A Supplement

In this Supplement we examine testing linearity against commonly applied STAR models and also provides simulation evidence of our methodology. We also demonstrate how Hansen's (1996) weighted bootstrap is applied to enhance the applicability of our methodology. Finally, we provide the proofs of the theoretical results in the paper

A.1 Monte Carlo Experiments and Application of the Weighted Bootstrap

A.1.1 Monte Carlo Experiments Using the ESTAR Model

To simplify our illustration, we assume that for all $t = 1, 2, \ldots$, $u_t \sim$ IID $N(0, \sigma_*^2)$ and $y_t = \pi_* y_{t-1} + u_t$ with $\pi_* = 0.5$. Under this DGP, we specify the following first-order ESTAR model: $\mathcal{M}_{ESTAR} := \{\pi y_{t-1} + \theta y_{t-1}\{1 - \exp[-\gamma y_{t-1}^2]\} : \pi \in \Pi,\ \theta \in \Theta, \text{ and } \gamma \in \Gamma\}$. The model does not contain an intercept, and the transition variable is y_{t-1}. The nonlinear function $f_t(\gamma) = 1 - \exp(-\gamma y_{t-1}^2)$ is defined on Γ which is compact and convex, and the exponential function is analytic. This means that the QLR test statistic is generically comprehensively revealing. To identify the model it is assumed that $\gamma_* > 0$. In our model set-up, we allow 0 to be included in Γ. The nonlinear function $f_t(\cdot)$ satisfies $f_t(0) = 0$. Given this model, the following hypotheses are of interest: $\mathcal{H}_0' : \exists \pi \in \mathbb{R}, \mathbb{P}(\mathbb{E}[y_t|y_{t-1}] = \pi y_{t-1}) = 1$; vs. $\mathcal{H}_1' : \forall \pi \in \mathbb{R}, \mathbb{P}(\mathbb{E}[y_t|y_{t-1}] = \pi y_{t-1}) < 1$. Two parameter restrictions make $\mathcal{H}_0'$ valid: either $\theta_* = 0$ or $\gamma_* = 0$. The sub-hypotheses are thus $\mathcal{H}_{01}' : \theta_* = 0$ and $\mathcal{H}_{02}' : \gamma_* = 0$.

We first examine the null distribution of the QLR test under $\mathcal{H}_{01}'$. By Theorem 1, the null limit distribution of this test statistic is given as $QLR_n^{(1)} = \sup_{\gamma \in \Gamma} \frac{1}{\widehat{\sigma}_{n,0}^2} (u' M F(\gamma) Z)^2 / \ Z' F(\gamma) M F(\gamma) Z \Rightarrow \sup_{\gamma \in \Gamma} \widetilde{\mathcal{G}}_1^2(\gamma)$, where $\widetilde{\mathcal{G}}_1(\cdot)$ is a mean-zero Gaussian process with the covariance structure $\widetilde{\rho}_1(\gamma, \widetilde{\gamma}) = c_1^{-1/2}(\gamma, \gamma) \widetilde{k}_1(\gamma, \widetilde{\gamma})$ $\widetilde{c}_1^{-1/2}(\widetilde{\gamma}, \widetilde{\gamma})$ with $\widetilde{k}_1(\gamma, \widetilde{\gamma}) = \widetilde{c}_1(\gamma, \widetilde{\gamma}) = \sigma_*^2 \{\mathbb{E}[y_t^2 \exp(-(\gamma + \widetilde{\gamma}) y_t^2)] - \mathbb{E}[y_t^2 \exp(-\gamma y_t^2)] \mathbb{E}[y_t^2 \exp(-\widetilde{\gamma} y_t^2)] / \mathbb{E}[y_t^2]\}$. Furthermore, under $\mathcal{H}_{01}'$, y_t is normally distributed with $\mathbb{E}[y_t] = 0$ and $\mathrm{var}[y_t] = \sigma_y^2 := \sigma_*^2/(1 - \pi_*^2)$, so that y_t^2 follows the gamma distribution with shape parameter $1/2$ and scale parameter $2\sigma_*^2/(1 - \pi_*^2)$. Define $\widetilde{m}(\gamma) := (1 + 2\sigma_*^2/1 - \pi_*^2 \gamma)^{-\frac{1}{2}}$, and $\widetilde{h}(\gamma, \widetilde{\gamma}) := \frac{1}{\sigma_y^2}([(1 + 2\sigma_y^2 \gamma)(1 + 2\sigma_y^2 \widetilde{\gamma}) / \{1 + 2\sigma_y^2(\gamma + \widetilde{\gamma})\}]^{3/2} - 1)$. Note that $\widetilde{m}(\gamma) = \mathbb{E}[\exp(-\gamma y_t^2)]$, so that $\mathbb{E}[y_t^2 \exp(-\gamma y_t^2)] = -\widetilde{m}'(\gamma)$. As a result, $\widetilde{\rho}_1(\gamma, \widetilde{\gamma})$ is further simplified to $\widetilde{k}_1(\gamma, \widetilde{\gamma}) = \sigma_*^2 \widetilde{m}'(\gamma) \widetilde{m}'(\widetilde{\gamma}) \widetilde{h}(\gamma, \widetilde{\gamma})$, and $\widetilde{\rho}_1(\gamma, \widetilde{\gamma}) = \widetilde{c}_1^{-1/2}(\gamma, \gamma) \widetilde{k}_1(\gamma, \widetilde{\gamma}) \widetilde{c}_1^{-1/2}(\widetilde{\gamma}, \widetilde{\gamma}) = \widetilde{h}^{-1/2}(\gamma, \gamma) \widetilde{h}(\gamma, \widetilde{\gamma}) \widetilde{h}^{-1/2}(\widetilde{\gamma}, \widetilde{\gamma})$.

We next examine the limit distribution of the QLR test statistic under $\mathcal{H}_{02}'$: $\gamma_* = 0$. The first-order derivative $(\partial/\partial\gamma)f_t(\gamma) = y_{t-1}^2\exp(-\gamma y_{t-1}^2)$, which is different from zero even when $\gamma = 0$, so that in this case $\kappa = 1$. Thus, we can apply the second-order Taylor expansion to obtain the limit distribution of the QLR test statistic under $\mathcal{H}'_{02}$. As a result, $\overline{QLR}_n^{(2)}(\theta) = \frac{1}{\widehat{\sigma}^2_{n,0}}(\theta' G'_\kappa u)^2/\theta' G'_\kappa G_\kappa\theta$, where $\theta' G'_\kappa u = \theta[\sum y_{t-1}^3 u_t - \sum y_{t-1}^4 \sum y_{t-1}u_t / \sum y_{t-1}^2]$ and $\theta' G'_\kappa G_\kappa \theta = \theta^2[\sum y_{t-1}^6 - (\sum y_{t-1}^4)^2/\sum y_{t-1}^2]$. Here, θ is a scalar, so that cancels out, and it follows that $QLR_n^{(2)} \Rightarrow \widetilde{\mathcal{G}}_2^2$, where $\widetilde{\mathcal{G}}_2 \sim N(0,1)$.

These two separate results can be combined, which means that we can examine the limit distribution of the QLR test under $\mathcal{H}'_0$. We have $QLR_n \Rightarrow \sup_{\gamma\in\Gamma}\widetilde{\mathcal{G}}^2(\gamma)$, where $\widetilde{\mathcal{G}}(\gamma) = \widetilde{\mathcal{G}}_1(\gamma)$, if $\gamma \neq 0$; and $\widetilde{\mathcal{G}}(\gamma) = \widetilde{\mathcal{G}}_2$, otherwise, and $\mathbb{E}[\widetilde{\mathcal{G}}(\gamma)\widetilde{\mathcal{G}}(\widetilde{\gamma})] = \widetilde{\rho}_1(\gamma,\widetilde{\gamma})$, if $\gamma \neq 0, \widetilde{\gamma} \neq 0$; $\mathbb{E}[\widetilde{\mathcal{G}}(\gamma)\widetilde{\mathcal{G}}(\widetilde{\gamma})] = 1$, if $\gamma = 0, \widetilde{\gamma} = 0$; and $\mathbb{E}[\widetilde{\mathcal{G}}(\gamma)\widetilde{\mathcal{G}}(\widetilde{\gamma})] = \widetilde{\rho}_3(\gamma)$, if $\gamma \neq 0, \widetilde{\gamma} = 0$ with $\widetilde{\rho}_3(\gamma) := \mathbb{E}[\widetilde{\mathcal{G}}_1(\gamma)\widetilde{\mathcal{G}}_2] = \sqrt{6}\sigma_y^2\gamma/\{\widetilde{h}^{1/2}(\gamma,\gamma)(1+2\sigma_y^2\gamma)\}$ such that $\widetilde{\rho}_3^2(\gamma) = \lim_{\widetilde{\gamma}\to 0}\widetilde{\rho}_1^2(\gamma,\widetilde{\gamma}) = (\sqrt{6}\sigma_y^2\gamma/\{\widetilde{h}^{1/2}(\gamma,\gamma)(1+2\sigma_y^2\gamma)\})^2$. Thus, we conclude that $QLR_n \Rightarrow \sup_\gamma \widetilde{\mathcal{G}}^2(\gamma)$, which agrees with Theorem 3.

