SUPPLEMENTARY MATERIAL ON "SPECIFICATION TESTS FOR MULTIPLICATIVE ERROR MODELS"

Indeewara Perera

Department of Econometrics and Business Statistics, Monash University

indeewara.perera@monash.edu

Mervyn J. Silvapulle

Department of Econometrics and Business Statistics, Monash University

mervyn.silvapulle@monash.edu

This online Supplementary Material provides omitted proofs (Appendix B), simulation results for evaluating density forecasts (Appendix C), and additional results for the empirical example in Section 6 (Appendix D).

APPENDIX B: Omitted proofs

Proof of Lemma 2. Let

$$K_{ni} := I[\eta_{ni} \leq x(1+\rho_{ni})] - I[\eta_{ni} \leq x] - H_{\vartheta_n}[x(1+\rho_{ni})] + H_{\vartheta_n}(x).$$

Then, for $\vartheta_n \in B$,

$$\frac{1}{n}\sum_{i=1}^{n} \gamma_{ni}^2 \mathbb{E}[K_{ni}^2|\mathcal{F}_{ni}] \leq \frac{1}{n}\sum_{i=1}^{n} \gamma_{ni}^2 \sup_{\vartheta_n \in B} \sup_{|z|\leq b} |H_{\vartheta_n}\{x(1+z)\} - H_{\vartheta_n}(x)| \leq \frac{a}{n}\sum_{i=1}^{n} \gamma_{ni}^2.$$

Now, from Lemma 1 with $D_{ni} = n^{-1}\gamma_{ni}K_{ni}$, it follows that for any $\eta, c > 0$,

$$\begin{aligned} &P_n\big([|\widetilde{U}_n(x) - U_n(x)| > \eta] \cap \Pi_n\big) \\ &\quad \leq P_n\Big(\Big[|n^{-1/2}\sum_{i=1}^{n} \gamma_{ni} I\{|\gamma_{ni}| \leq an^{-1/2}\} K_{ni}| > \eta\Big] \cap \Big[\sum_{i=1}^{n} \mathbb{E}(D_{ni}^2 \mid \mathcal{F}_{ni}) \leq ac\Big]\Big) \\ &\quad \leq \exp\{-\eta^2/2a(\eta + c)\}. \end{aligned}$$

■

Proof of Lemma 3. Fix $x \geq 0$ and $\varepsilon, \eta > 0$. Choose $c > 0$ and a positive integer n_1 such that $P_n(n^{-1}\sum_{i=1}^n \gamma_{ni}^2 > c) < \varepsilon$ for all $n \geq n_1$. Select $a > 0$ to have $\exp\{-\eta^2/2a(\eta + c)\} < \varepsilon$. Further, choose $\delta > 0$ and $b_0 > 0$ such that $a - \delta > 0$ and $Ł_0 := \sup_{|z|\leq b_0} |H_{\vartheta_0}\{x(1+z)\} - H_{\vartheta_0}(x)| < \delta$. Then, $Ł_n := \sup_{|z|\leq b_0} |H_{\vartheta_n}\{x(1+z)\} - H_{\vartheta_n}(x)|$ forms a real sequence such that $Ł_n \to Ł_0$ as $n \to \infty$. Therefore, there exists an $n_2 \in \mathbb{N}$ such that $Ł_n < Ł_0 + a - \delta < a$ for all $n \geq n_2$. We also have that $\vartheta_n \in B$ for all $n \geq n_0$ for some n_0. For each $n \in \{n_0, n_0+1, \ldots, \max(n_1, n_2)\}$, there exists a $b_n > 0$ such that $\sup_{|z|\leq b_n} |H_{\vartheta_n}\{x(1+z)\} - H_{\vartheta_n}(x)| < a$.

Now, let $b := \min\{b_{n_0}, b_{n_0+1}, \ldots, b_{\max(n_1,n_2)}, b_0\}$. Then, for each $\vartheta_n \in B$,

$$\sup_{|z|\leq b} |H_{\vartheta_0}\{x(1+z)\} - H_{\vartheta_0}(x)| < a.$$

Therefore, $\sup_{\vartheta_n \in B} \sup_{|z|\leq b} |H_{\vartheta_0}\{x(1+z)\} - H_{\vartheta_0}(x)| < a$. Thus, it follows from Lemma 2 that for $\vartheta_n \in B$,

$$\begin{aligned} P_n(|\widetilde{U}_n(x) - U_n(x)| > \eta) &\leq P_n([|\widetilde{U}_n(x) - U_n(x)| > \eta] \cap \Pi_n) + P_n(\Pi_n^c) \\ &\leq \exp\{-\eta^2/2a(\eta + c)\} + P_n(\max_i |\gamma_{ni}| > an^{-1/2}) \\ &\quad + P_n(n^{-1}\sum_{i=1}^n \gamma_{ni}^2 > c) + P_n((\max_i |\rho_{ni}| > b). \end{aligned}$$

The first term of the last upper bound is less than ε; the second and fourth terms are $o(1)$, and the third term is less than ε for all $n \geq n_1$. Consequently,

$$\limsup_{n\to\infty} P_n(|\widetilde{U}_n(x) - U_n(x)| > \eta) < 2\varepsilon.$$

Since ε is arbitrary, we have $|\widetilde{U}_n(x) - U_n(x)| = o_{p_n}(1)$. ■

Proof of Lemma 4. Fix δ and b such that $0 < \delta < 1$ and $0 \leq b < 1$. Choose $K_{1\delta}^B > 0$ such that $\sup_{\vartheta\in B} H_\vartheta(K_{1\delta}^B) \leq \delta^2/2$. Let $x_0 = K_{1\delta}^B/(1+b)$, and define $\|h_\vartheta\|_\infty^B := \sup_{\vartheta\in B}\sup_{x\geq 0} |h_\vartheta(x)|$, where h_ϑ is the density of H_ϑ. Choose an integer $N_\delta^B > 0$ to have $\delta^2/4 \leq \sup_{\vartheta\in B}\{1 - H_\vartheta(K_{2\delta}^B)\} \leq \delta^2/2$, where $K_{2\delta}^B := (1-b)[x_0 + N_\delta^B\delta^2/\{2\|h_\vartheta\|_\infty^B\}]$.

Now, partition $\mathbb{R}^+ \cup \{0\}$ as

$$[0, x_0] \cup (x_0, x_1] \cup \ldots \cup (x_{N_\delta^B}, \infty), \tag{B.1}$$

where $x_{N_\delta^B} = K_{2\delta}^B/(1-b) > 0$ and $x_k - x_{k-1} = \delta^2/\{2\|h_\vartheta\|_\infty^B\}$, $k = 1, \ldots, N_\delta^B$. In this partition, there are N_δ^B subintervals of length $\delta^2/\{2\|h_\vartheta\|_\infty^B\}$ covering the interval $(x_0, x_{N_\delta^B}]$. Because $x_0 = K_{1\delta}^B/(1+b) > 0$,

$$N_\delta^B = (1-b^2)^{-1}\{(1+b)K_{2\delta}^B - (1-b)K_{1\delta}^B\}2\|h_\vartheta\|_\infty^B/\delta^2.$$

Since $x_0 > 0$, we also have $\{x(1+b)\} \leq K_{1\delta}^B$ for $x \in [0, x_0]$. Therefore,

$$[\mu_b^B(x,y)]^2 \leq 2\sup_{\vartheta\in B} H_\vartheta(K_{1\delta}^B) \leq \delta^2 \text{ for } x, y \in [0, x_0].$$

For $x, y \in (x_{N_\delta^B}, \infty)$, we have $\{x(1-b)\} \geq K_{2\delta}^B$ and $\{y(1-b)\} \geq K_{2\delta}^B$, and hence,

$$[\mu_b^B(x,y)]^2 \leq 2\sup_{\vartheta\in B} |1 - H_\vartheta(K_{2\delta}^B)| \leq \delta^2 \text{ for } x, y \in (x_{N_\delta^B}, \infty).$$

