{\theta}_n)/q_t)$, by the mean value theorem, we have

$$\begin{aligned}
I_{8n} &= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[(\widehat{\lambda}_n + \widehat{\psi}_n|\varepsilon_t(\widehat{\theta}_n)|) - (\lambda_0 + \psi_0|\varepsilon_t(\theta_0)|)\Big]\frac{1}{\xi_t^*} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[\Big(\widehat{\lambda}_n + \widehat{\psi}_n\Big|\frac{\varepsilon_t}{v_t(\widehat{\theta}_n)}\Big|\Big) - \Big(\lambda_0 + \psi_0\Big|\frac{\varepsilon_t}{v_t(\theta_0)}\Big|\Big)\Big]\frac{1}{\xi_t^*} + O_p\Big(\frac{1}{\sqrt{n}}\Big) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[\Big(\widehat{\lambda}_n + \widehat{\psi}_n\Big|\frac{\varepsilon_t}{v_t(\widehat{\theta}_n)}\Big|\Big) - \Big(\lambda_0 + \psi_0\Big|\frac{\varepsilon_t}{v_t(\theta_0)}\Big|\Big)\Big]\frac{1}{\lambda_0 + \psi_0|\varepsilon_t(\theta_0)|} + O_p\Big(\frac{1}{\sqrt{n}}\Big) + o_p(1) \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\Big[\Big(\widehat{\lambda}_n + \widehat{\psi}_n\Big|\frac{\varepsilon_t}{v_t(\widehat{\theta}_n)}\Big|\Big) - \Big(\lambda_0 + \psi_0\Big|\frac{\varepsilon_t}{v_t(\theta_0)}\Big|\Big)\Big]\frac{1}{\lambda_0 + \psi_0|\varepsilon_t|} + O_p\Big(\frac{1}{\sqrt{n}}\Big) + o_p(1) \\
&\triangleq I_{8n,1} + O_p\Big(\frac{1}{\sqrt{n}}\Big) + o_p(1),
\end{aligned}$$

where $\xi_t^*$ lies between $\widehat{\lambda}_n + \widehat{\psi}_n|\varepsilon_t(\widehat{\theta}_n)|$ and $\lambda_0 + \psi_0|\varepsilon_t(\theta_0)|$, the second equality holds by Lemmas A.1(i) and A.2(i), the compactness of $\Theta$, and the fact that $\sqrt{n}(\widehat{\theta}_n - \theta_0) = O_p(1)$, the third equality holds by the similar arguments as for the second equality and the dominated convergence theorem, and the last equality holds by the similar arguments as for the second equality and the fact that $v_t(\theta_0) = 1$.

Let $\widehat{w}_n = \sqrt{n}\widehat{u}_n$. For $I_{8n,1}$, we rewrite it as $I_{8n,1} = I_{8n,2}(\widehat{\theta}_n)\widehat{w}_n + I_{8n,3}(\widehat{w}_n)$, where

$$\begin{aligned}
I_{8n,2}(\theta) &= \Big(\frac{1}{n}\sum_{t=1}^{n}\frac{|\varepsilon_t|/v_t(\theta)}{\lambda_0 + \psi_0|\varepsilon_t|}, \ \frac{1}{n}\sum_{t=1}^{n}\frac{1}{\lambda_0 + \psi_0|\varepsilon_t|}\Big), \\
I_{8n,3}(w) &= \frac{\psi_0}{\sqrt{n}}\sum_{t=1}^{n}\Big\{\Big|\frac{\varepsilon_t}{v_t(\theta_0 + w/\sqrt{n})}\Big| - |\varepsilon_t|\Big\}\frac{1}{\lambda_0 + \psi_0|\varepsilon_t|}.
\end{aligned}$$

