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

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

Then, for \( \theta_n \in B \),

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

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

\[
P_n\left( |\bar{U}_n(x) - U_n(x)| > \eta \right) \cap \Pi_n
\]

\[
\leq P_n\left( |\sum_{i=1}^{n} \gamma_{ni} I\{ |\gamma_{ni}| \leq an^{-1/2} \} K_{ni} | > \eta \right) \cap \left( \sum_{i=1}^{n} \mathbb{E} (D_{ni}^2 | F_{ni}) \leq ac \right)
\]

\[
\leq \exp\{-\eta^2/2a(\eta + c)\}.
\]

■
Proof of Lemma 3. Fix $x \geq 0$ and $\varepsilon, \eta > 0$. Choose $c > 0$ and a positive integer $n_1$ such that $P_n(n^{-1}\sum_{i=1}^{n} \gamma_i^2 > c) < \varepsilon$ 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 $L_0 := \sup_{|x| \leq b_0} |H_{\theta_0}(x)| \leq \Delta H_{\theta_0}(1)$. Then, $L_n := \sup_{|x| \leq b_0} |H_{\theta_0}(x)| \leq \Delta H_{\theta_0}(1)$. Therefore, there exists an $n_2 \in \mathbb{N}$ such that $L_n < L_0 + a - \delta < b$ 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 $\sup_{|x| \leq b_n} |H_{\theta_0}(x)| \leq a$.

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

$$
\sup_{|x| \leq b} |H_{\theta_0}(x)| \leq a.
$$

Therefore, $\sup_{\vartheta_n \in B} \sup_{|x| \leq b} |H_{\theta_0}(x)| < (\vartheta_n B)$. Thus, it follows from Lemma 2 that for $\vartheta_n \in B$, 

$$
P_n(|\tilde{U}_n(x) - U_n(x)| > \eta) \leq P_n(|\tilde{U}_n(x) - U_n(x)| > \eta) \cap \Pi_n) + P_n(\Pi_n^c)
$$

$$
\leq \exp\{-\eta^2/2a(\eta + c)\} + P_n(\max |\gamma_i| > an^{-1/2}) + P_n(\max |\alpha_i| > b).
$$

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

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

Since $\varepsilon$ is arbitrary, we have $|\tilde{U}_n(x) - U_n(x)| = o_p(1)$. \qed

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

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 \( x_{N_{\delta}^B} = K_{2\delta}/(1 - b) > 0 \) and \( x_k - x_{k-1} = \delta^2/\{2\|h_\vartheta\|_\infty^B\}, \ k = 1, \ldots, N_{\delta}^B. \) In this partition, there are \( N_{\delta}^B \) sub-intervals 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 \( \delta \) with respect to the pseudo-metric \( \mu_b^B. \) Therefore, \( \mathcal{N}(\delta, b) \leq 2 + N_{\delta}^B, \) and hence,

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

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

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

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 \( 0 \leq b < 1. \) \( \blacksquare \)
Proof of Lemma 6. Fix $M < \infty$. From (C2) and (E3) it follows that
\[
\sup_{\{\phi \in \Phi, \sqrt{n}\|\phi - \phi_n\| \leq M\}} \max_{1 \leq i \leq n} \{v_{ni}(\phi) - (\phi - \phi_n)^\top \lambda_i(\phi_n)\} = o_p(n^{-1/2}).
\]
Further, because max$_{1 \leq i \leq n} n^{-1/2} \lambda_i(\phi_n) = o_p(1)$ and $n^{-1} \sum_{i=1}^n \|\lambda_i(\phi_n)\| = O_p(1)$, one obtains that sup$_{\{\phi \in \Phi, \sqrt{n}\|\phi - \phi_n\| \leq M\}} n^{-1/2} \sum_{i=1}^n |v_{ni}(\phi)| = O_p(1).

Let $A_n = \{(x, \phi) : x \geq 0, \phi \in \Phi$ and $\sqrt{n}\|\phi - \phi_n\| \leq M\}$. Let $a > 0$ be as in Condition (C3). Then, there exists an $n_0 > 0$ such that for all $n > n_0$ and $(\phi, x) \in A_n$, $F_{\theta_n}(x + xv_n(\phi)) - F_{\theta_n}(x) = xv_n(\phi)f'_{\theta_n}(x(1 + \delta^*_n))$, where $\delta^*_n$ is a real number satisfying $|\delta^*_n| < a$.