Because $0 \leq b < 1$, by applying the mean value theorem,

$$[\mu_b^B(x,y)]^2 \leq \delta^2\{2\|h_\vartheta\|_\infty^B\}^{-1}\ (1+b)\|h_\vartheta\|_\infty^B \leq \delta^2, \text{ for } x, y \in (x_{k-1}, x_k],\ k = 1, \ldots, N_\delta^B.$$

Thus, each interval in the partition (B.1) has diameter less than δ with respect to the pseudo-metric μ_b^B. Therefore, $\mathcal{N}(\delta, b) \leq 2 + N_\delta^B$, and hence,

$$\mathcal{N}(\delta, b) \leq 2 + (1-b^2)^{-1}\{(1+b)K_{2\delta}^B - (1-b)K_{1\delta}^B\}2\|h_\vartheta\|_\infty^B/\delta^2. \qquad \text{(B.2)}$$

Now, let $\mu_\vartheta = \mathbb{E}(\epsilon_\vartheta)$ and $\mu_B = \sup_{\vartheta\in B}\mu_\vartheta$, where $\epsilon_\vartheta \sim H_\vartheta$. It follows by applying Markov's inequality that

$$\delta^2/4 \ \leq \ \sup_{\vartheta\in B}\{1 - H_\vartheta(K_{2\delta}^B)\} = \sup_{\vartheta\in B} Pr(\epsilon_\vartheta \geq K_{2\delta}^B) \leq \sup_{\vartheta\in B}\mathbb{E}(\epsilon_\vartheta)/K_{2\delta}^B \leq \mu_B/K_{2\delta}^B.$$

Thus, $K_{2\delta}^B \leq 4\mu_B/\delta^2$. Further, $K_{1\delta}^B > 0$ and $0 < \delta < 1$. Therefore, it follows from (B.2) that $\mathcal{N}(\delta, b) \leq D(b)/\delta^4$, where

$$D(b) := 2(1-b^2)^{-1}[(1-b^2) + 4(1+b)\mu_B\|h_\vartheta\|_\infty^B].$$

Because $D(b)$ is increasing in b, we have that $I(b) = \int_0^1[\log\mathcal{N}(u,b)]^{1/2}du < \infty$ for $0 \leq b < 1$. ■

Proof of Lemma 6. Fix $M < \infty$. From (C2) and (E3) it follows that

$$\sup_{\{\phi\in\Phi,\ \sqrt{n}\|\phi-\phi_n\|\le M\}} \max_{1\le i\le n}\{v_{ni}(\phi) - (\phi-\phi_n)^\top \lambda_i(\phi_n)\} = o_{p_n}(n^{-1/2}).$$

Further, because $\max_{1\le i\le n} n^{-1/2}\lambda_i(\phi_n) = o_{p_n}(1)$ and $n^{-1}\sum_{i=1}^n \|\lambda_i(\phi_n)\| = O_{p_n}(1)$, one obtains that $\sup_{\{\phi\in\Phi,\sqrt{n}\|\phi-\phi_n\|\le M\}} n^{-1/2}\sum_{i=1}^n |v_{ni}(\phi)| = O_{p_n}(1)$.

Let $A_n = \{(x,\phi) : x \ge 0, \phi \in \Phi \text{ and } \sqrt{n}\|\phi-\phi_n\| \le M\}$. Let $a > 0$ be as in Condition (C3). Then, there exists an $n_0 > 0$ such that for all $n > n_0$ and $(\phi, x) \in A_n$, $F_{\theta_n}(x + xv_{ni}(\phi)) - F_{\theta_n}(x) = xv_{ni}(\phi) f_{\theta_n}(x) + 2^{-1}\{xv_{ni}(\phi)\}^2 f'_{\theta_n}(x(1+\delta_x^*))$, where δ_x^* is a real number satisfying $|\delta_x^*| < a$.

Because $\sup_{\{\phi\in\Phi,\ \sqrt{n}\|\phi-\phi_n\|\le M\}} n^{-1/2}\sum_{i=1}^n |v_{ni}(\phi)| = O_{p_n}(1)$, it follows from the assumptions in Condition (C3) that $\sup_{x,\phi,M} n^{-1/2}\sum_{i=1}^n 2^{-1}\{xv_{ni}(\phi)\}^2 f'_{\theta_n}\{x(1+\delta_x^*)\} = o_{p_n}(1)$. By Conditions (C2), (C3), (C5), and Assumption (E3), one also obtains that $\sup_{x,\phi,M} |\{n^{-1/2}\sum_{i=1}^n xv_{ni}(\phi) f_{\theta_n}(x)\} - B_n(x)| = o_{p_n}(1)$. Thus, the proof follows. ■

Proof of Lemma 7. We indicate only the main idea of the proof. Recall that $\dot{g}_\theta(t) = (\partial/\partial\theta) g_\theta(t) = [\dot{g}_{\theta 1}(t), \ldots, \dot{g}_{\theta q}(t)]^\top$. Fix a $j \in \{1, \ldots, q\}$. Because $\hat{\theta} = \arg\max_{\theta\in\Theta}\sum_{i=1}^n g_\theta(\widetilde{\varepsilon}_i)$, under (E2),

$$0 = \textstyle\sum_{i=1}^n \dot{g}_{\hat{\theta}_j}(\widetilde{\varepsilon}_i) = \sum_{i=1}^n \dot{g}_{\theta_{0j}}(\widetilde{\varepsilon}_i) + (\hat{\theta}_j - \theta_{0j})\sum_{i=1}^n \left[(\partial/\partial\theta_j)\dot{g}_{\theta_j}(\widetilde{\varepsilon}_i)\,|_{\theta_j=\bar{\theta}_j}\right], \qquad \text{(B.3)}$$

for some $\bar{\theta}_j$, where $|\bar{\theta}_j - \theta_{0j}| < |\hat{\theta}_j - \theta_{0j}|$. Let

$$\widetilde{H}_n(\theta) = n^{-1}\sum_{i=1}^n \ddot{g}_\theta(\widetilde{\varepsilon}_i), \quad H_n(\theta) = n^{-1}\sum_{i=1}^n \ddot{g}_\theta(\varepsilon_i).$$

Suppose that H_0 holds. Then, under (C1), (E1) and (E2), there exist an $\alpha_L > 0$, $0 < K < \infty$, and an open neighbourhood B of θ_0, such that

$$\begin{aligned}
&\sup_{\theta\in B}\|\widetilde{H}_n(\theta) - H_n(\theta)\| \\
&\le\ Kn^{-1}\sum_{i=1}^n |\widetilde{\varepsilon}_i - \varepsilon_i| \ =\ Kn^{-1}\sum_{i=1}^n Z_i\left|\frac{1}{\widetilde{\Psi}_i(\hat{\phi})} - \frac{1}{\Psi_i(\hat{\phi})} + \frac{1}{\Psi_i(\hat{\phi})} - \frac{1}{\Psi_i(\phi_0)}\right| \\
&\le\ \frac{K}{\alpha_L n}\sum_{i=1}^n \varepsilon_i \sup_{\phi\in\Phi}|\widetilde{\Psi}_i(\phi) - \Psi_i(\phi)| + \frac{K}{\alpha_L n}\sum_{i=1}^n Z_i|\Psi_i(\hat{\phi}) - \Psi_i(\phi_0)|\Psi_i^{-1}(\phi_0) + o_p(1) \\
&=\ S_{n1} + S_{n2} + o_p(1), \quad \text{say.}
\end{aligned}$$

Since $\sup_{\phi\in\Phi}|\widetilde{\Psi}_i(\phi)-\Psi_i(\phi)| \stackrel{e.a.s.}{\to} 0$ by (C2) and $\{\varepsilon_i\}$ are iid, it follows from Lemma 2.1 of Straumann and Mikosch (2006) that $S_{n1}=O_p(n^{-1})$. By Condition (C2),

$$S_{n2} \quad\leq\quad \|n^{1/2}(\hat{\phi}-\phi_0)^\top\| \max_{1\leq i\leq n} n^{-1/2}\|\lambda_i(\phi_0)\|\frac{K}{\alpha_L n}\sum_{i=1}^{n} Z_i + o_p(n^{-1/2}).$$

Since $n^{1/2}(\hat{\phi}-\phi_0)=O_p(1)$, it follows from Condition (C5) and the Ergodic Theorem that the first term in the last upper bound is $o_p(1)$. Hence, $\sup_{\theta\in B}\|\widetilde{H}_n(\theta)-H_n(\theta)\| = o_p(1)$. Therefore, one obtains by (B.3) that $\hat{\theta}_j-\theta_{0j}=n^{-1}\sum_{i=1}^n h_{\theta_{0j}}(\widetilde{\varepsilon}_i)+o_p(n^{-1/2})$. Now, apply a two-term Taylor series expansion for $n^{-1}\sum_{i=1}^n h_{\theta_{0j}}(\widetilde{\varepsilon}_i)$ and verify that the remainder term is $o_p(n^{-1/2})$. This completes the proof of the first part of Lemma 7.