The null limit distribution can be approximated numerically by simulating a distributionally equivalent Gaussian process. To do this we present the following lemma:

Lemma A. 1. *If $\{z_k : k = 0,1,2,\ldots\}$ is an IID sequence of standard normal random variables, $\widetilde{\mathcal{G}}(\cdot) \stackrel{\mathrm{d}}{=} \overline{\mathcal{G}}(\cdot)$, where for each $\gamma \in \Gamma := \{\gamma \in \mathbb{R} : \gamma \geq 0\}$, $\overline{\mathcal{G}}(\gamma) := \sum_{k=1}^\infty c(\gamma)a^k(\gamma)\ [(-1)^k\binom{-3/2}{k}]^{1/2}z_k$, $c(\gamma) := \{\sum_{k=1}^\infty (-1)^k a^{2k}(\gamma)\binom{-3/2}{k}\ \}^{-1/2}$, and $a(\gamma) := 2\sigma_y^2\gamma/(1+2\sigma_y^2\ \gamma)$.* □

Note that the term $(-1)^k\binom{-3/2}{k}$ in Lemma A. 1 is always positive irrespective of k, and for any γ,

$$\lim_{k\to\infty}\mathrm{var}\left[a^k(\gamma)\left((-1)^k\binom{-3/2}{k}\right)^{1/2}z_k\right] = \lim_{k\to\infty}a^{2k}(\gamma)(-1)^k\binom{-3/2}{k} = 0 \qquad \text{(A.1)}$$

and $\widetilde{h}(\gamma,\gamma) = \sum_{k=1}^\infty a^{2k}(\gamma)(-1)^k\binom{-3/2}{k}$. Using these facts Lemma A. 1 shows that for any γ and $\widetilde{\gamma} \neq 0$, $\mathbb{E}[\overline{\mathcal{G}}(\gamma)\overline{\mathcal{G}}(\widetilde{\gamma})] = \widetilde{\rho}_1(\gamma,\widetilde{\gamma})$. Here, the non-negative parameter space condition for Γ is necessary for $\overline{\mathcal{G}}(\cdot)$ to be properly defined on Γ. Without this condition, $\overline{\mathcal{G}}(\gamma)$ cannot be properly generated. We note that $\lim_{\gamma\downarrow 0}\overline{\mathcal{G}}(\gamma) \stackrel{\text{a.s.}}{=} z_1$, so that if we let $z_1 = \overline{\mathcal{G}}_2$, $\mathbb{E}[\overline{\mathcal{G}}(\gamma)\overline{\mathcal{G}}_2] = \sqrt{6}\sigma_y^2\gamma\widetilde{h}^{-1/2}(\gamma,\gamma)(1+2\sigma_y^2\gamma)^{-1} = \widetilde{\rho}_3(\gamma)$. It follows that the distribution of $\widetilde{\mathcal{G}}(\cdot)$ can be simulated by iteratively generating $\overline{\mathcal{G}}(\cdot)$. In practice,

$$\overline{\mathcal{G}}(\gamma;K) := \sum_{k=1}^K a^k(\gamma)\left[(-1)^k\binom{-3/2}{k}\right]^{1/2} z_k \Bigg/ \sqrt{\sum_{k=1}^K a^{2k}(\gamma)(-1)^k\binom{-3/2}{k}}$$

is generated by choosing K to be sufficiently large . By (A.1), if this is the case, the difference between the distributions of $\overline{\mathcal{G}}(\cdot)$ and $\overline{\mathcal{G}}(\cdot;K)$ becomes negligible.

We now conduct Monte Carlo experiment and examine the empirical distributions of the QLR statistic under several different environments. First, we consider four different parameter spaces: $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, and $\Gamma_4 = [0, 5]$. They are selected to examine how the null limit distribution of the QLR test is influenced by the choice of Γ. We obtain the limit distribution by simulating $\sup_{\gamma \in \Gamma} \bar{\mathcal{G}}^2(\gamma; K)$ 5,000 times with $K = 2,000$, where Γ is in turn Γ_1, Γ_2, Γ_3, and Γ_4. Second, we study how the empirical distribution of the QLR test statistic changes with the sample size. We consider the sample sizes $n = 100$, $1,000$, $2,000$, and $5,000$.

Figure A.1 summarises the simulation results and shows that the empirical distribution approaches the null limit distribution under different parameter space conditions. We also provide the estimates of the probability density functions next to the empirical distributions. For every parameter space considered, the empirical rejection rates of the QLR test statistics are most accurate when $n = 2,000$. The empirical rejection rates are closer to the nominal levels when the parameter space is small. This result is significant when $n = 100$: the empirical rejection rates for $\Gamma = \Gamma_1$ are closer to the nominal ones than when $\Gamma = \Gamma_4$. Nonetheless, this difference becomes negligible as the sample size increases. The empirical rejection rates obtained using $n = 2,000$ are already satisfactorily close to the nominal levels, and this result is more or less similar to that from 5,000 observations. This suggests that the theory in Section 2 is effective for the ESTAR model. Considering even larger parameter spaces for γ yielded similar results, so they are not reported here.

A.1.2 Illustration Using the LSTAR Model

As another illustration, we consider testing against the first-order LSTAR model. We assume that the data-generating process is $y_t = \pi_* y_{t-1} + u_t$ with $\pi_* = 0.5$ and $u_t = \ell_t$ with probability $1 - \pi_*^2$; and $u_t = 0$ with probability π_*^2, where $\{\ell_t\}_{t=1}^n$ follows the Laplace distribution with mean 0 and variance 2. Under this assumption, y_t follows the same distribution as ℓ_t that makes the algebra associated with the LSTAR model straightforward. For example, the covariance kernel of the Gaussian process associated with the null limit distribution of the QLR test statistic is analytically obtained thanks to this distributional assumption. This data-generating process is a variation of the exponential autoregressive model in Lawrence and Lewis (1980). Their exponential distribution is replaced by the Laplace distribution to allow y_t to obtain negative values.

Given this DGP, the first-order LSTAR model for $\mathbb{E}[y_t|y_{t-1}, y_{t-2}, \ldots]$ is defined as follows: $\mathcal{M}^0_{LSTAR} := \{\pi y_{t-1} + \theta y_{t-1}\{1 + \exp(-\gamma y_{t-1})\}^{-1} : \pi \in \Pi,\ \theta \in \Theta, \text{ and } \gamma \in \Gamma := [0, \bar{\gamma}]\}$. The nonlinear logistic function $\{1 + \exp(-\gamma y_{t-1})\}^{-1}$ contains an exponential function. It is therefore analytic, and this fact delivers a consistent power for the QLR test statistic. Note, however, that for $\gamma = 0$ the value of the logistic function equals 1/2. This difficulty is avoided by subtracting 1/2 from the logistic function when carrying out the test, viz., $\mathcal{M}_{LSTAR} := \{\pi y_{t-1} + \theta y_{t-1}\{[1+\exp(-\gamma y_{t-1})]^{-1} - 1/2\} : \pi \in \Pi,\ \theta \in \Theta,$ and $\gamma \in \Gamma := [0, \bar{\gamma}]\}$. By the invariance principle, this shift does not affect the null limit distribution of the QLR test statistic. We here let $\gamma \geq 0$ so that transition function is bounded, which modifies the limit space of ς_n into $\mathbb{R}^+$. The null and the alternative hypotheses are identical to those in the ESTAR case.

Before proceeding, note that $\{1+\exp(-\gamma y_{t-1})\}^{-1} - \frac{1}{2} = \frac{1}{2}\tanh\left(\frac{\gamma y_{t-1}}{2}\right)$. Using the hyperbolic tangent function as in Bacon and Watts (1971) makes it easy to find a Gaussian process that is distributionally equivalent to the Gaussian process obtained under the null.

Using this fact, the limit distribution of QLR test statistic under $\mathcal{H}'_{01}$ is derived as in before. By Theorem 1, $QLR_n^{(1)} = \sup_{\gamma\in\Gamma} \frac{1}{\widehat{\sigma}^2_{n,0}}(u'MF^2(\gamma)Z)/Z'F(\gamma)MF(\gamma)Z \Rightarrow \sup_{\gamma\in\Gamma} \ddot{\mathcal{G}}_1^2(\gamma)$, where $\ddot{\mathcal{G}}_1(\cdot)$ is a mean-zero Gaussian process with the covariance structure $\ddot{\rho}_1(\gamma, \widetilde{\gamma}) := \ddot{c}_1^{-1/2}(\gamma,\gamma)\ddot{k}_1(\gamma,\ \widetilde{\gamma})\ddot{c}_1^{-1/2}(\widetilde{\gamma},\widetilde{\gamma})$. The function $\ddot{k}_1(\gamma,\tilde{\gamma})$ is equivalent to $\ddot{c}_1(\gamma,\tilde{\gamma})$ by the conditional homoskedasticity condition, and for each nonzero γ and $\widetilde{\gamma}$, we now obtain that $\ddot{k}_1(\gamma,\tilde{\gamma}) = \frac{1}{4}\mathbb{E}[y_t^2\tanh(\frac{\gamma y_t}{2})\tanh(\frac{\widetilde{\gamma} y_t}{2})] - \frac{1}{4}\mathbb{E}[y_t^2\tanh(\frac{\gamma y_t}{2})]\mathbb{E}[y_t^2\tanh(\frac{\widetilde{\gamma} y_t}{2})]/\ \mathbb{E}[y_t^2]$. In the proof of Lemma A. 2 given below, we further show that $\ddot{k}_1(\gamma,\widetilde{\gamma}) = \sum_{n=1}^{\infty} b_n(\gamma)b_n\,(\widetilde{\gamma})$, where $b_1(\gamma) := \frac{1}{\sqrt{2}}(1-2a(\gamma))$ with $a(\gamma) := \sum_{k=1}^{\infty}(-1)^{k-1}/(1+\gamma k)^3$ and for $n = 2, 3, \ldots$, $b_n(\gamma) := \sqrt{n(n+1)}\sum_{k=1}^{\infty}(-1)^{k-1}$ $(\gamma k)^{n-1}/(1+\gamma k)^{n+2}$.

Next, we derive the limit distribution of the QLR test statistic under $\mathcal{H}'_{02}$. Note that for $\gamma = 0$, $(\partial/\partial\gamma)f_t(\gamma)$ $= y_{t-1}\exp(\gamma y_{t-1})/[1+\exp(-\gamma y_{t-1})]^2 \neq 0$, implying that κ is unity as for the ESTAR case, so that we can apply a second-order Taylor expansion to obtain the limit distribution of the QLR test statistic under $\mathcal{H}'_{02}$: $\overline{QLR}_n^{(2)}(\theta) = \frac{1}{\widehat{\sigma}^2_{n,0}}(\theta' G'_\kappa u)^2/(\theta' G'_\kappa G_\kappa \theta)$, where, similarly to the ESTAR case, $\theta' G'_\kappa u = \frac{\theta}{4}[\sum y_{t-1}^2 u_t - \sum y_{t-1}^3 \sum y_{t-1}u_t/\sum y_{t-1}^2]$ and $\theta' G'_\kappa G_\kappa \theta = \frac{\theta^2}{16}[\sum y_{t-1}^4 - (\sum y_{t-1}^3)^2/\sum y_{t-1}^2]$. From this, it follows that $QLR_n^{(2)} \Rightarrow \ddot{\mathcal{G}}_2^2$, where $\ddot{\mathcal{G}}_2 \sim N(0,1)$.