By (A.2) with $\tau = 1/2$, we have $I_{8n,3}(w) = I_{8n,4}(w) + I_{8n,5}(w)$, where

$$\begin{aligned}
I_{8n,4}(w) &= -\frac{\psi_0}{\sqrt{n}}\sum_{t=1}^{n}\frac{\Delta(w)\mathrm{Sign}(\varepsilon_t)}{\lambda_0 + \psi_0|\varepsilon_t|}, \\
I_{8n,5}(w) &= \frac{2\psi_0}{\sqrt{n}}\sum_{t=1}^{n}\Big\{\int_0^{\Delta(w)} \mathrm{I}(\varepsilon_t \le s) - \mathrm{I}(\varepsilon_t \le 0)ds\Big\}\frac{1}{\lambda_0 + \psi_0|\varepsilon_t|},
\end{aligned}$$

where $\Delta(w)=\varepsilon_t\big(1-v_t^{-1}(\theta_0+w/\sqrt{n})\big)$. Since $\widehat{w}_n=O_p(1)$, we restrict our proofs below on the space $\Theta_w=\{w:|w|\le C\}$ for some $C>0$. By the mean value theory,

$$\Delta(w)=\frac{w'\varepsilon_t}{\sqrt{n}v_t(\theta_0+w/\sqrt{n})}\frac{\partial v_t(\xi)}{\partial\theta},$$

where $\xi$ lies between $\theta_0$ and $\theta_0+w/\sqrt{n}$. For $I_{8n,4}(w)$, it follows that

$$\begin{aligned}I_{8n,4}(w)&=-\frac{w'}{n}\sum_{t=1}^{n}\frac{\psi_0|\varepsilon_t|}{(\lambda_0+\psi_0|\varepsilon_t|)v_t(\theta_0+w/\sqrt{n})}\frac{\partial v_t(\xi)}{\partial\theta}\\&=-w'E\Big(\frac{\psi_0|\varepsilon_t|d_t(\theta_0)}{\lambda_0+\psi_0|\varepsilon_t|}\Big)+o_p(1),\end{aligned}$$

where $o_p(1)$ holds uniformly in $w\in\Theta_w$ by using the similar arguments as for (A.5). Since $E(d_t(\theta_0))=(\frac{1}{\psi_0},\frac{\nu_1}{\lambda_0(1-\nu_1)})'$ by direct calculations, we have $I_{8n,4}(\widehat{w}_n)=-\big(\frac{1-\nu_1}{\psi_0},\frac{\nu_1}{\lambda_0}\big)\widehat{w}_n+o_p(1)$. Similarly, since $\sqrt{n}\Delta(w)$ is uniformly bounded in probability, by the double expectation and Assumption 3.2, it is not hard to show that $I_{8n,5}(w)=O_p(\frac{1}{\sqrt{n}})$ uniformly in $w\in\Theta_w$. Hence, we can have

$$I_{8n,3}(\widehat{w}_n)=-\big(\frac{1-\nu_1}{\psi_0},\frac{\nu_1}{\lambda_0}\big)\widehat{w}_n+o_p(1).\tag{A.10}$$

Since $I_{8n,2}(\widehat{\theta}_n)=\big(\frac{1-\nu_1}{\psi_0},\frac{\nu_1}{\lambda_0}\big)+o_p(1)$ by the uniform ergodic theorem and dominated convergence theorem, it entails that $I_{8n,1}=o_p(1)$ and so $I_{8n}=o_p(1)$.

For $I_{7n}$, the standard central limit theorem implies that $I_{7n}\xrightarrow{d}N(0,\sigma^2_{\gamma_s})$ as $n\to\infty$. Hence, the conclusion holds by (A.9). □