For the second part, we establish the corresponding asymptotic expansion for $\hat{\theta}^*_j-\hat{\theta}_j$, in probability. To this end, let (a_n) be a subsequence of (n). Because $(\hat{\phi},\hat{\theta})$ converges in probability to (ϕ_0,θ_0), the subsequence (a_n) contains a further subsequence (r_n) such that $(\hat{\phi}_{r_n},\hat{\theta}_{r_n}) \stackrel{a.s.}{\to} (\phi_0,\theta_0)$. Now, choose a sample path along which $(\hat{\phi}_{r_n},\hat{\theta}_{r_n})\to(\phi_0,\theta_0)$. Then, it follows from (E3) that $\hat{\theta}^*_{r_n}-\hat{\theta}_{r_n}=o_{p^*_{r_n}}(1)$ and $\hat{\phi}^*_{r_n}-\hat{\phi}_{r_n}=O_{p^*_{r_n}}({r_n}^{-1/2})$ along the chosen fixed sample path.

Further, because the bootstrap is carried out under H_0, it follows by proceeding as in the proof of the first part that

$$\hat{\theta}^*_{r_nj}-\hat{\theta}_{r_nj}={r_n}^{-1}\sum_{i=1}^{r_n}\{h^*_{\hat{\theta}_{r_nj}}(\varepsilon^*_i)-(\hat{\phi}^*_{r_n}-\hat{\phi}_{r_n})^\top\varepsilon^*_i\lambda^*_i(\hat{\phi}_{r_n})h^{*'}_{\hat{\theta}_{r_nj}}(\varepsilon^*_i)\}+o_{p^*_{r_n}}({r_n}^{-1/2}).$$

Because this holds true for almost all sample paths, (A.5) holds in probability. ■

In the technical details involving the bootstrap method, we often need to show that certain terms are small, in the sense that they are $o_{p^*_n}(1)$, in probability. To establish this, it suffices to restrict the arguments to a subsequence (r_n) for which $(\hat{\phi}_{r_n},\hat{\theta}_{r_n})$ converges almost surely to (ϕ_0,θ_0), and to work along a fixed sample path for which $(\hat{\phi}_{r_n},\hat{\theta}_{r_n})\to(\phi_0,\theta_0)$. Hence, for showing that a quantity is $o_{p^*_n}(1)$ in probability, one may assume without loss of generality that $(\hat{\phi},\hat{\theta})\to(\phi_0,\theta_0)$ along almost all sample paths (see for example Theorem 5 of Salinetti, Vervaat and Wets, 1986). In the following lemmas, we restrict attention to such a fixed sample path. Thus, the terms '*e.a.s.*' and '*a.s.*' below may correspond to P^*_n probability, along a fixed sample path for which $(\hat{\phi},\hat{\theta})\to(\phi_0,\theta_0)$.

Proof of Lemma 8. Recall that $L_n^{*(m)}(\phi) = \sum_{i=1}^n \{\log \Psi_i^{*(m)}(\phi) + Z_i^{*(m)}/\Psi_i^{*(m)}(\phi)\}$ and $\hat{\phi}^{*(m)} = \arg\min_{\phi\in\Phi} L_n^{*(m)}(\phi)$. Let $\Delta_n^*(\phi) := (\partial/\partial\phi)\{L_n^{*(m)}(\phi) - L_n^*(\phi)\}$, and

$$\begin{aligned} a_i(\phi) &:= \{\Psi_i^{*(m)}(\hat{\phi})/\Psi_i^{*(m)}(\phi)\}\{\lambda_i^{*(m)}(\phi) - \lambda_i^*(\phi)\}, \\ b_i(\phi) &:= \lambda_i^*(\phi)\{\Psi_i^{*(m)}(\hat{\phi}) - \Psi_i^*(\hat{\phi})\}/\Psi_i^{*(m)}(\phi), \\ c_i(\phi) &:= \lambda_i^*(\phi)\{\Psi_i^*(\hat{\phi})/\Psi_i^*(\phi)\}\{\Psi_i^*(\phi) - \Psi_i^{*(m)}(\phi)\}/\Psi_i^{*(m)}(\phi). \end{aligned}$$

Then, $\Delta_n^*(\phi) = \sum_{i=1}^n \{\lambda_i^{*(m)}(\phi) - \lambda_i^*(\phi)\} - \{a_i(\phi) + b_i(\phi) + c_i(\phi)\}\varepsilon_i^*$. Now, let $e_{i1} = \|A_i(\cdot)\|$ and $e_{i2} = \|B_i(\cdot)\|$, where $A_i(\phi) = \{\Psi_i^{*(m)}(\phi) - \Psi_i^*(\phi)\}$ and $B_i(\phi) = \{\dot{\Psi}_i^{*(m)}(\phi) - \dot{\Psi}_i^*(\phi)\}$. Then, it follows from Conditions (C6) and (E3) that $e_{i1}, e_{i2} \overset{e.a.s.}{\to} 0$. Further,

$$\begin{aligned} \lambda_i^{*(m)}(\phi) - \lambda_i^*(\phi) &= \{\dot{\Psi}_i^{*(m)}(\phi)\}/\{\Psi_i^{*(m)}(\phi)\} - \{\dot{\Psi}_i^*(\phi)\}/\{\Psi_i^*(\phi)\} \\ &= \frac{1}{\Psi_i^{*(m)}(\phi)} B_i(\phi) - \lambda_i^*(\phi)\{\frac{1}{\Psi_i^{*(m)}(\phi)}\} A_i(\phi). \end{aligned}$$

Thus, for some fixed $\alpha > 0$, $\|\lambda_i^{*(m)}(\cdot) - \lambda_i^*(\cdot)\| \leq \alpha^{-1} e_{i2} + \alpha^{-1}\|\lambda_i^*(\cdot)\| e_{i1}$.

Because $\|\lambda_i^*(\cdot)\|$, $i = 1, 2\ldots$ are identically distributed and $e_{i1}, e_{i2} \overset{e.a.s.}{\to} 0$ as $i \to \infty$, then it follows from Lemma 2.1 of Straumann and Mikosch (2006) that $\|\lambda_i^{*(m)}(\cdot) - \lambda_i^*(\cdot)\| \overset{e.a.s.}{\to} 0$, as $i \to \infty$. Therefore, $\sum_{i=1}^n \|\lambda_i^{*(m)}(\cdot) - \lambda_i^*(\cdot)\|$ converges to a random variable (a.s.). Further, for some fixed $\alpha > 0$, we have that

$$\begin{aligned} &\|a_i(\cdot)\| \leq e_{i1}\alpha^{-1}|\Psi_i^*(\hat{\phi})|\|\lambda_i^{*(m)}(\cdot) - \lambda_i^*(\cdot)\|, \\ &\|b_i(\cdot)\| \leq e_{i1}\alpha^{-1}\|\lambda_i^*(\cdot)\|, \quad \|c_i(\cdot)\| \leq e_{i1}\alpha^{-2}|\Psi_i^*(\hat{\phi})|\|\lambda_i^*(\cdot)\|. \end{aligned}$$

Therefore, each of $\sup_{\phi\in\Phi}|a_i(\phi)|$, $\sup_{\phi\in\Phi}|b_i(\phi)|$, and $\sup_{\phi\in\Phi}|c_i(\phi)|$ is bounded by terms equal to the product of an identically distributed random variable and another term that $\overset{e.a.s}{\to} 0$ as $i \to \infty$. In view of Lemma 2.1 of Straumann and Mikosch (2006), if $\{v_i\}_{i\in\mathbb{Z}^+}$ is a sequence of identically distributed random elements with values in a separable Banach space (e.g. $\mathbb{R}^p$ with Euclidean norm) and $\zeta_i \overset{e.a.s}{\to} 0$, then $\sum_{i=1}^n \zeta_i\|v_i\|$ converges to a random variable (a.s.). Therefore, we obtain that

$$\sum_{i=1}^n \sup_{\phi\in\Phi} |a_i(\phi) + b_i(\phi) + c_i(\phi)|\varepsilon_i^* = O_{p_n^*}(1),$$

and hence, $\sup_{\phi\in\Phi}|\Delta_n^*(\phi)| = O_{p_n^*}(1)$. This completes the proof of part (a).