Therefore, we conclude that $QLR_n \Rightarrow \sup_\gamma \ddot{\mathcal{G}}^2(\gamma)$, where $\ddot{\mathcal{G}}(\gamma) := \ddot{\mathcal{G}}_1(\gamma)$, if $\gamma \neq 0$; and $\ddot{\mathcal{G}}(\gamma) := \mathcal{G}_2$, otherwise. The limit variance of $\ddot{\mathcal{G}}(\gamma)$ is given as $\ddot{\rho}(\gamma,\widetilde{\gamma}) := \mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})]$ such that $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \ddot{\rho}_1(\gamma,\widetilde{\gamma})$ if $\gamma \neq 0$ and $\widetilde{\gamma} \neq 0$; $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = 1$, if $\gamma = 0$ and $\widetilde{\gamma} = 0$; and $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \ddot{\rho}_3(\gamma)$, if $\gamma \neq 0$ and $\widetilde{\gamma} = 0$, where $\ddot{\rho}_3(\gamma) := \mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}_2] = \ddot{k}_1^{-1/2}(\gamma,\gamma)\ddot{r}_1(\gamma)\ddot{q}^{-1/2}$ with $\ddot{r}_1(\gamma) := \frac{1}{2}\mathbb{E}\left[y_{t-1}^3\tanh\left(\frac{\gamma y_{t-1}}{2}\right)\right]$ and $\ddot{q} := \mathbb{E}[y_t^4] - \mathbb{E}[y_t^3]^2/\mathbb{E}[y_t^2]$. From this it follows that $QLR_n \Rightarrow \sup_{\gamma\in\Gamma}\ddot{\mathcal{G}}^2(\gamma)$. Furthermore, $\mathbb{E}[y_t^3] = 0$ and $\mathbb{E}[y_t^4] = 24$ given our DGP, so that

$$\ddot{\rho}_3(\gamma) = \frac{\mathbb{E}[y_t^3\tanh(\gamma y_t/2)]}{4\sqrt{6}\ddot{k}_1^{1/2}(\gamma,\gamma)}. \tag{A.2}$$

Here, we note that

$$\mathbb{E}[y_t^3 \tanh(\gamma y_t/2)] = \frac{1}{8\gamma^4}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right] \tag{A.3}$$

by some tedious algebra assisted by Mathematica, where $P_G(n,x)$ is the polygamma function: $P_G(n,x) := d^{n+1}/d\,x^{n+1}\log(\Gamma(x))$. Inserting (A.3) into (A.2) yields

$$\ddot{\rho}_3(\gamma) = \frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1^{1/2}(\gamma,\gamma)}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right]. \tag{A.4}$$

In addition, we show in Lemma A. 3 given below that applying L'Hôpital's rule iteratively yields that

$$\lim_{\widetilde{\gamma}\downarrow 0}\ddot{\rho}_1^2(\gamma,\widetilde{\gamma}) = \left[\frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1^{1/2}(\gamma,\gamma)}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right]\right]^2. \tag{A.5}$$

This fact implies that $\mathrm{plim}_{\gamma\downarrow 0}\ddot{\mathcal{G}}_1^2(\gamma) = \ddot{\mathcal{G}}_2^2$. That is, the weak limit of the QLR test statistic under $\mathcal{H}'_{02}$ can be obtained from $\ddot{\mathcal{G}}_1^2(\cdot)$ by letting γ converging to zero, so that $QLR_n \Rightarrow \sup_{\gamma\in\Gamma}\ddot{\mathcal{G}}_1^2(\gamma)$ under $\mathcal{H}'_0$.

Next, we derive another Gaussian process that is distributionally equivalent to $\ddot{\mathcal{G}}(\cdot)$ and conduct Monte Carlo simulations using it. The process is presented in the following lemma.

Lemma A. 2. *If $\{z_k\}_{k=1}^{\infty}$ is an IID sequences of standard normal random variables, then for each γ and $\widetilde{\gamma} \in \Gamma := \{\gamma \in \mathbb{R} : \gamma \geq 0\}$, $\ddot{\mathcal{G}}(\cdot) \overset{\mathrm{d}}{=} \dot{\mathcal{G}}(\cdot)$, where $\ddot{\mathcal{Z}}_1(\gamma) := \sum_{n=1}^{\infty} b_n(\gamma) z_n$ and $\dot{\mathcal{G}}(\gamma) := \left(\sum_{n=1}^{\infty} b_n^2(\gamma)\right)^{-1/2}\ddot{\mathcal{Z}}_1(\gamma)$.* □

We prove Lemma A. 2 by showing that the Gaussian process $\dot{\mathcal{G}}(\cdot)$ given in Lemma A. 2 has the same covariance structure as $\ddot{\mathcal{G}}(\cdot)$, and for this purpose, we focus on proving that for all $\gamma, \widetilde{\gamma} \geq 0$, $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}(\widetilde{\gamma})]$. If $\gamma, \widetilde{\gamma} > 0$, the desired equality trivially follows from the definition of $\dot{\mathcal{G}}(\cdot)$. On the other hand, applying L'Hôpital's rule iterative shows that $\mathrm{plim}_{\gamma\downarrow 0}\dot{\mathcal{G}}(\gamma) = \frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2 \sim N(0,1)$, so that if we let $\dot{\mathcal{G}}_2 := \lim_{\gamma\downarrow 0}\dot{\mathcal{G}}(\gamma)$, then for $\gamma \neq 0$, $\mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}_2] = \left[\sqrt{3}b_1(\gamma) + b_2(\gamma)\right]/\{2\ddot{k}_1^{1/2}(\gamma,\gamma)\}$. We show in the proof of Lemma A. 2 that the term on the right side is identical to $\ddot{\rho}_3(\gamma)$ in (A.4), so that the covariance kernel of $\dot{\mathcal{G}}(\cdot)$ is identical to $\ddot{\rho}(\cdot,\cdot)$. This fact implies that $\ddot{\mathcal{G}}(\cdot)$ has the same distribution as $\dot{\mathcal{G}}(\cdot)$, and $\dot{\mathcal{G}}_2$ can be regarded as the weak limit obtained under $\mathcal{H}'_{02}$.

Lemma A. 2 can be used to obtain the approximate null limit distribution of the QLR test statistic. We cannot generate $\dot{\mathcal{G}}(\cdot)$ using the infinite number of $b_n(\cdot)$, but we can simulate the following process to approximate the distribution of $\dot{\mathcal{G}}(\cdot)$: $\dot{\mathcal{G}}(\gamma; K) := (\sum_{n=1}^{K} b_{K,n}^2(\gamma))^{-1/2} \sum_{n=1}^{K} b_{K,n}(\gamma) z_k$, where for $n = 2, 3, \ldots$, $b_{K,1}(\gamma) := (1 - 2a_K(\gamma))/\sqrt{2}$, $a_K(\gamma) := \sum_{k=1}^{K} (-1)^{k-1}/(1+\gamma k)^3$ and $b_{K,n}(\gamma) := \sqrt{n(n+1)} \sum_{k=1}^{K} (-1)^{k-1}(\gamma k)^{n-1}/(1+\gamma k)^3$. If K is sufficiently large, the distribution of $\dot{\mathcal{G}}(\cdot; K)$ is close to that of $\dot{\mathcal{G}}(\cdot)$ as can be easily affirmed by simulations.

We conduct Monte Carlo Simulations for the LSTAR case as in the ESTAR case. The only aspect different from the ESTAR case is that the DGP is the one defined in the beginning of this section. Simulation results are summarised into Figure A.2. We use the same parameter spaces $\Gamma = \Gamma_i$, $i = 1, ..., 4$, as before, and we can see that the empirical distribution and PDF estimate of the QLR test approach the null limit distribution and its PDF that are obtained using $\dot{\mathcal{G}}(\cdot; K)$ with $K = 2,500$. This shows that the theory in Section 2 is also valid for the LSTAR model. When the parameter space Γ for γ becomes even larger, we obtain similar results. To save space, they are not reported.

A.1.3 Application of the Weighted Bootstrap

The standard approach to obtaining the null limit distribution of the QLR test is not applicable for empirical analysis because it requires knowledge of the error distribution. Without this information it is not possible in practice to obtain a distributionally equivalent Gaussian process. Hansen's (1996) weighted bootstrap is useful for this case. We apply it to our models as in Cho and White (2010), Cho, Ishida, and White (2011, 2014), and Cho, Cheong, and White (2011).

Although the relevant weighted bootstrap is available in Cho, Cheong, and White (2011), we provide here a version adapted to the structure of the STAR model. We consider the previously studied ESTAR and LSTAR models, in which the transition function $f_t(\gamma)$ is respectively given by $1 - \exp(-\gamma y_{t-1}^2)$ for the ESTAR model and $\{1 + \exp(\gamma y_{t-1})\}^{-1} - 1/2$ for the LSTAR model. However, the discussions made in this section can be applied to other transition functions as well. Specifically, we first let $\widetilde{\theta}_n$ denote the least squares estimator under the null and let $\widetilde{u}_{n,t} := y_t - y_{t-1}\widetilde{\theta}_n$. Then, using the residuals $(\widetilde{u}_{n,1}, \ldots, \widetilde{u}_{n,n})$, we compute the score $s_{n,t}(\gamma) = \{\widetilde{W}_n(\gamma)\}^{-1/2}\, \widetilde{d}_{n,t}(\gamma)$ for each grid point of $\gamma \in \Gamma$, where

$$\widetilde{W}_n(\gamma) := n^{-1}\left(\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t^2(\gamma) z_t z_t' - \sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t(\gamma) z_t z_t' [\sum_{t=1}^{n} \widetilde{u}_{n,t}^2 z_t z_t']^{-1} \sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t(\gamma) z_t z_t'\right)$$

and

$$\widetilde{d}_{n,t}(\gamma) := z_t f_t(\gamma) \widetilde{u}_{n,t} - n^{-1} \sum_{t=1}^{n} \widetilde{u}_{n,t}^2 f_t(\gamma) z_t z_t' [n^{-1} \sum_{t=1}^{n} \widetilde{u}_{n,t}^2 z_t z_t']^{-1} z_t \widetilde{u}_{n,t}.$$

Given the score function $s_{n,t}(\gamma)$, we construct the following pseudo-QLR test statistic:

$$\overline{QLR}_{b,n} := \sup_{\gamma \in \Gamma} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \widetilde{s}_{n,t}(\gamma)' \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \widetilde{s}_{n,t}(\gamma)$$

and $\widetilde{s}_{n,t}(\gamma) := s_{n,t}(\gamma) z_{b,t}$, where $z_{b,t} \sim$ IID $(0,1)$ with respect to b and t, $b = 1, 2, \ldots, B$, and B is the number of bootstrap replications. For example, we can resample $z_{b,t}$ from the standard normal distribution. For possible two-point distributions, see Davidson *et al.* (2007). Finally, we estimate the empirical p-value by $\widehat{p}_n := B^{-1} \sum_{b=1}^{B} \mathbb{I}[QLR_n < \overline{QLR}_{b,n}]$, where $\mathbb{I}[\cdot]$ is the indicator function. When the null hypothesis holds, this proportion converges to α.