Because sup$_{\{\phi \in \Phi, \sqrt{n}\|\phi - \phi_n\| \leq M\}} n^{-1/2} \sum_{i=1}^n |v_{ni}(\phi)| = O_p(1)$, it follows from the assumptions in Condition (C3) that sup$_{x, \phi, M} n^{-1/2} \sum_{i=1}^n 2^{-1}\{xv_n(\phi)\}^2 f'_{\theta_n}(x(1 + \delta^*_n)) = o_p(n)$. By Conditions (C2), (C3), (C5), and Assumption (E3), one also obtains that sup$_{x, \phi, M} \{|n^{-1/2} \sum_{i=1}^n xv_n(\phi)f_{\theta_n}(x) - B_n(x)| = o_p(1).$ 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 $\hat{\theta} = \arg \max_{\theta \in \Theta} \sum_{i=1}^n g_\theta(\tilde{\xi}_i)$, under (E2),
\[
0 = \sum_{i=1}^n \dot{g}_{\theta_j}(\tilde{\xi}_i) = \sum_{i=1}^n \dot{g}_{\theta_{0j}}(\tilde{\xi}_i) + (\hat{\theta}_j - \theta_{0j}) \sum_{i=1}^n \left[ (\partial/\partial \theta_j)g_{\theta_{0j}}(\tilde{\xi}_i) |_{\theta_{0j}=\hat{\theta}_j} \right], \tag{B.3}
\]
for some $\hat{\theta}_j$, where $|\hat{\theta}_j - \theta_{0j}| < |\hat{\theta}_j - \theta_{0j}|$. Let
\[
\tilde{H}_n(\theta) = n^{-1} \sum_{i=1}^n \tilde{g}_\theta(\tilde{\xi}_i), \quad H_n(\theta) = n^{-1} \sum_{i=1}^n g_\theta(\xi_i).
\]
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 $\theta_0$, such that
\[
\sup_{\theta \in B} \| \tilde{H}_n(\theta) - H_n(\theta) \|
\leq Kn^{-1} \sum_{i=1}^n |\tilde{\xi}_i - \xi_i| = Kn^{-1} \sum_{i=1}^n Z_i \left| \frac{1}{\psi_i(\hat{\phi})} - \frac{1}{\psi_i(\hat{\phi})} + \frac{1}{\psi_i(\phi_n)} - \frac{1}{\psi_i(\phi_n)} \right|
\leq \frac{K}{\alpha_L n} \sum_{i=1}^n \xi_i \sup_{\phi \in \Phi} \left| \tilde{\psi}_i(\hat{\phi}) - \psi_i(\phi) \right| + \frac{K}{\alpha_L n} \sum_{i=1}^n Z_i \left| \psi_i(\hat{\phi}) - \psi_i(\phi_n) \psi_i^{-1}(\phi_n) + o_p(1) \right|
= S_{n1} + S_{n2} + o_p(1), \text{ say.}
\]
Since \( \sup_{\phi \in \Phi} |\hat{Y}_i(\phi) - \Psi_i(\phi)| \xrightarrow{\text{e.a.s.}} 0 \) 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),

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

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, \( \sup_{\theta \in B} \| \hat{H}_n(\theta) - H_n(\theta) \| = o_p(1) \). Therefore, one obtains by (B.3) that \( \hat{\theta}_j - \theta_0 = n^{-1} \sum_{i=1}^n h_{\theta_0}(\varepsilon_i) + o_p(n^{-1/2}) \).