We indicate the main idea for the proof for part (b). To this end, it suffices to consider the simpler case when ϕ is a scalar parameter. Let $0.5 < \eta < 1$. Note that for a given $\delta > 0$, the curve $n^{-1}(\partial/\partial\phi)L_n^{*(m)}(\phi)$ lies in the band $n^{-1}(\partial/\partial\phi)L_n^*(\phi) \pm \delta n^{-\eta}$ with probability approaching 1 as $n \to \infty$. Let $S(\phi) = -(\partial/\partial\phi)n^{-1}L_n^*(\phi)$ and $J(\phi) = -(\partial/\partial\phi)n^{-1}L_n^{*(m)}(\phi)$. Let ϕ_a and ϕ_b be chosen such that $S(\phi_a) = n^{-\eta}\delta$ and $S(\phi_b) = -n^{-\eta}\delta$. Then, there exists a $K > 0$ such that $n^{1/2}|\phi_a - \hat{\phi}^*| < K$ and $n^{1/2}|\phi_b - \hat{\phi}^*| < K$. Let $B = \{\phi : n^{1/2}|\phi - \hat{\phi}^*| < K\}$. In view of Assumptions (E1) and (E3), there exists $c_0 > 0$, such that $P_n^*[\sup_{\phi\in B}|\dot{S}(\phi)| > c_0] \to 1$ as $n \to \infty$. In view of the mean value theorem, for some $\bar{\phi}_a$ and $\bar{\phi}_b$ satisfying $| \bar{\phi}_a - \hat{\phi}^* | \leq | \phi_a - \hat{\phi}^* |$ and $|\bar{\phi}_b - \hat{\phi}^*| \leq |\phi_b - \hat{\phi}^*|$,

$$\begin{aligned} n^{-\eta}\delta &= S(\phi_a) = S(\hat{\phi}^*) + (\phi_a - \hat{\phi}^*)\dot{S}(\bar{\phi}_a) = 0 + (\phi_a - \hat{\phi}^*)\dot{S}(\bar{\phi}_a), \\ -n^{-\eta}\delta &= S(\phi_b) = S(\hat{\phi}^*) + (\phi_b - \hat{\phi}^*)\dot{S}(\bar{\phi}_b) = 0 + (\phi_b - \hat{\phi}^*)\dot{S}(\bar{\phi}_b). \end{aligned}$$

Since $\hat{\phi}^* - \hat{\phi} = o_{p_n^*}(1)$, then with P_n^* probability $\to 1$ as $n \to \infty$, $|\phi_a - \phi_b| \leq 2n^{-\eta}\delta c_0^{-1}$.

Because $P_n^*[-\delta < n^\eta \sup_{\phi\in\Phi}(\partial/\partial\phi)\{n^{-1}L_n^{*(m)}(\phi) - n^{-1}L_n^*(\phi)\} < \delta] \to 1$, we have $J(\phi_a) \geq S(\phi_a) - n^{-\eta}\delta = 0$ and $J(\phi_b) \leq S(\phi_b) + n^{-\eta}\delta = 0$, with P_n^* probability $\to 1$. Thus, $P_n^*[\hat{\phi}^{*(m)} \in (\phi_a, \phi_b)] \to 1$, and hence, $|\hat{\phi}^{*(m)} - \hat{\phi}^*| \leq |\phi_a - \phi_b| \leq 2n^{-\eta}\delta c_0^{-1}$, with P_n^* probability $\to 1$. Therefore, part (b) follows.

To prove part (c), let $\ell(\theta) = \sum_{i=1}^n g_\theta(\widetilde{\varepsilon}_i)$. Then, the corresponding bootstrap terms are $\ell^*(\theta) = \sum_{i=1}^n g_\theta(\widetilde{\varepsilon}_i^*)$ and $\ell^{*(m)}(\theta) = \sum_{i=1}^n g_\theta(\widetilde{\varepsilon}_i^{*(m)})$. Further, for some $\bar{\varepsilon}_i$ between $\widetilde{\varepsilon}_i^{*(m)}$ and $\widetilde{\varepsilon}_i^*$, $|\ell^{*(m)}(\theta) - \ell^*(\theta)| = |\sum_{i=1}^n(\widetilde{\varepsilon}_i^{*(m)} - \widetilde{\varepsilon}_i^*)g'_\theta(\bar{\varepsilon}_i;\theta)|$. Hence, $|\ell^{*(m)}(\theta) - \ell^*(\theta)|$ is bounded from above by

$$\begin{aligned} & K_0\sum_{i=1}^n |Z_i^{*(m)}/\widetilde{\Psi}_i^{*(m)}(\hat{\phi}^{*(m)}) - Z_i^*/\widetilde{\Psi}_i^*(\hat{\phi}^*)| \\ &= K_0\sum_{i=1}^n \varepsilon_i^*|\Psi_i^{*(m)}(\hat{\phi})/\widetilde{\Psi}_i^{*(m)}(\hat{\phi}^{*(m)}) - \Psi_i^*(\hat{\phi})/\widetilde{\Psi}_i^{*(m)}(\hat{\phi}^{*(m)}) \\ &\quad + \Psi_i^*(\hat{\phi})/\widetilde{\Psi}_i^{*(m)}(\hat{\phi}^{*(m)}) - \Psi_i^*(\hat{\phi})/\widetilde{\Psi}_i^*(\hat{\phi}^*)| \\ &\leq K_0\sum_{i=1}^n \varepsilon_i^*[K_1|\Psi_i^{*(m)}(\hat{\phi}) - \Psi_i^*(\hat{\phi})| + K_2\Psi_i^*(\hat{\phi})|\widetilde{\Psi}_i^*(\hat{\phi}^*) - \widetilde{\Psi}_i^{*(m)}(\hat{\phi}^{*(m)})|, \end{aligned}$$

where K_0, K_1 and K_2 are fixed constants. In view of Assumption (E3), the terms $\sup_{\phi\in\Phi}|\Psi_i^{*(m)}(\hat{\phi}) - \Psi_i^*(\hat{\phi})|$ and $\sup_{\phi\in\Phi}|\widetilde{\Psi}_i^*(\hat{\phi}^*) - \widetilde{\Psi}_i^{*(m)}(\hat{\phi}^{*(m)})|$ converge to zero (e.a.s) as

$i \to \infty$. Therefore, by Lemma 2.1 of Straumann and Mikosch (2006), for every $\eta < 1$, $n^\eta \sup_{\theta\in\Theta}\{n^{-1}\ell^{*(m)}(\theta) - n^{-1}\ell^*(\theta)\} \overset{a.s.}{\to} 0$ as $n \to \infty$. Thus, one obtains by proceeding as in the proof of part (b) that $\hat\theta^{*(m)} - \hat\theta^* = o_{p_n^*}(n^{-\eta})$ for every $0.5 < \eta < 1$. ■