The intuition of the weighted bootstrap is straightforward. Note that if the null hypothesis is valid, the QLR test statistic is bounded in probability, and its null limit distribution can be revealed by the covariance structure of $\widetilde{s}_{n,t}(\cdot)$ asymptotically. That is, for each γ and $\widetilde{\gamma}$, $\mathbb{E}[\widetilde{s}_{n,t}(\gamma)\widetilde{s}_{n,t}(\widetilde{\gamma})']$ converges to $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})']$ from the fact that $z_{b,t}$ is independent of $\widetilde{d}_{n,t}(\cdot)$ such that its population mean is zero and variance is unity. This means that $\mathbb{E}[\widetilde{s}_{n,t}(\gamma)\widetilde{s}_{n,t}(\widetilde{\gamma})']$ is asymptotically equivalent to $\widetilde{W}_n^{-1/2}(\gamma)\mathbb{E}[\widetilde{d}_{n,t}(\gamma)\widetilde{d}_{n,t}(\widetilde{\gamma})']\widetilde{W}_n^{-1/2}(\widetilde{\gamma})$, which is asymptotically equivalent to $\mathbb{E}[\mathcal{G}_1(\gamma)\mathcal{G}_1(\widetilde{\gamma})']$ as given in Theorem 3. Therefore, the null limit distribution can be asymptotically revealed by the resampling distribution of $\overline{QLR}_{b,n}$. On the contrary, if the alternative hypothesis is valid, the QLR test statistic is not bounded in probability, but $\overline{QLR}_{b,n}$ is bounded in probability from the fact that $z_{b,t}$ is distributed around zero, so that the chance for the QLR test statistic to be bounded by the critical value obtained by the resampling distribution of $\overline{QLR}_{b,n}$ gets smaller, as n increases. This aspect implies that the weighted bootstrap is asymptotically consistent.

We conduct a small-scale Monte Carlo experiment to study the performance of the empirical p-values. The DGP is given in Section A.1.1 and A.1.2. To compute the empirical p-values, we set $B = 300$ and obtain $\widehat{p}_n^{(i)}$ for $i = 1, 2, \ldots, 2{,}000$. Then, for a specified nominal value of α, we compute $\frac{1}{2000}\sum_{i=1}^{2000} \mathbb{I}[\widehat{p}_n^{(i)} < \alpha]$.

The simulation results are displayed in the percentile-percentile (PP) plots for the ESTAR and LSTAR models in Figures A.3 (ESTAR) and A.4 (LSTAR). The horizonal unit interval stands for α, and the vertical unit interval is the space of p-values. As a function of α, the aforementioned proportion should converge to the 45-degree line under the null hypothesis. As before, the four parameter spaces are considered: $\Gamma = \Gamma_i$, $i = 1, ..., 4$. The results are summarised as follows. First, as a function of α, the proportion $\frac{1}{2000}\sum_{i=1}^{2000} \mathbb{I}[\widehat{p}_n^{(i)} < \alpha]$ does converge to the 45-degree line. Second, the empirical rejection rates estimated by the weighted bootstrap are closest to the nominal levels when the parameter space is small. Although the overall finite sample level distortions are smaller for the ESTAR model than the LSTAR model, the empirical rejection rate is close to the nominal significance level if α is close to zero. Finally, as the size of the parameter space increases, more observations are needed to better approximate the 45-degree line in the PP plots. We have conducted simulations using even larger parameter spaces and obtained similar results. We omit reporting them for brevity.

A.2 Proofs

Proof of Lemma 1. (i) Given Assumptions 1, 2, 3, and 5, it is trivial to show that $\widehat{\sigma}^2_{n,0} \stackrel{\text{a.s.}}{\to} \sigma^2_*$ by the ergodic theorem.

(ii) The null limit distribution of $QLR^{(1)}_n$ is determined by the two terms in $QLR^{(1)}_n$: $Z'F(\cdot)Mu$ and $Z'F(\cdot)M F(\cdot)Z$. We examine their null limit behaviour one by one and combine the limit results using the converging-together lemma in Billingsley (1999, p. 39).

(a) We show the weak convergence part of $n^{-1/2}Z'F(\cdot)Mu$. Using the definition of $M := I - Z(Z'Z)^{-1} Z'$ we have $Z'F(\gamma)Mu = Z'F(\gamma)u - Z'F(\gamma)Z(Z'Z)^{-1}Z'u$, and we now examine the components on the right-hand side of this equation separately. For each $\gamma \in \Gamma$, we define $\widehat{f}_{n,t}(\gamma) := f_t(\gamma)u_t z_t - (\sum_{t=1}^n f_t(\gamma) z_t z_t')(\sum_{t=1}^n z_t z_t')^{-1} \sum_{t=1}^n z_t u_t$, $\widetilde{f}_{n,t}(\gamma) := f_t(\gamma)u_t z_t - \mathbb{E}[f_t(\gamma)z_t z_t']\mathbb{E}[z_t z_t']^{-1}\sum_{t=1}^n z_t u_t$ and show that

$$\sup_{\gamma\in\Gamma(\epsilon)} \left\| n^{-1/2}\sum_{t=1}^n \left[\widehat{f}_{n,t}(\gamma) - \widetilde{f}_{n,t}(\gamma)\right]\right\|_\infty = o_{\mathbb{P}(1)}, \tag{A.6}$$

where $\Gamma(\epsilon) := \{\gamma \in \Gamma : |\gamma| \geq \epsilon\}$ and $\|\cdot\|_\infty$ is the uniform matrix norm. We have

$$\begin{aligned}
&\sup_{\gamma\in\Gamma(\epsilon)} \left\| \frac{1}{\sqrt{n}}\sum_{t=1}^n \left[\widehat{f}_{n,t}(\gamma) - \widetilde{f}_{n,t}(\gamma)\right]\right\|_\infty \\
&\quad \leq \sup_{\gamma\in\Gamma(\epsilon)} \left\| \left(\frac{1}{n}\sum_{t=1}^n f_t(\gamma)z_t z_t'\right)\left\{\left(\frac{1}{n}\sum_{t=1}^n z_t z_t'\right)^{-1} - \mathbb{E}[z_t z_t']^{-1}\right\} n^{-1/2}\sum_{t=1}^n z_t u_t \right\|_\infty \\
&\quad + \sup_{\gamma\in\Gamma(\epsilon)} \left\| \left\{\left(n^{-1}\sum_{t=1}^n f_t(\gamma)z_t z_t'\right) - \mathbb{E}[f_t(\gamma)z_t z_t']\right\}\mathbb{E}[z_t z_t']^{-1} n^{-1/2}\sum_{t=1}^n z_t u_t \right\|_\infty .
\end{aligned} \tag{A.7}$$

We show that each term on the right-hand side of (A.7) is $o_{\mathbb{P}}(1)$. Now, $\{z_t u_t, \mathcal{F}_t\}$ is a martingale difference sequence, where $\mathcal{F}_t$ is the smallest sigma-field generated by $\{z_t u_t, z_{t-1}u_{t-1}, \ldots\}$. Therefore, $\mathbb{E}[z_t u_t|\mathcal{F}_{t-1}] = 0$, $\mathbb{E}[|Z_{t,j}u_t|^2] = \mathbb{E}[u_t^4]^{1/2}\mathbb{E}[|Z_{t,j}|^4]^{1/2} \leq \mathbb{E}[m_t^4]^{1/2}\mathbb{E}[Z_{t,j}^4]^{1/2} < \infty$, and $\mathbb{E}[u_t^2 z_t z_t']$ is positive definite. Thus, $n^{-1/2}\sum_{t=1}^n z_t u_t$ is asymptotically normal. Next, we note that $n^{-1/2}\sum_{t=1}^n f_t(\gamma)u_t z_t$ is also asymptotically normal. This follows from the fact that $\{f_t(\gamma)u_t z_t, \mathcal{F}_t\}$ is a martingale difference sequence, and for each j,

$|f_t(\gamma)u_t z_{t,j}|^2 \leq m_t^6$, and $\mathbb{E}[m_t^6] < \infty$ by Assumptions 4 and 5. Furthermore, $\sup_{\gamma\in\Gamma} \|n^{-1}\sum_{t=1}^n f_t(\gamma)z_t z_t' - \mathbb{E}[f_t(\gamma)z_t z_t']\|_\infty = o_{\mathbb{P}}(1)$ by Ranga Rao's (1962) uniform law of large numbers. Thus,

$$\sup_{\gamma\in\Gamma(\epsilon)} \left\| \left\{ n^{-1}\sum_{t=1}^n f_t(\gamma)z_t z_t' - \mathbb{E}[f_t(\gamma)z_t z_t'] \right\} \mathbb{E}[z_t z_t']^{-1} n^{-1/2} \sum_{t=1}^n z_t u_t \right\|_\infty = o_{\mathbb{P}}(1). \tag{A.8}$$

This shows that the second term of (A.7) is $o_{\mathbb{P}}(1)$. We now demonstrate that the first term of (A.7) is also $o_{\mathbb{P}}(1)$. By Assumption 4 and the ergodic theorem, we note that $\|n^{-1}\sum_{t=1}^n z_t z_t' - \mathbb{E}[z_t z_t']\|_\infty = o_{\mathbb{P}}(1)$, and $|\sum_{t=1}^n f_t(\gamma)z_{t,j}z_{t,i}| \leq \sum_{t=1}^n m_t^3 = O_{\mathbb{P}}(n)$, so that (A.8) follows, leading to (A.6). Therefore, $n^{-1/2}Z'F(\gamma)M u \overset{\text{A}}{\sim} N[0, B_1(\gamma,\gamma)]$ by noting that $\mathbb{E}[\widetilde{f}_{n,t}(\gamma)\widetilde{f}_{n,t}(\gamma)'] = B_1(\gamma,\gamma)$. Using the same methodology, we can show that for each $\gamma, \widetilde{\gamma} \in \Gamma(\epsilon)$,

$$\frac{1}{\sqrt{n}} \begin{bmatrix} Z'F(\gamma)Mu \\ Z'F(\widetilde{\gamma})Mu \end{bmatrix} \overset{\text{A}}{\sim} N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} B_1(\gamma,\gamma) & B_1(\gamma,\widetilde{\gamma}) \\ B_1(\widetilde{\gamma},\gamma) & B_1(\widetilde{\gamma},\widetilde{\gamma}) \end{bmatrix} \right].$$

Finally, we have to show that $\{\widetilde{f}_{n,t}(\cdot)\}$ is tight. First note that by Assumptions 1, 2, and 4, it follows that $|f_t(\gamma)z_{t,j}u_t - f_t(\widetilde{\gamma})z_{t,j}u_t| \leq m_t\,|z_{t,j}u_t|\;|\gamma-\widetilde{\gamma}|$ for each j. From this we obtain that $\sup_{|\gamma-\widetilde{\gamma}|<\eta}|f_t(\gamma)z_{t,j}u_t - f_t(\widetilde{\gamma})z_{t,j}u_t|^{2+\omega} \leq m_t^{2+\omega}|z_{t,j}u_t|^{2+\omega}\eta^{2+\omega} \leq m_t^{6+3\omega}\eta^{2+\omega}$, so that it follows that $\mathbb{E}[\sup_{|\gamma-\widetilde{\gamma}|<\eta}|f_t(\gamma)z_{t,j}u_t - f_t(\widetilde{\gamma})z_{t,j}u_t|^{2+\omega}]^{\frac{1}{2+\omega}} \leq \mathbb{E}[m_t^{6+3\omega}]^{\frac{1}{2+\omega}}\eta$ for each j. This implies that $\{n^{-1/2}f_t(\cdot)z_{t,j}u_t\}$ is tight because Ossiander's $L^{2+\omega}$ entropy is finite.