Proof of Theorem 4.2. Rewrite

$$\sqrt{n}(\widehat{\gamma}_{m,n}-\gamma_m)=I_{9n}\widehat{w}_n+\psi_0I_{10n}+\psi_0I_{11n},\tag{A.11}$$

where $I_{9n}=\big(\frac{1}{n}\sum_{t=1}^{n}|\varepsilon_t(\widehat{\theta}_n)|,1\big)$, $I_{10n}=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\big(|\varepsilon_t(\widehat{\theta}_n)|-|\varepsilon_t|\big)$, and $I_{11n}=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\big(|\varepsilon_t|-E|\varepsilon_t|\big)$. By the similar arguments as for (A.10), we can show

$$I_{10n}=-\Big(\frac{E|\varepsilon_t|}{\psi_0},\frac{\nu_1E|\varepsilon_t|}{\lambda_0(1-\nu_1)}\Big)\widehat{w}_n+o_p(1).$$

Noth that $I_{9n}=(E|\varepsilon_t|,1)+o_p(1)$ by the uniform ergodic theorem and dominated convergence theorem. By (A.8) and (A.11), it entails that $\sqrt{n}(\widehat{\gamma}_{m,n}-\gamma_m)=\Lambda\Sigma_{\theta_0,|\varepsilon|,n}+o_p(1)$, where

$$\Sigma_{\theta_0,|\varepsilon|,n}=\frac{1}{\sqrt{n}}\sum_{t=1}^{n}\left(\begin{array}{c}-\frac{\pi_\tau(\varepsilon_t+1)}{f(-1)}\Sigma_{\theta_0}^{-1}d_t(\theta_0)\\|\varepsilon_t|-E|\varepsilon_t|\end{array}\right).$$

Now, the conclusion follows, since $\Sigma_{\theta_0,|\varepsilon|,n} \xrightarrow{d} N(0, \Sigma_{\theta_0,|\varepsilon|})$ as $n \to \infty$ by the martingale central limit theorem. $\square$

PROOF OF THEOREM 4.3. Since $\widehat{\varepsilon}_t = \varepsilon_t/(Q_t(\widehat{\theta}_n)/q_t)$ and $\widehat{\theta}_n = \theta_0 + o_p(1)$, by the mean value theorem, Lemma A.1, and the dominated convergence theorem, it is not hard to see that under $H_0$ in (4.5),

$$n(\widehat{\tau}_n - 1) = \frac{n^{-1}\sum_{t=2}^n \widehat{z}_t(\widehat{z}_t - \widehat{z}_{t-1})}{n^{-2}\sum_{t=2}^n \widehat{z}_{t-1}^2} = \frac{n^{-1}\sum_{t=2}^n z_t(z_t - z_{t-1})}{n^{-2}\sum_{t=2}^n z_{t-1}^2} + o_p(1),$$
$$n[\mathrm{se}(\widehat{\tau}_n)]\sqrt{\frac{n^{-1}\sum_{t=2}^n(\widehat{z}_t - \widehat{z}_{t-1})^2}{n^{-2}\sum_{t=2}^n \widehat{z}_t^2}} + o_p(1) = \sqrt{\frac{n^{-1}\sum_{t=2}^n(z_t - z_{t-1})^2}{n^{-2}\sum_{t=2}^n z_t^2}} + o_p(1).$$

Now, since $z_t = z_{t-1} + e_t$ under $H_0$, the conclusion holds by standard arguments. $\square$

PROOF OF THEOREM 5.1. Let $1 \le l \le k$ be any given integer. First, since $\sqrt{n}(\widehat{\theta}_n - \theta_0) = O_p(1)$ by Theorem 3.2, the similar arguments as for (A.4) entail

$$\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \widehat{Hit}_t|\widehat{\varepsilon}_{t-l}|^p = \frac{1}{\sqrt{n}}\sum_{t=k+1}^n \widehat{Hit}_t|\varepsilon_{t-l}|^p + o_p(1).$$

Rewrite $\widehat{Hit}_t = \big[\mathrm{I}(\varepsilon_t < -Q_t(\widehat{\theta}_n)/q_t) - \mathrm{I}(\varepsilon_t < -Q_t(\theta_0)/q_t)\big] + \big[\mathrm{I}(\varepsilon_t < -Q_t(\theta_0)/q_t) - \tau\big]$. Then, by using the similar arguments as for Lemma A.4(ii), we have