Proof of Lemma 9. Let $v_i^* = \{\Psi_i^*(\hat\phi)\}^{-1}\{\Psi_i^*(\hat\phi^*) - \Psi_i^*(\hat\phi)\}$. Because $n^{1/2}(\hat\phi^* - \hat\phi) = O_{p_n^*}(1)$, by Condition (C2), $\max_{1\le i\le n}\{v_i^* - (\hat\phi^* - \hat\phi)^\top\lambda_i^*(\hat\phi)\} = o_{p_n^*}(n^{-1/2})$. Because $\max_{1\le i\le n} n^{-1/2}\|\lambda_i^*(\hat\phi)\| = o_{p_n^*}(1)$, one obtains that $\max_{1\le i\le n}|v_i^*| = o_{p_n^*}(1)$. Let us write $\hat F_n^*(x) = n^{-1}\sum_{i=1}^n I(\varepsilon_i^* \le x + xv_i^*)$. Then, it follows from Lemma 5 with $\gamma_{ni} = 1$ and $\rho_{ni} = v_i^*$ that $\sup_{x\ge0}|U_n^*(x) - n^{-1/2}\sum_{i=1}^n\{F_{\theta_n}(x+xv_i^*) - F_{\theta_n}(x)\}| = o_{p_n^*}(1)$. Now, apply Lemma 6 to complete the proof. ■

Proof of Lemma 11. Let $v_{ni}^{*(m)} = n^{1/2}[\Psi_i^{*(m)}(\hat\phi^{*(m)}) - \Psi_i^{*(m)}(\hat\phi)]/\Psi_i^{*(m)}(\hat\phi)$. Then, it follows from a one-term Taylor expansion that $v_{ni}^{*(m)} = n^{1/2}(\hat\phi^{*(m)} - \hat\phi)^\top\lambda_i^{*(m)}(\hat\phi) + r_{ni}^{*(m)}$, for some random array $\{r_{ni}^{*(m)}\}$ satisfying $n^{-1}\sum_{i=1}^n r_{ni}^{*(m)} = O_{p_n^*}(n^{-1/2})$.

From the proof of Lemma 8, for some $\eta > 1/2$, $n^{\eta-1}\sum_{i=1}^n\|\lambda_i^{*(m)}(\hat\phi) - \lambda_i^*(\hat\phi)\| \to 0$ (a.s.). By assumption, $\max_{1\le i\le n} n^{-1/2}\|\lambda_i^*(\hat\phi)\| = o_{p_n^*}(1)$, and from Lemma 8, $n^{1/2}(\hat\phi^{*(m)} - \hat\phi) = O_{p_n^*}(1)$. Hence, $\max_{1\le i\le n}|n^{-1/2}v_{ni}^{*(m)}| = o_{p_n^*}(1)$. Further, with $\gamma_{ni} = 1$ and $\rho_{ni} = n^{-1/2}v_{ni}^{*(m)}$, it follows from Lemma 5 that uniformly in $y \ge 0$,

$$
\begin{aligned}
n^{1/2}\hat F_n^{*(m)}(y) \;&=\; n^{-1/2}\sum_{i=1}^n I(\hat\varepsilon_i^{*(m)} \le y) \;=\; n^{-1/2}\sum_{i=1}^n I\{\varepsilon_i^* \le y + yn^{-1/2}v_{ni}^{*(m)}\} \\
&= n^{1/2}F_n^{*(m)}(y) + n^{-1/2}\sum_{i=1}^n\{F_{\hat\theta}(y + yn^{-1/2}v_{ni}^{*(m)}) - F_{\hat\theta}(y)\} + o_{p_n^*}(1) \\
&= n^{1/2}F_n^{*(m)}(y) + n^{-1}\sum_{i=1}^n v_{ni}^{*(m)} y f_{\hat\theta}(y) + o_{p_n^*}(1).
\end{aligned}
$$

This equality continues to hold with ${}^{*(m)}$ replaced by *. Further, because $F_n^{*(m)} = F_n^*$, by triangle inequality,

$$
\sup_{y\ge0} n^{1/2}|\hat F_n^{*(m)}(y) - \hat F_n^*(y)| \le \sup_{y\ge0}\left|n^{-1}\sum_{i=1}^n(v_{ni}^{*(m)} - v_{ni}^*)yf_{\hat\theta}(y)\right| + o_{p_n^*}(1). \qquad \text{(B.4)}
$$

By direct substitution, we obtain that

$$
\begin{aligned}
v_{ni}^{*(m)} - v_{ni}^* \;&=\; n^{1/2}(\hat\phi^{*(m)} - \hat\phi)^\top\lambda_i^{*(m)}(\hat\phi) - n^{1/2}(\hat\phi^* - \hat\phi)^\top\lambda_i^*(\hat\phi) + [r_{ni}^{*(m)} - r_{ni}^*] \\
&=\; \{n^{1/2}(\hat\phi^{*(m)} - \hat\phi^*)^\top\lambda_i^{*(m)}(\hat\phi) + n^{1/2}(\hat\phi^* - \hat\phi)^\top[\lambda_i^{*(m)}(\hat\phi) - \lambda_i^*(\hat\phi)]\} + [r_{ni}^{*(m)} - r_{ni}^*].
\end{aligned}
$$

Because $n^{1/2}(\hat{\phi}^{*(m)} - \hat{\phi}^*) = o_{p_n^*}(1)$, $n^{-1}\sum_{i=1}^n \|\lambda_i^{*(m)}(\hat{\phi}) - \lambda_i^*(\hat{\phi})\| = o_{p_n^*}(n^{-\eta})$, and $n^{-1}\sum_{i=1}^n [r_{ni}^{*(m)} - r_{ni}^*] = O_{p_n^*}(n^{-1/2})$, this yields that $n^{-1}\sum_{i=1}^n \{v_{ni}^{*(m)} - v_{ni}^*\} = o_{p_n^*}(1)$. Since $\sup_{\theta\in B, y\geq 0}(1+y)f_\theta(y) < \infty$ for some open neighbourhood B of θ_0, the first part follows from (B.4). Because $\hat{\theta}^{*(m)} - \hat{\theta}^* = o_{p_n^*}(n^{-\eta})$, the second part follows from a one-term Taylor expansion. ■

Proof of Lemma 14. Let $f_{(n)}$, f_{θ_0} and $\tilde{f}$ denote the densities corresponding to $F_{(n)}$, F_{θ_0} and $\tilde{F}$, respectively. Let $\ell_n := \sum_{i=1}^n \log\{f_{(n)}(\varepsilon_i)/f_{\theta_0}(\varepsilon_i)\}$. It follows from Theorem 7.2 in van der Vaart (1998) that

$$\ell_n = \delta n^{-1/2}\textstyle\sum_{i=1}^n \{[\tilde{f}(\varepsilon_i) - f_{\theta_0}(\varepsilon_i)]f_{\theta_0}^{-1}(\varepsilon_i)\} - 2^{-1}\delta^2\sigma^2 + o_p(1),$$

where $\sigma^2 = \int_{x\geq 0}\{\tilde{f}(x) - f_{\theta_0}(x)\}^2 f_{\theta_0}^{-1}(x)dx$. Hence, by the central limit theorem, $\ell_n \xrightarrow{d} N(-2^{-1}\delta^2\sigma^2, \delta^2\sigma^2)$ under H_0. Therefore, by Le Cam's first lemma (see van der Vaart and Wellner, 1996, Theorem 3.10.2) H_{an} is contiguous with respect to H_0.

Let $G_n(t) = n^{-1/2}\sum_{i=1}^n g_i(t)$, where $g_i(t)$ is defined in equation (9) of the main text as $g_i(t) = a_i(t) - b_i(t) + c_i(t)$ with functions $a_i(\cdot), b_i(\cdot)$, and $c_i(\cdot)$ as defined in Section 4. Then, $G_n(\cdot)$ is the same as $G_n^*(\cdot)$ in the proof of Lemma 10, except that it is now defined for the original sample instead of the bootstrapped sample.