Next, for some $c > 0$, it holds that $\|\mathbb{E}[f_t(\gamma)z_t z_t']\mathbb{E}[z_t z_t']^{-1}z_t u_t - \mathbb{E}[f_t(\widetilde{\gamma})z_t z_t']\mathbb{E}[z_t z_t']^{-1}z_t u_t\|_\infty = \|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\}z_t z_t']\mathbb{E}[z_t\ z_t']^{-1}z_t u_t\|_\infty \leq cm_t^2 \left\|\mathbb{E}[z_t z_t']^{-1}\right\|_\infty \|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\}z_t z_t']\|_\infty$ by the property of the uniform norm and Assumption 5. Also note that $\|\mathbb{E}[f_t(\gamma)z_t z_t' - f_t(\widetilde{\gamma})z_t z_t']\|_\infty \leq \|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\}z_t z_t']\|_1$ and by Assumption 4, for each $i, j = 1, 2, \ldots, m+1$, $|z_{t,j}z_{t,i}[f_t(\gamma) - f_t(\widetilde{\gamma})]| \leq m_t^3|\gamma-\widetilde{\gamma}|$, where $\|[g_{i,j}]\|_1 := \sum_i\sum_j |g_{i,j}|$. Therefore,

$$\begin{aligned} &\left\|\mathbb{E}[f_t(\gamma)z_t z_t']\mathbb{E}[z_t z_t']^{-1}z_t u_t - \mathbb{E}[f_t(\widetilde{\gamma})z_t z_t']\mathbb{E}[z_t z_t']^{-1}z_t u_t\right\|_\infty \\ &\quad \leq cm_t^2\left\|\mathbb{E}[z_t z_t']^{-1}\right\|_\infty \left\|\mathbb{E}[\{f_t(\gamma) - f_t(\widetilde{\gamma})\}z_t z_t']\right\|_\infty \leq c^2(m+1)m_t^2\left\|\mathbb{E}[z_t z_t']^{-1}\right\|_\infty \mathbb{E}[m_t^3]|\gamma-\widetilde{\gamma}|. \end{aligned} \tag{A.9}$$

This inequality (A.9) implies that $\{n^{-1/2}\mathbb{E}[f_t(\cdot)z_t z_t']\mathbb{E}[z_t z_t']^{-1}z_t u_t\}$ is also tight. Hence, it follows that for some $b < \infty$, $\mathbb{E}[\sup_{|\gamma-\widetilde{\gamma}|<\eta}|\widetilde{f}_t(\gamma) - \widetilde{f}_t(\widetilde{\gamma})|^{2+\omega}] \leq b\cdot\eta$. That is, $\{n^{-1/2}\sum_{t=1}^n \widetilde{f}_{n,t}(\cdot)\}$ is tight. From this and the fact that the finite-dimensional multivariate CLT holds, the weak convergence of $\{n^{-1/2}\sum_{t=1}^n \widetilde{f}_{n,t}(\cdot)\}$ is established.

(b) Next, we examine the limit behaviour of $n^{-1}Z'F(\cdot)F(\cdot)Z$. Note that $n^{-1}Z'F(\gamma)F(\gamma)Z = n^{-1}\sum_{t=1}^{n} f_t^2(\gamma)z_t z_t' - \{n^{-1}\sum_{t=1}^{n} f_t(\gamma)z_t z_t'\}\{n^{-1}\sum_{t=1}^{n} z_t z_t'\}^{-1}\{n^{-1}\sum_{t=1}^{n} f_t(\gamma)z_t z_t'\}$ and, given Assumptions 1, 2, 3, 4, and 6, $\sup_{\gamma\in\Gamma(\epsilon)}\|n^{-1}\sum_{t=1}^{n} f_t^2(\gamma)\ z_t z_t' - \mathbb{E}[f_t^2(\gamma)z_t z_t']\| \overset{\text{a.s.}}{\to} 0$ and $\sup_{\gamma\in\Gamma(\epsilon)}\|n^{-1}\sum_{t=1}^{n} f_t(\gamma)z_t z_t' - \mathbb{E}[f_t(\gamma)z_t\ z_t']\| \overset{\text{a.s.}}{\to} 0$ by Ranga Rao's (1962) uniform law of large numbers. Therefore, from the fact that $\|n^{-1}\sum_{t=1}^{n} z_t z_t' - \mathbb{E}[z_t z_t']\|_\infty = o_{\mathbb{P}}(1)$, it follows that $\sup_{\gamma\in\Gamma(\epsilon)}|n^{-1}Z'F(\gamma)MF(\gamma)Z - \{\mathbb{E}[f_t^2(\gamma)z_t z_t'] - \mathbb{E}[f_t(\gamma)z_t z_t']\mathbb{E}[z_t z_t']^{-1}\,\mathbb{E}[f_t(\gamma)z_t z_t']\}| = o_{\mathbb{P}}(1)$. Applying the converging-together lemma yields the desired result.

(iii) This result trivially follows from the fact that $\mathbb{E}[u_t^2|z_t] = \sigma_*^2$. ■

Proof of Lemma 2. Given Assumption 2, $\mathcal{H}_{02}$, and the definition of $H_j(\gamma)$, the j-th order derivative of $\mathcal{L}_n^{(2)}(\,\cdot\,,\theta)$ is obtained as

$$\begin{aligned}\frac{\partial^j}{\partial\gamma^j}\mathcal{L}_n^{(2)}(\gamma,\theta) &= -\sum_{k=0}^{j}\binom{j}{k}\left\{\frac{\partial^k}{\partial\gamma^k}(y - F(\gamma)Z\theta)'\right\}M\left\{\frac{\partial^{j-k}}{\partial\gamma^{j-k}}(y - F(\gamma)Z\theta)\right\}\\ &= 2\theta'Z'H_j(\gamma)Mu - \sum_{k=1}^{j-1}\binom{j}{k}\theta'Z'H_j(\gamma)MH_{j-k}(\gamma)Z\theta \qquad \text{(A.10)}\end{aligned}$$

by iteratively applying the general Leibniz rule. We now evaluate this derivative at $\gamma = 0$. Note that $H_j(0) = 0$ if $j < \kappa$ by the definition of κ. This implies that $(\partial^j/\partial\gamma^j)\mathcal{L}_n^{(2)}(0,\theta) = 0$ for $j = 1, 2, \ldots, \kappa - 1$. This also implies that $\binom{j}{k}\theta'Z'H_j(0)MH_{j-k}(0)Z\theta = 0$ for $j = \kappa,\ \kappa+1,\ \ldots,\ 2\kappa - 1$. Therefore, $\frac{\partial^j}{\partial\gamma^j}\mathcal{L}_n^{(2)}(0,\theta) = 2\theta'Z'H_j(0)Mu$. Finally, we examine the case in which $j = 2\kappa$. For each $j < 2\kappa$, $H_j(0) = 0$ and $H_\kappa(0) \neq 0$, so that the summand of the second term in the right side of (A.10) is different from zero only when $j = 2\kappa$ and $k = \kappa$: $\frac{\partial^{2\kappa}}{\partial\gamma^{2\kappa}}\mathcal{L}_n^{(2)}(0,\theta) = 2\theta'Z'H_{2\kappa}(\gamma)Mu - \binom{2\kappa}{\kappa}\theta'Z'H_\kappa(\gamma)MH_\kappa(\gamma)Z\theta$. This completes the proof. ■

Proof of Lemma 3. Given Assumptions 1, 2, 7, and $\mathcal{H}_{02}$, we note that

$$\begin{aligned}QLR_n^{(2)} &:= \sup_\theta \overline{QLR}_n^{(2)}(\theta)\\ &= \sup_\theta \sup_\varsigma \frac{1}{\widehat{\sigma}_{n,0}^2}\left[\frac{2\{\theta'G_\kappa'u\}\varsigma^\kappa}{\kappa!\sqrt{n}} - \frac{1}{(2\kappa)!n}\left\{\binom{2\kappa}{\kappa}\theta'G_\kappa'G_\kappa\theta\right\}\varsigma^{2\kappa}\right] + o_{\mathbb{P}}(n). \qquad \text{(A.11)}\end{aligned}$$

Then, the FOC with respect to ς implies that $\widehat{\varsigma}_n^\kappa(\theta) = W_n(\theta)$, κ is odd; and $\widehat{\varsigma}_n^\kappa(\theta) = \max[0, W_n(\theta)]$, if κ is even by noting that $\widehat{\varsigma}_n^\kappa(\theta)$ cannot be negative. If we plug $\widehat{\varsigma}_n^\kappa(\theta)$ back into the right side of (A.11), the desired result follows. ■