Now, apply a two-term Taylor series expansion for \( n^{-1} \sum_{i=1}^n h_{\theta_0}(\varepsilon_i) \) 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 \( (\phi_0, \theta_0) \), the subsequence \( (a_n) \) contains a further subsequence \( (r_n) \) such that \( (\hat{\phi}_{r_n}, \hat{\theta}_{r_n}) \xrightarrow{\text{a.s.}} (\phi_0, \theta_0) \). Now, choose a sample path along which \( (\hat{\phi}_{r_n}, \hat{\theta}_{r_n}) \rightarrow (\phi_0, \theta_0) \). Then, it follows from (E3) that \( \hat{\theta}_{r_n}^* - \hat{\theta}_{r_n} = o_p(r_n^{-1}) \) and \( \hat{\phi}_{r_n}^* - \hat{\phi}_{r_n} = O_p(r_n^{-1/2}) \) 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

\[
\hat{\theta}_{r_n,j}^* - \hat{\theta}_{r_n,j} = r_n^{-1} \sum_{i=1}^{r_n} \{h_{\hat{\theta}_{r_n,j}}^*(\varepsilon_i^*) - (\hat{\phi}_{r_n}^* - \hat{\phi}_{r_n})^\top \lambda_i^*(\hat{\phi}_{r_n}^*) h_{\hat{\theta}_{r_n,j}}^*(\varepsilon_i^*)\} + o_p(r_n^{-1/2}).
\]

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 \( o_p(1) \), 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 \( (\phi_0, \theta_0) \), and to work along a fixed sample path for which \( (\hat{\phi}_{r_n}, \hat{\theta}_{r_n}) \rightarrow (\phi_0, \theta_0) \). Hence, for showing that a quantity is \( o_p(1) \) in probability, one may assume without loss of generality that \( (\hat{\phi}, \hat{\theta}) \rightarrow (\phi_0, \theta_0) \) 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 \( (\hat{\phi}, \hat{\theta}) \rightarrow (\phi_0, \theta_0) \).
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

\[ a_i(\phi) := \{\Psi_i^{(m)}(\hat{\phi})/\Psi_i^{(m)}(\phi)\}\{\lambda_i^{(m)}(\phi) - \lambda_i^*(\phi)\}, \]

\[ b_i(\phi) := \lambda_i^*(\phi)\{\Psi_i^{(m)}(\hat{\phi}) - \Psi_i^*(\phi)\}/\Psi_i^{(m)}(\phi), \]

\[ c_i(\phi) := \lambda_i^*(\phi)\{\Psi_i^{(m)}(\hat{\phi})/\Psi_i^*(\phi)\}\{\Psi_i^*(\phi) - \Psi_i^{(m)}(\phi)/\Psi_i^{(m)}(\phi)\}. \]

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 $B_i(\phi) = \{\hat{\Psi}_i^{(m)}(\phi) - \hat{\Psi}_i^*(\phi)\}$. Then, it follows from Conditions (C6) and (E3) that $e_{i1}, e_{i2} \xrightarrow{e.a.s} 0$. Further,

\[ \lambda_i^{(m)}(\phi) - \lambda_i^*(\phi) = \{\hat{\Psi}_i^{(m)}(\phi)/\Psi_i^{(m)}(\phi)\} - \{\hat{\Psi}_i^*(\phi)/\Psi_i^*(\phi)\} = \frac{1}{\Psi_i^{(m)}(\phi)} B_i(\phi) - \lambda_i^*(\phi) \{1/\Psi_i^{(m)}(\phi)\} A_i(\phi). \]

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 $e_{i1}, e_{i2} \xrightarrow{e.a.s} 0$ as $i \to \infty$, then it follows from Lemma 2.1 of Straumann and Mikosch (2006) that $||\lambda_i^{(m)}(\cdot) - \lambda_i^*(\cdot)|| \xrightarrow{e.a.s} 0$, 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

\[ ||a_i(\cdot)|| \leq e_{i1} \alpha^{-1}||\Theta_i^*(\phi)|| ||\lambda_i^{(m)}(\cdot) - \lambda_i^*(\cdot)||, \]

\[ ||b_i(\cdot)|| \leq e_{i1} \alpha^{-1}||\lambda_i^*(\cdot)||, \quad ||c_i(\cdot)|| \leq e_{i1} \alpha^{-2}||\Theta_i^*(\phi)|| ||\lambda_i^*(\cdot)||. \]