It may be verified that each of $n^{-1/2}\sum a_i(t)$, $n^{-1/2}\sum b_i(t)$, and $n^{-1/2}\sum c_i(t)$ is asymptotically equicontinuous, under H_0, by applying Markov's inequality and using Assumption (E2) and Conditions (C3) and (C5). Because $g_i(\cdot)$ forms a martingale difference sequence, by a martingale CLT, the finite dimensional distributions of $G_n(t)$ converge to those of the centered Gaussian process $G(\cdot)$ in Theorem 1 [under H_0]. Consequently, $G_n(t)$ converges weakly to $G(\cdot)$, $\sup_t |\widetilde{W}_n\{F_{\theta_0}^{-1}(t)\} - G_n(t)| = o_p(1)$, and $\mathbb{E}[G_n(t)\ell_n] = m(t,\theta_0) + o(1)$ under H_0, $t \in [0,1]$, where

$$\begin{aligned} m(t,\theta) &= \delta\int\{I(\varepsilon \leq F_\theta^{-1}(t)) - t\}\ d\tilde{F}(\varepsilon) \\ &\quad +\delta\Big[\int\{(\partial/\partial\theta)\dot{g}_\theta(y)\}^{-1}\ dF_\theta(y)\int\dot{g}_\theta(\varepsilon)\ d\tilde{F}(\varepsilon)\Big]^\top \dot{F}_\theta(F_\theta^{-1}(t)) \\ &= \delta[\tilde{F}(F_\theta^{-1}(t)) - t] + \delta\Big[\int\{\ddot{g}_\theta(y)\}^{-1}\ dF_\theta(y)\int\dot{g}_\theta(\varepsilon)\ d\tilde{F}(\varepsilon)\Big]^\top \dot{F}_\theta(F_\theta^{-1}(t)). \end{aligned}$$

Thus, $m(\cdot,\theta_0)$ is the same as $m_a(\cdot)$ in (A.15). By applying a general version of Le Cam's third lemma for sequences of probability measures in metric spaces (see van der

Vaart and Wellner, 1996, Theorem 3.10.7), we obtain that $\widetilde{W}_n \circ F_{\theta_0}^{-1}(\cdot)$ converges weakly to $W_a(\cdot)$ in $D[0,1]$, under H_{an}, where $W_a(\cdot) = m_a(\cdot) + G(\cdot)$. By assumption, $\int \dot{g}_{\theta_0}(\varepsilon)\, d\tilde{F}(\varepsilon) = 0$ only if $\tilde{F} = F_{\theta_0}$. Further, $\tilde{F} \neq F_{\theta_0}$ and

$$[t - \tilde{F}(F_{\theta_0}^{-1}(t))] \neq \Big[\int \{\ddot{g}_{\theta_0}(y)\}^{-1}\, dF_{\theta_0}(y) \int \dot{g}_{\theta_0}(\varepsilon)\, d\tilde{F}(\varepsilon)\Big]^{\top} \dot{F}_{\theta_0}(F_{\theta_0}^{-1}(t)), \quad t \in [0,1].$$

Hence, $m_a \not\equiv 0$ for $\delta > 0$.

Under H_{an}, $\hat{\theta} \to \theta_0$ in probability, and θ_0 is the true value satisfying $F^0 = F_{\theta_0}$ under H_0. Therefore, one may proceed as in the proof of Lemma 10 and show that $\widehat{W}_n^* \circ F_{\hat{\theta}}^{-1}(\cdot)$ converges weakly to $G(\cdot)$, under H_{an} [in probability]. Hence, the proof follows from Lemmas 11 and 13. ■

Proof of Proposition 2. Let

$$\ell_n = -\sum_{i=1}^{n} \left[\log\{\Psi_i(\phi_0) + r_i/\sqrt{n}\} - \log\{\Psi_i(\phi_0)\} - Z_i/\{\Psi_i(\phi_0) + r_i/\sqrt{n}\} + Z_i/\Psi_i(\phi_0)\right].$$

By arguing as in the proof of Lemma 14 we obtain the following: (a) H_{bn} is contiguous with respect to H_0, (b) $G_n(\cdot)$ converges weakly to $G(\cdot)$ and $\sup_t |\widetilde{W}_n\{F_{\theta_0}^{-1}(t)\} - G_n(t)| = o_p(1)$ under H_0, where $G_n = n^{-1/2}\sum_{i=1}^n g_i$, and (c) $\mathbb{E}[G_n(t)\ell_n] = m_b(t) + o(1)$.

Therefore, part (i) follows from Le Cam's third lemma (see van der Vaart and Wellner, 1996, Theorem 3.10.7) and the continuous mapping theorem.

To prove part (ii), note that under H_{bn}, $\hat{\phi} = \arg\min_{\phi \in \Phi} \sum_{i=1}^n \ell_i(\phi)$, where

$$\ell_i(\phi) = \log \widetilde{\Psi}_i(\phi) + \frac{\left[\Psi_i(\phi_0) + n^{-1/2} r_i\right]}{\widetilde{\Psi}_i(\phi)} \varepsilon_i = \log \widetilde{\Psi}_i(\phi) + \frac{[\Psi_i(\phi_0)\varepsilon_i]}{\widetilde{\Psi}_i(\phi)} + \frac{n^{-1/2} r_i \varepsilon_i}{\widetilde{\Psi}_i(\phi)}.$$

Let $\hat{\phi}^{(1c)} = \arg\min_{\phi \in \Phi} \sum_{i=1}^n \kappa_i(\phi)$, where $\kappa_i(\phi) = \log \widetilde{\Psi}_i(\phi) + [\Psi_i(\phi_0)\varepsilon_i]/\widetilde{\Psi}_i(\phi)$. In view of Assumption (E1), $\hat{\phi}^{(1c)} \xrightarrow{p} \phi_0$ as $n \to \infty$. Further, we have that

$$\sum_{i=1}^{n} (\partial/\partial\phi)\{\ell_i(\phi) - \kappa_i(\phi)\} = -\sum_{i=1}^{n} n^{-1/2} r_i \varepsilon_i \left[\widetilde{\lambda}_i(\phi)/\widetilde{\Psi}_i(\phi)\right].$$

Therefore, under H_{bn}, for any $1/4 < \eta < 1/2$,

$$n^{\eta} \sup_{\phi \in \Phi} \|(\partial/\partial\phi)\, \{n^{-1}\textstyle\sum_{i=1}^n \ell_i(\phi) - n^{-1}\sum_{i=1}^n \kappa_i(\phi)\}\| \xrightarrow{a.s.} 0 \;\text{ as } n \to \infty.$$

Consequently, $\hat{\phi} - \hat{\phi}^{(1c)} \xrightarrow{p} 0$ as $n \to \infty$, and hence, $\hat{\phi} \xrightarrow{p} \phi_0$. Further, $\hat{\theta} \equiv \theta_0 = 1$ for each n because the error distribution is standard exponential and the parameter space $\Theta = \{1\}$. Therefore, by arguments similar to those of the proof of Lemma 10, $\widehat{W}_n^* \circ F_{\hat{\theta}}^{-1}(\cdot)$ converges weakly to $G(\cdot)$ under H_{bn} [in probability]. Because $T_j = \mathfrak{h}_j(\widetilde{W}_n \circ F_{\theta_0}^{-1}) + o_p(1)$ and $T_j^{*(m)} = \mathfrak{h}_j(\widetilde{W}_n^{*(m)} \circ F_{\hat{\theta}}^{-1}) + o_{p^*}(1)$ [in probability] $(j = 1, \dots, 5)$, the proof follows from Lemmas 11 and 13, and the continuous mapping theorem. ■

Proof of Proposition 3. Let

$$\ell_n := -\textstyle\sum_{i=1}^n \left[\log\{\Psi_i(\phi_0) + r_i/\sqrt{n}\} - \log\{\Psi_i(\phi_0)\} + \log\{f_{(n)}(\varepsilon_i)/f_{\theta_0}(\varepsilon_i)\}\right].$$

One obtains by arguing as in the proof of Lemma 14 that H_{cn} is contiguous with respect to H_0. From the proof of Proposition 2 we have that $G_n(\cdot)$ converges weakly to $G(\cdot)$ and that $\sup_t |\widetilde{W}_n\{F_{\theta_0}^{-1}(t)\} - G_n(t)| = o_p(1)$ under H_0. We also have that $\lim_{n\to\infty} \mathbb{E}[G_n(t)\ell_n] = m_c(t) + m_a(t) + o(1)$. Therefore, part (i) follows as in the proof of Proposition 2.