Proof of Lemma 4. Before proving Lemma 4, we first show that for each j, $Z'H_j(0)Mu = O_{\mathbb{P}}(n^{1/2})$, so that $j = \kappa+1,\dots,2\kappa-1$, $Z'H_j(0)Mu = o_{\mathbb{P}}(n^{j/2\kappa})$. Note that for $j = \kappa+1,\dots,2\kappa$, $Z'H_jMu = \sum_{t=1}^n z_t h_{t,j}(0)u_t - \sum_{t=1}^n z_t h_{t,j}(0)z_t' \left(\sum_{t=1}^n z_t z_t'\right)^{-1} \sum_{t=1}^n z_t u_t$. First, we apply the ergodic theorem to $n^{-1}\sum_t z_t h_{t,j}(0)z_t'$ and $n^{-1}\sum_t z_t z_t'$, respectively. Second, given Assumptions 1, 2, 3, 7, and 8, following the proof of Lemma 1, we have that $n^{-1/2}\sum_t z_t u_t$ is asymptotically normal. Furthermore, for all $j = \kappa+1,\dots,2\kappa$, $n^{-1/2}\sum_t z_t h_{t,j}(0)u_t$ is asymptotically normal. For this verification, note that $\{z_t h_{t,j}(0)u_t, \mathcal{F}_t\}$ is a martingale difference sequence, so that for each j, $\mathbb{E}[z_t h_{t,j}(0)u_t|\mathcal{F}_{t-1}] = 0$. Next, we prove that for each j, $\mathbb{E}[z_{t,i}^2 h_{t,j}^2(0)u_t^2] < \infty$. First note that using Assumption 7, $\mathbb{E}[|z_{t,i}^2 h_{t,j}^2(0)u_t^2|] \le \mathbb{E}[|u_t|^4]^{1/2}\mathbb{E}[|h_{t,j}^2(0)z_{t,i}^2|^2]^{1/2} \le \mathbb{E}[|u_t|^4]^{1/2}\mathbb{E}[|h_{t,j}(0)|^8]^{1/4}\,\mathbb{E}[|z_{t,i}|^8]^{1/4} < \infty$ by the Cauchy-Schwarz's inequality. For the same reason, $\mathbb{E}[|z_{t,i}^2 h_{t,j}^2(0)u_t^2|] \le \mathbb{E}[|u_t\, h_{t,j}(0)|^4]^{1/2}\mathbb{E}[|z_{t,i}|^4]^{1/2} \le \mathbb{E}[|u_t|^8]^{1/2}\mathbb{E}[|h_{t,j}(0)|^8]^{1/2}\,\mathbb{E}[|z_{t,i}|^4]^{1/2} < \infty$. By Assumption 8, $\mathbb{E}[u_t^2 z_t h_{t,j}^2(0)z_t']$ is positive definite. It then follows by Theorem 5.25 of White (2001) that $n^{-1/2}\sum_t z_t h_{t,j}(0)u_t$ is asymptotically normal. Thus, $Z'H_j(0)Mu = O_{\mathbb{P}}(n^{1/2})$.

We now consider the statements (*i*)–(*iii*).

(*i*) First, we show that $\theta'Z'H_\kappa(0)Mu = O_{\mathbb{P}}(n^{1/2})$. By the definition of M,

$$Z'H_\kappa(0)Mu = \sum_{t=1}^n z_t h_{t,\kappa}(0)u_t - \sum_{t=1}^n z_t h_{t,\kappa}(0)z_t'\left[\sum_{t=1}^n z_t z_t'\right]^{-1}\sum_{t=1}^n z_t u_t. \tag{A.12}$$

We examine all sums on the right-hand side of (A.12). First, $h_{t,\kappa}(0)$ is a function of z_t, which implies that, given the moment condition in Assumption 7, $n^{-1}\sum z_t h_{t,\kappa}(0)z_t'$ obeys the ergodic theorem. Second, similarly under Assumptions 1, 2, 3, 7, 8, and $\mathcal{H}_{02}$, $n^{-1}\sum z_t z_t'$ also obeys the ergodic theorem. Third, given the assumptions and the proof of Lemma 1, we have already proved that $n^{-1/2}\sum z_t u_t$ is asymptotically normally distributed. Finally, $n^{-1/2}\sum z_t h_{t,\kappa}(0)u_t$ is asymptotically normal, and the proof is similar to that of the asymptotic normality of $n^{-1/2}\sum_t z_t h_{t,j}(0)u_t$ ($j = \kappa+1,\dots,2\kappa$). All these facts imply that $Z'H_\kappa(0)Mu = O_{\mathbb{P}}(n^{1/2})$.

(*ii*) $n^{-1}G_\kappa' G_\kappa \stackrel{\text{a.s.}}{\to} A_2$ by the ergodic theorem.

(*iii*) Note that

$$Z'H_\kappa(0)MH_\kappa(0)Z = \sum_{t=1}^n z_t h_{t,\kappa}^2(0)z_t' - \sum_{t=1}^n z_t h_{t,\kappa}(0)z_t'\left[\sum_{t=1}^n z_t z_t'\right]^{-1}\sum_{t=1}^n z_t h_{t,\kappa}(0)z_t'. \tag{A.13}$$

The limit of (A.13) is revealed by applying the ergodic theorem to each term on the right-hand side of this expression. Consequently, $n^{-1}Z'H_\kappa(0)MH_\kappa(0)Z \stackrel{\text{a.s.}}{\to} \mathbb{E}[g_{t,\kappa}g_{t,\kappa}']$, where $\mathbb{E}[g_{t,\kappa}g_{t,\kappa}'] := \mathbb{E}[z_t H_{2\kappa}^2(0)z_t'] -$

$\mathbb{E}[z_t H_{2\kappa}(0) z_t']\mathbb{E}[z_t z_t']^{-1}\ \mathbb{E}[z_t H_{2\kappa}(0)\ z_t']$. This completes the proof. ■

Proof of Lemma A. 2. The distributional equivalence between $\dot{\mathcal{G}}(\cdot)$ and $\ddot{\mathcal{G}}(\cdot)$ can be established by showing that for all $\gamma, \widetilde{\gamma} \geq 0$, $\mathbb{E}[\ddot{\mathcal{G}}(\gamma)\ddot{\mathcal{G}}(\widetilde{\gamma})] = \mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}(\widetilde{\gamma})]$. We will proceed in three steps. First, we derive the functional form of $\ddot{\rho}(\gamma, \widetilde{\gamma})$. We show that if $\gamma, \widetilde{\gamma} > 0$, then $\ddot{k}_1(\gamma, \widetilde{\gamma}) = \sum_{n=1}^{\infty} b_n(\gamma)b_n(\widetilde{\gamma})$. This in turn implies that for $\gamma, \widetilde{\gamma} > 0$, $\ddot{\rho}(\gamma, \widetilde{\gamma}) = \ddot{k}_1^{-1/2}(\gamma, \gamma)\sum_{n=1}^{\infty} b_n(\gamma)b_n(\widetilde{\gamma})\ddot{k}_1^{-1/2}(\widetilde{\gamma}, \widetilde{\gamma})$. It follows that the specific functional form of $\ddot{\rho}(\gamma, \widetilde{\gamma})$ can be obtained from this result and (A.4).

Second, similarly for all $\gamma, \widetilde{\gamma} \geq 0$, we derive the functional form of $\dot{\rho}(\gamma, \widetilde{\gamma})$ and compare it to $\ddot{\rho}(\gamma, \widetilde{\gamma})$. To do all this, we first note that for all $\gamma, \widetilde{\gamma} > 0$,

$$\begin{aligned}
\ddot{k}_1(\gamma, \tilde{\gamma}) =& \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right)\tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right] - \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right)\right]\mathbb{E}\left[y_t^2 \tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right] \Big/ \mathbb{E}[y_t^2] \\
=& \frac{1}{4}\mathbb{E}\left[y_t^2 \tanh\left(\frac{\gamma y_t}{2}\right)\tanh\left(\frac{\widetilde{\gamma} y_t}{2}\right)\right]. \qquad \text{(A.14)}
\end{aligned}$$

This follows from that fact that for any $x \in \mathbb{R}$, $\tanh(x) = -\tanh(-x)$ and that y_t follows the Laplace distribution with mean zero and variance 2, so that $\mathbb{E}\left[y_t^2 \tanh\left(\ \gamma y_t/2\right)\right] = 0$. Given this, we can apply the Dirichlet series to $\tanh(\cdot)$ to obtain the functional form of $\ddot{k}_1(\cdot,\cdot)$. Thus, for any $x \in \mathbb{R}$, $\tanh(x) = \mathrm{sgn}(x)(1 - 2\sum_{k=0}^{\infty}(-1)^k \exp(-2|x|(k+1)))$ and, furthermore, that $\mathbb{E}\left[s_t^2 \exp(-s_t\gamma k)\right] = 2/\ (1+\gamma k)^3$ and $\mathbb{E}[s_t^2] = 2$, where $s_t := |y_t|$ follows the exponential distribution with mean 1 and variance 2. Applying these to (A.14) yields

$$\begin{aligned}
\ddot{k}_1(\gamma, \widetilde{\gamma}) &= \mathbb{E}\left[\frac{y_t^2}{4}\tanh\left(\frac{\gamma y_t}{2}\right)\tanh\left(\frac{\tilde{\gamma} y_t}{2}\right)\right] \\
&= \mathbb{E}\left[\frac{s_t^2}{4}\right] - \mathbb{E}\left[\frac{s_t^2}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\exp(-s\gamma k)\right] - \mathbb{E}\left[\frac{s_t^2}{2}\sum_{k=1}^{\infty}(-1)^{k-1}\exp(-s\widetilde{\gamma} k)\right] \\
&\quad + \mathbb{E}\left[s_t^2 \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{k+j-2}\exp(-s_t(\gamma k + \widetilde{\gamma} j))\right] \\
&= \frac{1}{2} - \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3} - \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\widetilde{\gamma} k)^3} + \sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k-2}\frac{2}{(1+\gamma k + \widetilde{\gamma} j)^3}.
\end{aligned}$$

Next, for $|x| < 1$ we have $(1-x)^{-3} = \sum_{n=1}^{\infty}\frac{n(n+1)}{2}x^{n-1}$, so that $(1+\gamma k + \widetilde{\gamma} j)^{-3} = (1+\gamma k)^{-3}(1 + \widetilde{\gamma} j)^{-3}\left(1 - \frac{\gamma k}{1+\gamma k}\frac{\widetilde{\gamma} j}{1+\widetilde{\gamma} j}\right)^{-3}$, where we note that $(1 - \frac{\gamma k}{1+\gamma k}\frac{\widetilde{\gamma} j}{1+\widetilde{\gamma} j})^{-3} = \sum_{n=1}^{\infty}\frac{n(n+1)}{2}\left(\frac{\gamma k}{1+\gamma k}\frac{\widetilde{\gamma} j}{1+\widetilde{\gamma} j}\right)^{n-1}$. Therefore, it follows that

$$\ddot{k}_1(\gamma,\widetilde{\gamma}) = \frac{1}{2} - \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3} - \sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\widetilde{\gamma}k)^3} + \sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k-2}\frac{n(n+1)(\gamma k)^{n-1}(\widetilde{\gamma}j)^{n-1}}{(1+\gamma k)^{n+2}(1+\widetilde{\gamma}j)^{n+2}}.$$