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 $\xrightarrow{e.a.s} 0$ 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 $\xi_i \xrightarrow{e.a.s} 0$, then $\sum_{i=1}^n \xi_i\|v_i\|$ 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 $\phi$ is a scalar parameter. Let $0.5 < \eta < 1$. Note that for a given $\delta > 0$, the curve $n^{-1}(\partial/\partial \phi)L_n^\phi(\phi)$ 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 $J(\phi) = -(\partial/\partial \phi)n^{-1}L_n^\phi(\phi)$. Let $\phi_a$ and $\phi_b$ 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 $P_n^*[\sup_{\phi \in B} |\hat{\phi}(\phi)| > c_0] \to 1$ as $n \to \infty$. In view of the mean value theorem, for some $\tilde{\phi}_a$ and $\tilde{\phi}_b$ satisfying $|\tilde{\phi}_a - \hat{\phi}^*| \leq |\phi_a - \hat{\phi}^*|$ and $|\tilde{\phi}_b - \hat{\phi}^*| \leq |\phi_b - \hat{\phi}^*|$, we have

$$n^{-\eta}\delta = S(\phi_a) = S(\hat{\phi}^*) + (\phi_a - \hat{\phi}^*)\hat{S}(\phi_a) = 0 + (\phi_a - \hat{\phi}^*)\hat{S}(\phi_a),$$

$$-n^{-\eta}\delta = S(\phi_b) = S(\hat{\phi}^*) + (\phi_b - \hat{\phi}^*)\hat{S}(\phi_b) = 0 + (\phi_b - \hat{\phi}^*)\hat{S}(\phi_b).$$

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

Because $P_n^*[-\delta < n^\eta\sup_{\phi \in \phi}(\partial/\partial \phi)\{n^{-1}L_n^\phi(\phi) - n^{-1}L_n(\phi)\} < \delta] \to 1$, 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, $P_n^*|\hat{\phi}(\phi) - \phi| \to 1$, and hence, $|\hat{\phi}(\phi) - \phi| \leq |\phi_a - \phi_b| \leq 2n^{-\eta}\delta c_0^{-1}$, with $P_n^*$ probability $\to 1$. Therefore, part (b) follows.

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

$$K_0 \sum_{i=1}^n |Z_i^*(\theta)/\tilde{Z}_i^*(\theta)|$$

$$= K_0 \sum_{i=1}^n |\tilde{Z}_i^*(\theta)|/\tilde{Z}_i^*(\theta) - Z_i^*(\theta)/\tilde{Z}_i^*(\theta)|$$

$$\leq K_0 \sum_{i=1}^n |K_i|/\tilde{Z}_i^*(\theta) - Z_i^*(\theta) + K_2 \tilde{Z}_i^*(\theta)/\tilde{Z}_i^*(\theta) - \tilde{Z}_i^*(\theta)/\tilde{Z}_i^*(\theta)|,$$

where $K_0, K_1$ and $K_2$ are fixed constants. In view of Assumption (E3), the terms $\sup_{\phi \in \phi} |\Psi^*(\phi) - \Psi^*(\phi)|$ and $\sup_{\phi \in \phi} |\tilde{\Psi}^*(\phi) - \tilde{\Psi}^*(\phi)|$ converge to zero (e.a.s) as
Let apply Lemma 6 to complete the proof. ■

By direct substitution, we obtain that follows from a one-term Taylor expansion that

This equality continues to hold with by triangle inequality,

From the proof of Lemma 8, for some \( \rho_{ni} = v_i^* \) that \( |v_i^*| = o_{p_n}(1) \). Now, apply Lemma 6 to complete the proof. ■

Proof of Lemma 11. Let \( v_{ni}^*(m) = n^{1/2} [\Psi_i^*(\hat{\phi}) - \Psi_i^*(\hat{\phi})] / \Psi_i^*(\hat{\phi}) \). 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, \( \max_{1 \leq i \leq n} n^{-1/2} \|\lambda_i^{(m)}(\hat{\phi}) - \lambda_i^*(\hat{\phi})\| = o_{p_n}(1) \), and from Lemma 8, \( n^{1/2} (\hat{\phi}^{(m)} - \hat{\phi}) = O_{p_n}(1) \). Hence, \( \max_{1 \leq i \leq n} |n^{-1/2} v_{ni}^*(m)| = o_{p_n}(1) \). Further, with \( \gamma_{ni} = 1 \) and \( \rho_{ni} = n^{-1/2} v_{ni}^*(m) \), it follows from Lemma 5 that uniformly in \( y \geq 0 \),