To prove part (ii), note that under H_{cn}, $\hat{\phi} = \arg\min_{\phi\in\Phi} \sum_{i=1}^n \ell_i(\phi)$, where

$$\ell_i(\phi) = \log\widetilde{\Psi}_i(\phi) + \frac{\left[\Psi_i(\phi_0) + n^{-1/2}r_i\right]}{\widetilde{\Psi}_i(\phi)}\varepsilon_i = \log\widetilde{\Psi}_i(\phi) + \frac{[\Psi_i(\phi_0)\varepsilon_i]}{\widetilde{\Psi}_i(\phi)} + \frac{n^{-1/2}r_i\varepsilon_i}{\widetilde{\Psi}_i(\phi)}.$$

Now, let $\hat{\phi}^{(2c)} = \arg\min_{\phi\in\Phi} \sum_{i=1}^n \kappa_i(\phi)$, where $\kappa_i(\phi) = \log\widetilde{\Psi}_i(\phi) + [\Psi_i(\phi_0)\varepsilon_i]/\widetilde{\Psi}_i(\phi)$. In view of Assumption (E1), $\hat{\phi}^{(2c)} \xrightarrow{p} \phi_0$ as $n \to \infty$. Further, we have that

$$\sum_{i=1}^n (\partial/\partial\phi)\{\ell_i(\phi) - \kappa_i(\phi)\} = -\sum_{i=1}^n n^{-1/2} r_i\varepsilon_i \left[\widetilde{\lambda}_i(\phi)/\widetilde{\Psi}_i(\phi)\right].$$

In view of these, under H_{cn}, for any $1/4 < \eta < 1/2$,

$$n^\eta \sup_{\phi\in\Phi} \left\|(\partial/\partial\phi)\left\{n^{-1}\sum_{i=1}^n \ell_i(\phi) - n^{-1}\sum_{i=1}^n \kappa_i(\phi)\right\}\right\| \xrightarrow{a.s.} 0 \quad \text{as } n \to \infty.$$

Further, $\hat{\phi} - \hat{\phi}^{(2c)} \xrightarrow{p} 0$ as $n \to \infty$, and hence, $\hat{\phi} \xrightarrow{p} \phi_0$ as $n \to \infty$. Note that under H_{cn}, $\hat{\theta} \to \theta_0$ as $n \to \infty$, where θ_0 is the true value satisfying $F^0 = F_{\theta_0}$ under H_0. As in the proof of Lemma 14, $\widehat{W}_n^* \circ F_{\hat{\theta}}^{-1}(\cdot)$ converges weakly to $G(\cdot)$ under H_{cn} [in probability], and hence the rest of the arguments follow as in the proof of Proposition 2. ■

APPENDIX C: Simulation study on the importance of the tests in density forecasting

One important area of application of the tests proposed in the paper is forecasting the conditional distribution and/or density of Z_{i+1}. The role of these tests in forecasting is that they can be used for testing the goodness-of-fit of the specified parametric model. To evaluate the potential contribution of these tests, this simulation study estimates different measures of 'loss' in using an incorrect parametric family when the tests have adequate power to reject the incorrect family.

Design of the simulation study:

Let $\{G_\beta : \beta \in B\}$ and $\{F_\theta : \theta \in \Theta\}$ denote two distinct families of cumulative distribution functions (cdfs). Let $\Psi_i(\phi) = \phi_1 + \phi_2 Z_{i-1} + \phi_3 \Psi_{i-1}(\phi)$, $\phi = (\phi_1, \phi_2, \phi_3)^\top \in \Phi$, denote the parametric specification for Ψ_i. The design of the simulation is based on the following scenario: The true DGP is $\Psi_i = \Psi_i(\phi_0)$ for some $\phi_0 \in \Phi$ and $F^0 \in \{G_\beta : \beta \in B\}$. The parametric model being considered for use in forecasting, and hence defines the null hypothesis is $H_0 : \Psi_i \in \{\Psi_i(\phi) : \phi \in \Phi\}$ and $F^0 \in \{F_\theta : \theta \in \Theta\}$. We are interested in estimating some measures of 'loss' in using the incorrect parametric family $\{F_\theta : \theta \in \Theta\}$ for forecasting and the extent to which the goodness-of-fit tests could be expected to help in reducing such losses.

We use three different measures to estimate the 'loss'. Let G denote the true cdf $G_{\beta_0}(x/\Psi_{i+1}(\phi_0))$ of Z_{i+1}, $\hat{G}(x)$ denote its forecast $G_{\hat{\beta}}(x/\widetilde{\Psi}_{i+1}(\hat{\phi}))$, $\hat{F}(x)$ denote the forecast $F_{\hat{\theta}}(x/\widetilde{\Psi}_{i+1}(\hat{\phi}))$ when the model in H_0 is used. Let $A1 = (1/2)\int |\widehat{g}(x) - g(x)|dx$ and $A2 = (1/2)\int |\widehat{f}(x) - g(x)|dx$. By using the term 'misallocation' in a broad sense, we may interpret $A2$ as the proportion of the total probability of 1 that is misallocated by $\widehat{f}$ when the true target is g. We use $L_P := \mathbb{E}(A2)/\mathbb{E}(A1)$ as the first measure of the loss resulting from using the incorrect parametric model specified by H_0.

We also estimated the following two measures of loss: (a) $L_O := \mathbb{E}(B2)/\mathbb{E}(B1)$, and (b) $L_T := \mathbb{E}(C2)/\mathbb{E}(C1)$, where

$$B1 := \max_{a_L \le x \le a_U} \frac{|\widehat{g}(x) - g(x)|}{g(x)}, \quad B2 := \max_{a_L \le x \le a_U} \frac{|\widehat{f}(x) - g(x)|}{g(x)},$$
$$C1 := \max_{b_L \le x \le b_U} \frac{|\widehat{G}(x) - G(x)|}{\{1 - G(x)\}}, \quad C2 := \max_{b_L \le x \le b_U} \frac{|\widehat{F}(x) - G(x)|}{\{1 - G(x)\}},$$
$$a_L = G^{-1}(0.025),\ a_U = G^{-1}(0.975), \quad b_L = G^{-1}(0.9),\ b_U = G^{-1}(0.99).$$

The first quantity L_O measures the extent to which the *ordinate* of the forecast pdf $\widehat{f}(x)$ deviates from the true pdf g, relative to the deviation $|\widehat{g}(x) - g(x)|$, which is due to purely random error. Similarly, the second quantity, L_T, measures the extent to which the forecast of the upper *tail* quantiles of $\widehat{F}(x)$ deviates from the true quantiles of G, relative to the deviation $|\widehat{G}(x) - G(x)|$ in the upper tail of the distribution G.

Results:

Estimates of L_P, L_O, and L_T are given in Table 3. As an example, consider the first entry of 3.1 for L_P in that table. It says that on an average, the probability misallocated by the forecast density because the use of the incorrect parametric family is 3.1 times (= 310%) of what would be incurred had the true parametric family been used. Therefore, the loss in terms of L_P is large. Since the goodness-of-fit tests, for example, the A^2 test, have nearly 100% power, the tests almost certainly point us to the fact that the use of the null model would result in loss.

Table 3 also shows that L_P increases with the power of A^2. Consequently, if the power of the test is low, then the null and the true models are likely to be close, and hence, the loss in terms of L_P is also likely to be low. The estimated values of L_O and L_T are also of the same order of magnitude as those of L_P. Therefore, the use of these tests can be expected to reduce such losses in density/quantile forecasting.

APPENDIX D: Additional results for the empirical example in Section 6

Figure 2 contains empirical cumulative distribution functions of the probability integral transforms of density forecasts based on the six multiplicative error models for the UTX realized volatility series considered in the empirical example in Section 6.

Figure 3 provides the summary plots for the UTX realized volatility series. The first panel gives a plot of UTX realized volatility expressed on a percent annualized scale. An Autocorrelogram is in the second panel. The last panel displays a residual correlogram for the MEM(1,1) model.

References

Salinetti, G., W. Vervaat, & R.J.-B. Wets (1986) On the convergence in probability of random sets (measurable multifunctions). *Mathematics of Operations Research* 11, 420–422.

Straumann, D. & T. Mikosch (2006) Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *The Annals of Statistics* 34, 2449–2495.

van der Vaart, A.W. (1998) *Asymptotic Statistics*, volume 3 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge.

van der Vaart, A.W. & J.A. Wellner (1996) *Weak Convergence and Empirical Processes: With Applications to Statistics*. Springer Series in Statistics. Springer-Verlag, New York.