Furthermore,

$$\begin{aligned}&\sum_{n=1}^{\infty}\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}(-1)^{j+k-2}n(n+1)\frac{(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}\frac{(\widetilde{\gamma}j)^{n-1}}{(1+\widetilde{\gamma}j)^{n+2}}\\ &=2\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{(1+\gamma k)^3}\sum_{j=1}^{\infty}\frac{(-1)^{j-1}}{(1+\widetilde{\gamma}j)^3} + \sum_{n=2}^{\infty}n(n+1)\sum_{k=1}^{\infty}\frac{(-1)^{k-1}(\gamma k)^{n-1}}{(1+\gamma k)^{n+2}}\sum_{j=1}^{\infty}\frac{(-1)^{j-1}(\widetilde{\gamma}j)^{n-1}}{(1+\widetilde{\gamma}j)^{n+2}},\end{aligned}$$

which is equal to $2a(\gamma)a(\widetilde{\gamma}) + \sum_{n=2}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})$, where for $n = 2, 3, \ldots$, $a(\gamma) := \sum_{k=1}^{\infty}(-1)^{k-1}/(1+\gamma k)^3$ and $b_n(\gamma) := \sqrt{n(n+1)}\sum_{k=1}^{\infty}(-1)^{k-1}(\gamma k)^{n-1}/(1+\gamma k)^{n+2}$. In particular, $b_1(\gamma) := 2^{-1/2}(1-2a(\gamma))$, so that $\ddot{k}_1(\gamma,\widetilde{\gamma}) = \frac{1}{2} - a(\gamma) - a(\widetilde{\gamma}) + 2a(\gamma)a(\widetilde{\gamma}) + \sum_{n=2}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma}) = \frac{1}{2}(1-2a(\gamma))(1-2a(\widetilde{\gamma})) + \sum_{n=2}^{\infty}b_n(\gamma)\ b_n(\widetilde{\gamma}) = \sum_{n=1}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})$. Then, for each γ and $\widetilde{\gamma} > 0$, $\ddot{\rho}_1(\gamma,\widetilde{\gamma}) := \mathbb{E}[\ddot{\mathcal{G}}_1(\gamma)\ddot{\mathcal{G}}_1(\widetilde{\gamma})] = \ddot{k}_1^{-1/2}(\gamma,\gamma)\sum_{n=1}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})\ddot{k}_1^{-1/2}(\widetilde{\gamma},\widetilde{\gamma})$. In addition, for $\gamma > 0$, we examine $\ddot{\rho}_3(\gamma) := \mathbb{E}[\ddot{\mathcal{G}}_1(\gamma)\ddot{\mathcal{G}}_2]$. Note that from (A.4), $\ddot{\rho}_3(\gamma) = \{\mathbb{E}[y_t^3\tanh(\gamma y_t/2)]\}/\{4\sqrt{6}\ddot{k}_1^{1/2}(\gamma,\gamma)\} = [48\gamma^4 + P_G(3, 1+\frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})]/\{32\sqrt{6}\gamma^4\ddot{k}_1^{1/2}(\gamma,\gamma)\}$ as affirmed by Mathematica. It follows that the specific functional form of $\ddot{\rho}(\gamma,\widetilde{\gamma})$ is given as

$$\ddot{\rho}(\gamma,\widetilde{\gamma}) = \begin{cases}\frac{\ddot{k}_1(\gamma,\widetilde{\gamma})}{\ddot{k}_1^{1/2}(\gamma,\gamma)\ddot{k}_1^{1/2}(\widetilde{\gamma},\widetilde{\gamma})}, & \text{if } \gamma > 0 \text{ and } \widetilde{\gamma} > 0;\\ 1, & \text{if } \gamma = 0 \text{ and } \widetilde{\gamma} = 0;\\ \frac{48\gamma^4 + P_G(3,1+\frac{1}{2\gamma}) - P_G(3,\frac{1+\gamma}{2\gamma})}{32\sqrt{6}\gamma^4\ddot{k}_1^{1/2}(\gamma,\gamma)}, & \text{if } \gamma > 0 \text{ and } \widetilde{\gamma} = 0,\end{cases} \cdot \tag{A.15}$$

Third, we examine the covariance kernel of $\dot{\mathcal{G}}(\cdot)$, viz., $\dot{\rho}(\cdot,\cdot)$. If we let $\gamma, \widetilde{\gamma} > 0$, $\dot{\rho}(\gamma,\widetilde{\gamma}) := \mathbb{E}[\dot{\mathcal{G}}(\gamma)\cdot\mathcal{G}(\widetilde{\gamma})] = \ddot{k}_1^{-1/2}(\gamma,\gamma)\sum_{n=1}^{\infty}b_n(\gamma)b_n(\widetilde{\gamma})\ddot{k}_1^{-1/2}(\widetilde{\gamma},\widetilde{\gamma}) = \ddot{k}_1^{-1/2}(\gamma,\gamma)\ddot{k}_1(\gamma,\widetilde{\gamma})\ddot{k}_1^{-1/2}(\widetilde{\gamma},\widetilde{\gamma}) = \ddot{\rho}_1(\gamma,\widetilde{\gamma})$. Furthermore, by some tedious algebra, $\mathrm{plim}_{\gamma\downarrow 0}\ddot{\mathcal{Z}}_1^2(\gamma) = 0$, $\mathrm{plim}_{\gamma\downarrow 0}\frac{\partial}{\partial\gamma}\ddot{\mathcal{Z}}_1^2(\gamma) = 0$, $\mathrm{plim}_{\gamma\downarrow 0}\frac{\partial^2}{\partial\gamma^2}\ddot{\mathcal{Z}}_1^2(\gamma) = \frac{1}{8}\{3\sqrt{2}Z_1 + \sqrt{6}Z_2\}^2$, $\mathrm{plim}_{\gamma\downarrow 0}\ddot{k}_1(\gamma,\gamma) = 0$, $\mathrm{plim}_{\gamma\downarrow 0}\frac{\partial}{\partial\gamma}\ddot{k}_1(\gamma,\gamma) = 0$, and $\mathrm{plim}_{\gamma\downarrow 0}\frac{\partial^2}{\partial\gamma^2}\ddot{k}_1(\gamma,\gamma) = 3$, so that $\mathrm{plim}_{\gamma\downarrow 0}\dot{\mathcal{G}}^2(\gamma) = (\frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2)^2$, which implies $\dot{\mathcal{G}}_2 := \mathrm{plim}_{\gamma\downarrow 0}\dot{\mathcal{G}}(\gamma) = \frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2 \sim N(0,1)$. Consequently, if $\gamma > 0$,

$$\begin{aligned}\mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}_2] &= \ddot{k}_1^{-1/2}(\gamma,\gamma)\mathbb{E}\left[\ddot{\mathcal{Z}}_1(\gamma)\left(\frac{\sqrt{3}}{2}z_1 + \frac{1}{2}z_2\right)\right] = \ddot{k}_1^{-1/2}(\gamma,\gamma)\left[\frac{\sqrt{3}}{2}b_1(\gamma) + \frac{1}{2}b_2(\gamma)\right]\\ &= \frac{1}{32\sqrt{6}\gamma^4\ddot{k}_1^{1/2}(\gamma,\gamma)}\left[48\gamma^4 + P_G\left(3, 1+\frac{1}{2\gamma}\right) - P_G\left(3, \frac{1+\gamma}{2\gamma}\right)\right].\end{aligned} \tag{A.16}$$

The last equality follows from the fact that $b_1(\gamma) = \frac{1}{8\sqrt{2}\gamma^3}[8\gamma^3 - P_G(2, 1 + \frac{1}{2\gamma}) + P_G(2, \frac{1+\gamma}{2\gamma})]$, $b_2(\gamma) = \frac{1}{16\sqrt{6}\gamma^4}[6\gamma P_G(2, \frac{1}{2\gamma}) - 6\gamma P_G(2, \frac{1+\gamma}{2\gamma}) + P_G(3, \frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})]$, $P_G(2, \frac{1}{2\gamma}) - P_G(2, 1 + \frac{1}{2\gamma}) = -16\gamma^3$, and $P_G(3, \frac{1}{2\gamma}) - P_G(3, 1 + \frac{1}{2\gamma}) = 96\gamma^4$, as obtained by Mathematica. Equation (A.16) then leads to the following functional form for $\dot{\rho}(\gamma, \widetilde{\gamma}) := \mathbb{E}[\dot{\mathcal{G}}(\gamma)\dot{\mathcal{G}}(\widetilde{\gamma})]$:

$$\dot{\rho}(\gamma, \widetilde{\gamma}) = \begin{cases} \frac{\ddot{k}_1(\gamma,\widetilde{\gamma})}{\ddot{k}_1^{1/2}(\gamma,\gamma)\ddot{k}_1^{1/2}(\widetilde{\gamma},\widetilde{\gamma})}, & \text{if } \gamma > 0 \text{ and } \widetilde{\gamma} > 0; \\ 1, & \text{if } \gamma = 0 \text{ and } \widetilde{\gamma} = 0; \\ \frac{48\gamma^4 + P_G(3, 1+\frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})}{32\sqrt{6}\gamma^4 \ddot{k}_1^{1/2}(\gamma,\gamma)}, & \text{if } \gamma > 0 \text{ and } \widetilde{\gamma} = 0, \end{cases}$$

which is identical to the functional form of $\ddot{\rho}(\cdot, \cdot)$ in (A.15). This allows the conclusion that $\ddot{\mathcal{G}}(\cdot)$ has the same distribution as $\dot{\mathcal{G}}(\cdot)$. ■

In the following, we provide additional supplementary claim in (A.5) that is given in the following lemma:

Lemma A. 3. *Given the DGP and Model conditions in Section A.1.2,* $\lim_{\widetilde{\gamma}\downarrow 0} \ddot{\rho}_1^2(\gamma, \widetilde{\gamma}) = [48\gamma^4 + P_G(3, 1 + \frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})]^2 / \{32\sqrt{6}\gamma^4 \ddot{k}_1^{1/2}(\gamma, \gamma)\}^2$. □

Lemma A. 3 implies that $\text{plim}_{\gamma\downarrow 0} \ddot{\mathcal{G}}_1^2(\gamma) = \ddot{\mathcal{G}}_2^2$, so that $\sup_{\gamma\in\Gamma} \ddot{\mathcal{G}}_1^2(\gamma) \geq \ddot{\mathcal{G}}_2^2$ and $QLR_n \Rightarrow \sup_{\gamma\in\Gamma} \ddot{\mathcal{G}}_1^2(\gamma)$.