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

By direct substitution, we obtain that

\[
\begin{align*}
v_{ni}^*(m) - v_{ni}^* &= n^{1/2} (\hat{\phi}^{(m)} - \hat{\phi})^\top \lambda_i^{(m)}(\hat{\phi}) - n^{1/2} (\hat{\phi}^* - \hat{\phi})^\top \lambda_i^*(\hat{\phi}) + [r_{ni}^*(m) - r_{ni}^*] \\
&= \{n^{1/2} (\hat{\phi}^{(m)} - \hat{\phi})^\top \lambda_i^{(m)}(\hat{\phi}) + n^{1/2} (\hat{\phi}^* - \hat{\phi})^\top [\lambda_i^{(m)}(\hat{\phi}) - \lambda_i^*(\hat{\phi})]\} + [r_{ni}^*(m) - r_{ni}^*].
\end{align*}
\]
Because $n^{1/2}(\hat{\theta}^{(m)} - \hat{\theta}^*) = o_p(n)$, $n^{-1} \sum_{i=1}^n \| \Lambda_i^{(m)}(\hat{\theta}) - \Lambda_i^{*}(\hat{\theta}) \| = o_p(n^{-n})$, and $n^{-1} \sum_{i=1}^n [r^{(m)}_{ni} - r^{*}_{ni}] = O_p(n^{-1/2})$, this yields that $n^{-1} \sum_{i=1}^n [v^{(m)}_{ni} - v^{*}_{ni}] = o_p(n)$. Since $\sup_{\theta \in B, y \geq 0} (1 + y) f_\theta(y) < \infty$ for some open neighbourhood $B$ of $\theta_0$, the first part follows from (B.4). Because $\hat{\theta}^{(m)} - \hat{\theta}^* = o_p(n^{-n})$, the second part follows from a one-term Taylor expansion.

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

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

where $\sigma^2 = \int_{x \geq 0} \{ \hat{f}(x) - f_{\theta_0}(x) \}^2 f_{\theta_0}^{-1}(x) \, dx$. Hence, by the central limit theorem, $\ell_n \overset{d}{\rightarrow} N(-2^{-1} \delta^2 \sigma^2, \delta^2 \sigma^2)$ 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 $G(\cdot)$, $\sup_t |W_n \{ F_{\theta_0}^{-1}(t) \} - G_n(t) | = o_p(1)$, and $\mathbb{E}[G_n(t) \ell_n] = m(t, \theta_0) + o(1)$ under $H_0$, $t \in [0, 1]$, where

$$m(t, \theta) = \delta \int \{ I(\epsilon \leq F^{-1}_\theta(t)) - t \} \, d\hat{F}(\epsilon)$$

$$+ \delta \left[ \int \{ (\partial/\partial \theta) \hat{g}_\theta(y) \}^{-1} \, dF_\theta(y) \right] \int \hat{g}_\theta(\epsilon) \, d\hat{F}(\epsilon) \right\] ^T \hat{F}_\theta(F^{-1}_\theta(t))$$

$$= \delta [\hat{F}(F^{-1}_\theta(t)) - t] + \delta \left[ \int \{ \hat{g}_\theta(y) \}^{-1} \, dF_\theta(y) \right] \int \hat{g}_\theta(\epsilon) \, d\hat{F}(\epsilon) \right\] ^T \hat{F}_\theta(F^{-1}_\theta(t)).$$