Figure 2: Empirical cumulative distribution functions of the probability integral transforms of density forecasts of the UTX realized volatility series for the MEM(1,1): $\Psi_i(\phi) = \phi_1 + \phi_2 Z_{i-1} + \phi_3 \Psi_{i-1}(\phi)$, $\phi = (\phi_1, \phi_2, \phi_3)^\top$, when F_θ is Weibull [— —], and mixture of Burr and Generalized Gamma [• • •].

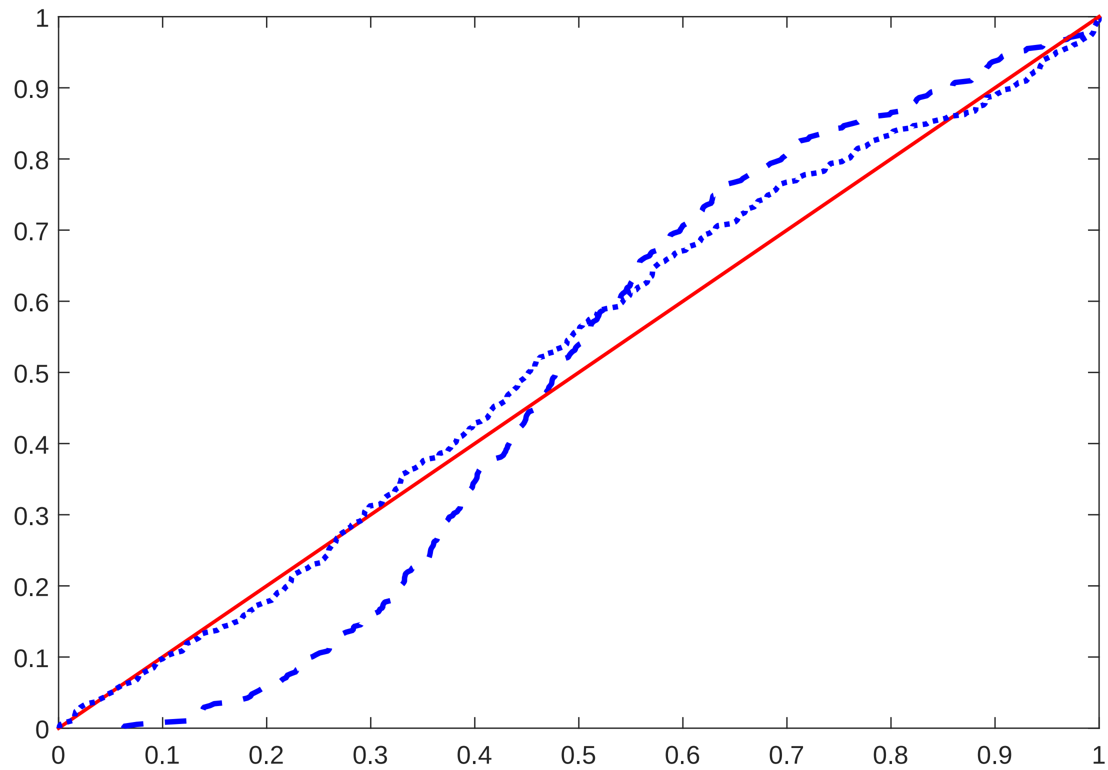

Figure 3: Summary figures for the UTX realized volatility series and the corresponding residuals estimated by the MEM(1,1) model.

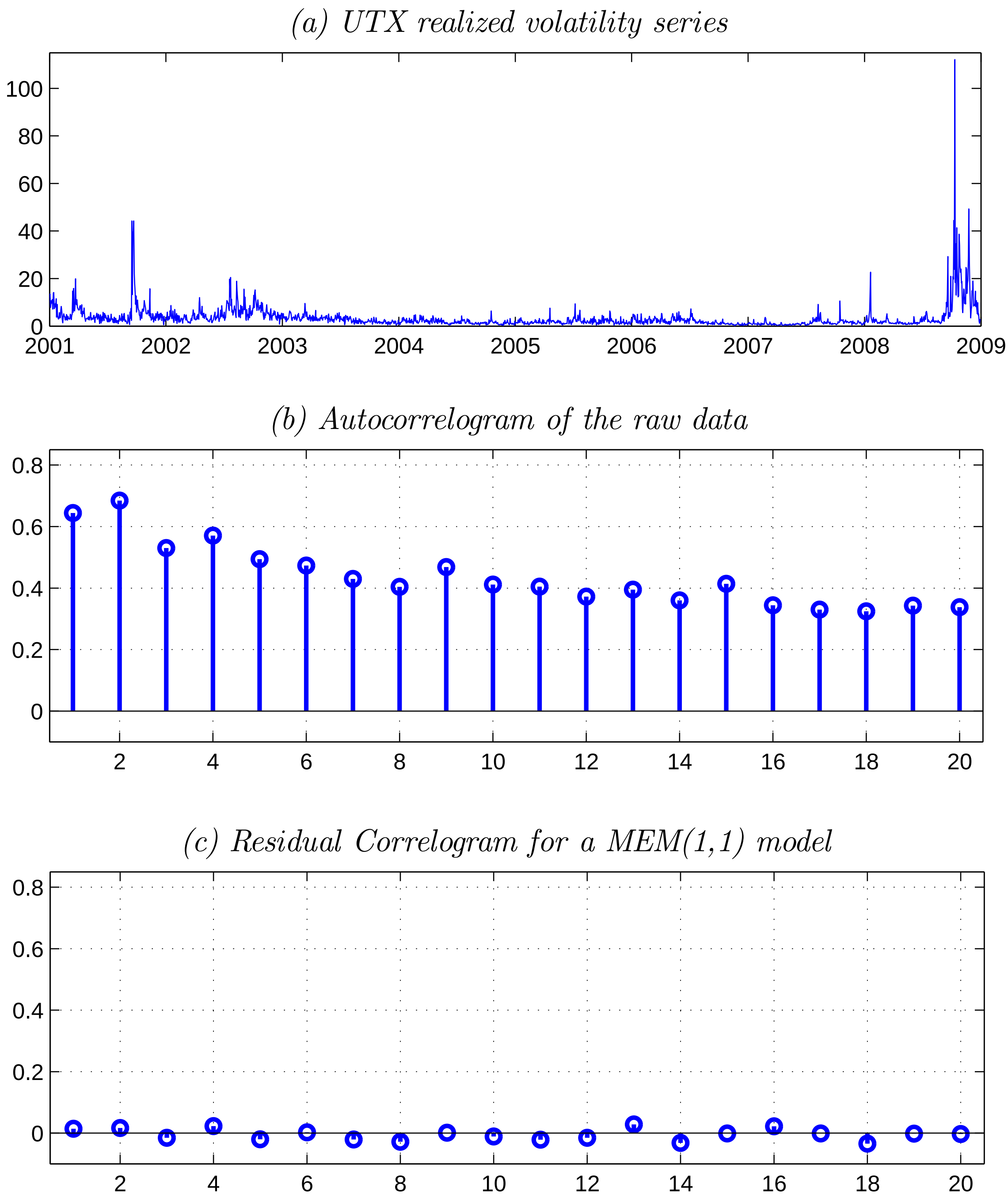

Table 3: Estimated losses of using an incorrect parametric MEM for forecasting.

G_β	F_θ	L_P	L_O	L_T	Power of A^2
Gen. Gamma	Gamma	3.1	2.1	2.9	0.99
	Weibull	2.6	2.0	2.3	0.96
	Exp	3.4	2.4	3.2	0.99
	Burr	1.8	1.2	1.2	0.17
Burr	Gamma	2.7	2.5	3.1	0.99
	Weibull	3.0	3.0	2.9	0.99
	Exp	3.0	3.0	2.8	0.98
	Gen. Gamma	1.5	1.6	1.6	0.43

Note: The results are based on 1000 Monte Carlo replications and the sample size was $n = 1000$. The true DGP is MEM(1,1) for the mean function and G_β for the error distribution. The model under consideration for forecasting, and hence, H_0 is MEM(1,1) for the mean function and F_θ for the error distribution.