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \Big[\mathrm{I}\Big(\varepsilon_t < -\frac{Q_t(\widehat{\theta}_n)}{q_t}\Big) - \mathrm{I}\Big(\varepsilon_t < -\frac{Q_t(\theta_0)}{q_t}\Big)\Big]|\varepsilon_{t-l}|^p \\ &= -f(-1)E\big(|\varepsilon_{t-l}|^p d_t(\theta_0)'\big)[\sqrt{n}(\widehat{\theta}_n - \theta_0)] + o_p(1).\end{aligned}$$

By (A.8), it follows that

$$\begin{aligned}&\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \Big[\mathrm{I}\Big(\varepsilon_t < -\frac{Q_t(\widehat{\theta}_n)}{q_t}\Big) - \mathrm{I}\Big(\varepsilon_t < -\frac{Q_t(\theta_0)}{q_t}\Big)\Big]|\varepsilon_{t-l}|^p \\ &= E\big(|\varepsilon_{t-l}|^p d_t(\theta_0)'\big)\Sigma_{\theta_0}^{-1}\Big(\frac{1}{\sqrt{n}}\sum_{t=1}^n \pi_\tau(\varepsilon_t + 1)d_t(\theta_0)\Big) + o_p(1).\end{aligned}$$

Moreover, since $\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \big[\mathrm{I}(\varepsilon_t < -Q_t(\theta_0)/q_t) - \tau\big]|\varepsilon_{t-l}|^p = -\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \pi_\tau(\varepsilon_t + 1)|\varepsilon_{t-l}|^p + o_p(1)$ by using the similar arguments as for (A.4), we can obtain

$$\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \widehat{Hit}_t|\widehat{\varepsilon}_{t-l}|^p = \frac{1}{\sqrt{n}}\sum_{t=1}^n \pi_\tau(\varepsilon_t + 1)\Big[E\big(|\varepsilon_{t-l}|^p d_t(\theta_0)'\big)\Sigma_{\theta_0}^{-1}d_t(\theta_0) - |\varepsilon_{t-l}|^p\Big] + o_p(1),$$

which implies

$$\frac{1}{\sqrt{n}}\sum_{t=k+1}^n \widehat{Hit}_t\widehat{X}_{t,k}^{(p)} = \frac{1}{\sqrt{n}}\sum_{t=1}^n \pi_\tau(\varepsilon_t + 1)\Big[E\big(X_{t,k}^{(p)} d_t(\theta_0)'\big)\Sigma_{\theta_0}^{-1}d_t(\theta_0) - X_{t,k}^{(p)}\Big] + o_p(1).$$

Finally, the conclusion holds by using the martingale central limit theorem and the fact that $\widehat{\Upsilon}_{n,k}^{(p)} \xrightarrow{p} E\big(X_{t,k}^{(p)}[X_{t,k}^{(p)}]'\big) - E\big(X_{t,k}^{(p)}d_t(\theta_0)'\big)\Sigma_{\theta_0}^{-1}E\big(d_t(\theta_0)[X_{t,k}^{(p)}]'\big)$ as $n \to \infty$. $\square$

# References


[1] Francq, C. and Zakoïan, J.-M. (2012) Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models. *Econometrica* **80**, 821–861.

[2] Li, D., Zhang, X., Zhu, K. and Ling, S. (2018) The ZD-GARCH model: A new way to study heteroscedasticity. *Journal of Econometrics* **202**, 1–17.

[3] Zhu, K. and Li, W. K. (2015) A new Pearson-type QMLE for conditionally heteroskedastic models. *Journal of Business & Economic Statistics* **33**, 552-565.

[4] Zhu, K. and Ling, S. (2012) The global weighted LAD estimators for finite/infinite variance ARMA$(p, q)$ models. *Econometric Theory* **28**, 1065–1086.