Thus, $m(\cdot, \theta_0)$ is the same as $m_\mathfrak{a}(\cdot)$ 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 \( \tilde{W}_n \circ F_{\theta_0}^{-1}(\cdot) \) converges weakly to \( W_0(\cdot) \) in \( D[0,1] \), under \( H_{an} \), where \( W_0(\cdot) = m_\alpha(\cdot) + G(\cdot) \). By assumption, \( \int \tilde{g}_{\theta_0}(\varepsilon) \, d\tilde{F}(\varepsilon) = 0 \) only if \( \tilde{F} = F_{\theta_0} \). Further, \( \tilde{F} \neq F_{\theta_0} \) and
\[
[t - \tilde{F}(F_{\theta_0}^{-1}(t))] \neq \left[ \int \{\tilde{g}_{\theta_0}(y)\}^{-1} \, dF_{\theta_0}(y) \right] \, \tilde{F}_{\theta_0}(F_{\theta_0}^{-1}(t)), \quad t \in [0,1].
\]
Hence, \( m_\alpha \neq 0 \) for \( \delta > 0 \).

Under \( H_{an} \), \( \hat{\theta} \to \theta_0 \) in probability, and \( \theta_0 \) 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 \( \tilde{W}_n \circ F_{\theta_0}^{-1}(\cdot) \) 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 \( \sup_t |\tilde{W}_n(\cdot) - G_n(t)| = o_p(1) \) 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
\[
\ell_i(\phi) = \log \tilde{\Psi}_i(\phi) + \frac{[\psi_i(\phi) + n^{-1/2}r_i]/\psi_i(\phi)}{\psi_i(\phi)} \varepsilon_i = \log \tilde{\Psi}_i(\phi) + \frac{[\psi_i(\phi_0)]/\psi_i(\phi)}{\psi_i(\phi)} + n^{-1/2} r_i \varepsilon_i.
\]
Let \( \hat{\phi}^{(1c)} = \arg \min_{\phi \in \Phi} \sum_{i=1}^n \kappa_i(\phi) \), where \( \kappa_i(\phi) = \log \tilde{\Psi}_i(\phi) + [\psi_i(\phi_0)]/\tilde{\Psi}_i(\phi) \). In view of Assumption (E1), \( \hat{\phi}^{(1c)} \overset{p}{\to} \phi_0 \) as \( n \to \infty \). Further, we have that
\[
\sum_{i=1}^n (\partial/\partial \phi) \{\ell_i(\phi) - \kappa_i(\phi)\} = -n^{-1/2} \sum_{i=1}^n r_i \varepsilon_i \tilde{\lambda}_i(\phi)/\tilde{\Psi}_i(\phi) = 0 \quad \text{as } n \to \infty.
\]
Therefore, under \( H_{bn} \), for any \( 1/4 < \eta < 1/2 \),
\[
n^n \sup_{\phi \in \Phi} \| (\partial/\partial \phi) \{n^{-1} \sum_{i=1}^n \ell_i(\phi) - n^{-1} \sum_{i=1}^n \kappa_i(\phi)\} \| \overset{a.s.}{\to} 0 \quad \text{as } n \to \infty.
\]
Consequently, \( \hat{\phi} - \hat{\phi}^{(1c)} \overset{p}{\to} 0 \) as \( n \to \infty \), and hence, \( \hat{\phi} \overset{p}{\to} \phi_0 \). 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, $\hat{W}_n^* \circ F_{\hat{\theta}}^{-1(\cdot)}$ converges weakly to $G(\cdot)$ under $H_{bn}$ [in probability]. Because $T_j = h_j(\hat{W}_n \circ F_{\hat{\theta}_b}^{-1}) + o_p(1)$ and $T_j^{(m)} = h_j(\hat{W}_n^{* \cdot \cdot \cdot} \circ F_{\hat{\theta}}^{-1}) + o_p(1)$ [in probability] $(j = 1, \ldots, 5)$, the proof follows from Lemmas 11 and 13, and the continuous mapping theorem.

**Proof of Proposition 3.** Let

$$
\ell_n := -\sum_{i=1}^n \left[ \log \{\psi_i(\phi_0) + r_i/\sqrt{n} \} - \log \{\psi_i(\phi_0)\} + \log \{f_{(n)}(\varepsilon_i)/f_{b_0}(\varepsilon_i)\} \right].
$$

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 $\sup_t |\hat{W}_n \{F_{\hat{\theta}_b}^{-1}(t)\} - G_n(t)| = o_p(1)$ 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

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

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

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

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

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

Further, $\hat{\phi} - \hat{\phi}^{(2c)} \overset{p}{\to} 0$ as $n \to \infty$, and hence, $\hat{\phi} \overset{p}{\to} \phi_0$ as $n \to \infty$. Note that under $H_{cn}$, $\hat{\theta} \to \theta_0$ as $n \to \infty$, where $\theta_0$ is the true value satisfying $F^0 = F_{\theta_0}$ under $H_0$. As in the proof of Lemma 14, $\hat{W}_n^* \circ F_{\hat{\theta}}^{-1(\cdot)}$ 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 $\Psi_i$. 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 $G_{\hat{\beta}}(x/\hat{\Psi}_{i+1}(\hat{\phi}))$, $\hat{F}(x)$ denote the forecast $F_{\hat{\theta}}(x/\hat{\Psi}_{i+1}(\hat{\phi}))$ when the model in $H_0$ is used. Let $A1 = (1/2) \int |\hat{g}(x) - g(x)|dx$ and $A2 = (1/2) \int |\hat{f}(x) - g(x)|dx$. 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 $\hat{f}$ 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) \( L_T := \frac{\mathbb{E}(C2)}{\mathbb{E}(C1)} \), where

\[
B1 := \max_{a_L \leq x \leq a_U} \frac{|\hat{g}(x) - g(x)|}{g(x)}, \quad B2 := \max_{a_L \leq x \leq a_U} \frac{|\hat{f}(x) - g(x)|}{g(x)},
\]

\[
C1 := \max_{b_L \leq x \leq b_U} \frac{|\hat{G}(x) - G(x)|}{1 - G(x)}, \quad C2 := \max_{b_L \leq x \leq b_U} \frac{|\hat{F}(x) - G(x)|}{1 - G(x)}.
\]

\[
a_L = G^{-1}(0.025), \quad a_U = G^{-1}(0.975), \quad b_L = G^{-1}(0.9), \quad b_U = G^{-1}(0.99).
\]

The first quantity \( L_O \) measures the extent to which the ordinate of the forecast pdf \( \hat{f}(x) \) deviates from the true pdf \( g \), relative to the deviation \( |\hat{g}(x) - g(x)| \), 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 \( \hat{F}(x) \) deviates from the true quantiles of \( G \), relative to the deviation \( |\hat{G}(x) - G(x)| \) 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 Autocorrelationogram is in the second panel. The last panel displays a residual correlogram for the MEM(1,1) model.

References


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 \( F_\theta \) 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.

(a) UTX realized volatility series

(b) Autocorrelogram of the raw data

(c) Residual Correlogram for a MEM(1,1) model
Table 3: Estimated losses of using an incorrect parametric MEM for forecasting.

<table>
<thead>
<tr>
<th>$G_\beta$</th>
<th>$F_\theta$</th>
<th>$L_P$</th>
<th>$L_O$</th>
<th>$L_T$</th>
<th>Power of $A^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gen. Gamma</td>
<td>Gamma</td>
<td>3.1</td>
<td>2.1</td>
<td>2.9</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Weibull</td>
<td>2.6</td>
<td>2.0</td>
<td>2.3</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>Exp</td>
<td>3.4</td>
<td>2.4</td>
<td>3.2</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Burr</td>
<td>1.8</td>
<td>1.2</td>
<td>1.2</td>
<td>0.17</td>
</tr>
<tr>
<td>Burr</td>
<td>Gamma</td>
<td>2.7</td>
<td>2.5</td>
<td>3.1</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Weibull</td>
<td>3.0</td>
<td>3.0</td>
<td>2.9</td>
<td>0.99</td>
</tr>
<tr>
<td></td>
<td>Exp</td>
<td>3.0</td>
<td>3.0</td>
<td>2.8</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>Gen. Gamma</td>
<td>1.5</td>
<td>1.6</td>
<td>1.6</td>
<td>0.43</td>
</tr>
</tbody>
</table>

*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 $G_\beta$ for the error distribution. The model under consideration for forecasting, and hence, $H_0$ is MEM(1,1) for the mean function and $F_\theta$ for the error distribution.