Proof of Lemma A. 3. From the definition of $\ddot{\rho}_1(\gamma, \widetilde{\gamma})$, note that $\ddot{\rho}_1^2(\gamma, \widetilde{\gamma}) := \ddot{k}_1^{-1}(\gamma, \gamma)\, \ddot{k}_1^2(\gamma, \widetilde{\gamma}) \ddot{k}_1^{-1}(\widetilde{\gamma}, \widetilde{\gamma})$. Furthermore, we have $\text{plim}_{\widetilde{\gamma}\downarrow 0} \ddot{k}_1^2(\gamma, \widetilde{\gamma}) = 0$, $\text{plim}_{\widetilde{\gamma}\downarrow 0} \frac{\partial}{\partial\widetilde{\gamma}} \ddot{k}_1^2(\gamma, \widetilde{\gamma}) = 0$, $\text{plim}_{\widetilde{\gamma}\downarrow 0} \ddot{k}_1(\widetilde{\gamma}, \widetilde{\gamma}) = 0$, $\text{plim}_{\widetilde{\gamma}\downarrow 0}(\partial/\partial\widetilde{\gamma}) \ddot{k}_1(\widetilde{\gamma}, \widetilde{\gamma}) = 0$, $\text{plim}_{\widetilde{\gamma}\downarrow 0}(\partial^2/\partial\widetilde{\gamma}^2)\ddot{k}_1(\widetilde{\gamma}, \widetilde{\gamma}) = 3$, and $\text{plim}_{\widetilde{\gamma}\downarrow 0} \frac{\partial^2}{\partial\widetilde{\gamma}^2} \ddot{k}_1^2(\gamma, \widetilde{\gamma}) = (\{48\gamma^4 + P_G(3, 1+\frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})\} / \{32\sqrt{2}\gamma^4\})^2$ by some algebra using Mathematica. This property implies that $\lim_{\widetilde{\gamma}\downarrow 0} \ddot{\rho}_1^2(\gamma, \widetilde{\gamma}) = (\{48\gamma^4 + P_G(3, 1 + \frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})\} / \{32\sqrt{2}\gamma^4\})^2 / \{3\ddot{k}_1(\gamma, \gamma)\} = [48\gamma^4 + P_G(3, 1 + \frac{1}{2\gamma}) - P_G(3, \frac{1+\gamma}{2\gamma})]^2 / \{32\sqrt{6}\gamma^4 \ddot{k}_1^{1/2}(\gamma, \gamma)\}^2$. This completes the proof. ■

References

BILLINGSLEY, P. (1999): *Convergence of Probability Measures*. New York: Wiley.

CHO, J.S. AND WHITE, H. (2010): "Testing for Unobserved Heterogeneity in Exponential and Weibull Duration Models," *Journal of Econometrics*, 157, 458–480.

DAVIDSON, J., MONTICINI, A., AND PEEL, D. (2007): "Implementing the Wild Bootstrap Using a Two-Point Distribution," *Economics Letters*, 96, 309–315.

ESCRIBANO, Á AND JORDÀ, Ò. (1999): "Testing Nonlinearity: Decision Rules for Selecting between Logistic and Exponential STAR Models," *Spanish Economic Review*, 3, 193–209.

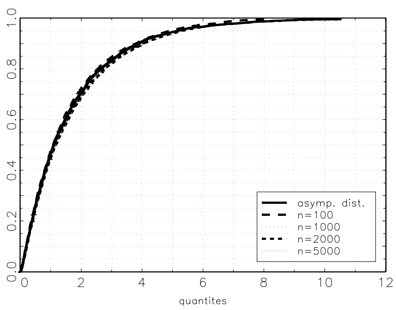
KOOP, G., PESARAN, M.H. AND POTTER, S.M. (1996): "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of econometrics*, 74, 119–147.

LAWRANCE, A. J., AND LEWIS, P. A. W. (1980): "The Exponential Autoregressive-Moving Average EARMA (p, q) Process," *Journal of the Royal Statistical Society. Series B*, 47, 150–161.

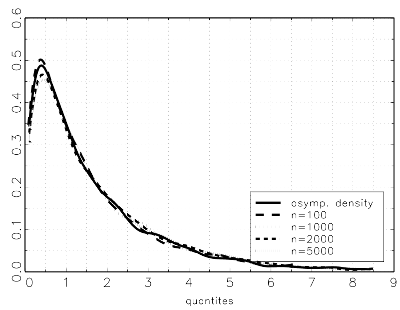
RANGA RAO, R. (1962): "Relations Between Weak and Uniform Convergence of Measures with Applications," *Annals of Mathematical Statistics*, 33, 659–680.

VAN DIJK, D., TERÄSVIRTA, T., AND FRANSES, P. H. (2002): "Smooth Transition Autoregressive Models–a Survey of Recent Developments," *Econometric Reviews*, 21, 1–47.

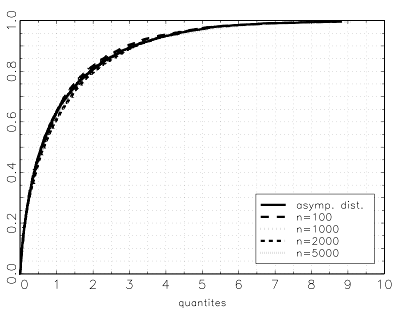
WHITE, H. (2001): *Asymptotic Theory for Econometricians*. Orlando, FL: Academic Press.



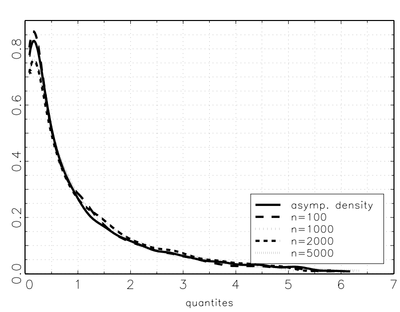
$\Gamma_1 = [0, 2]$



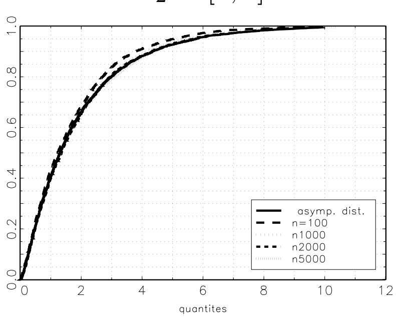
$\Gamma_1 = [0, 2]$



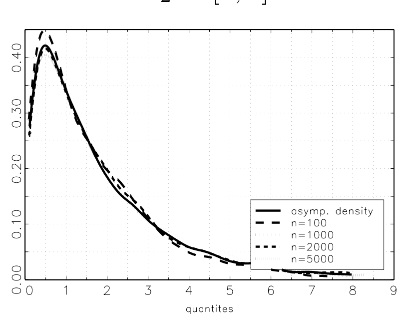
$\Gamma_2 = [0, 3]$



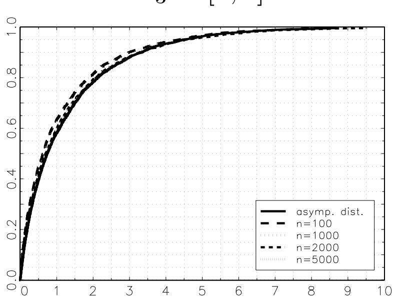
$\Gamma_2 = [0, 3]$



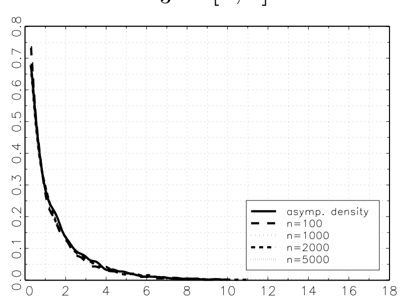
$\Gamma_3 = [0, 4]$



$\Gamma_3 = [0, 4]$

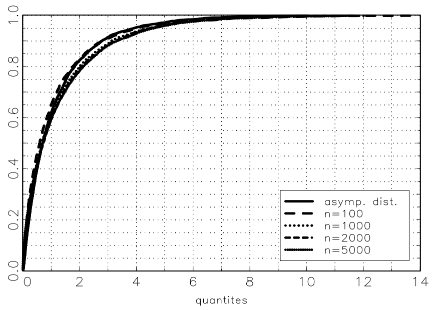


$\Gamma_4 = [0, 5]$

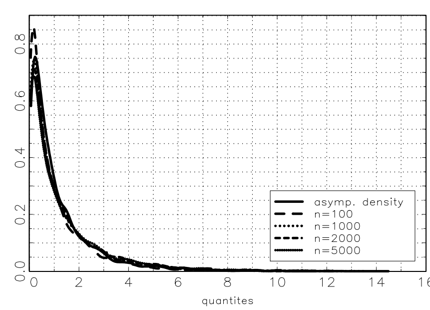


$\Gamma_4 = [0, 5]$

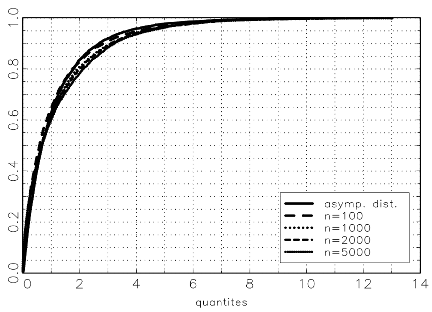
Figure A.1: EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTIC AND ITS NULL LIMIT DISTRIBUTION (ESTAR MODEL CASE). Notes: (*i*) Number of Iterations: 5,000; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t \sim$ IID $N(0,1)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{1 - \exp(-\gamma y_{t-1}^2)\} + u_t$ and $u_t \sim$ IID $N(0,1)$; and (*iv*) $\Gamma_1 = [0,2]$, $\Gamma_2 = [0,3]$, $\Gamma_3 = [0,4]$, $\Gamma_4 = [0,5]$.



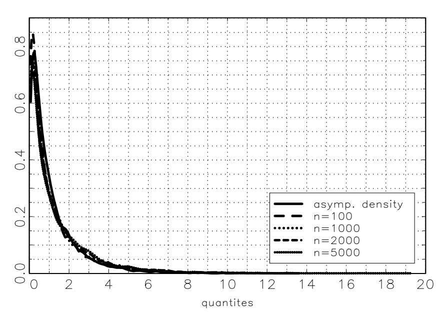
$\Gamma_1 = [0, 2]$



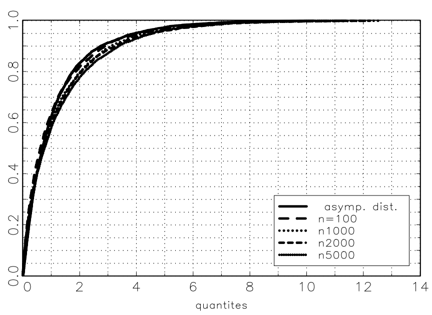
$\Gamma_1 = [0, 2]$



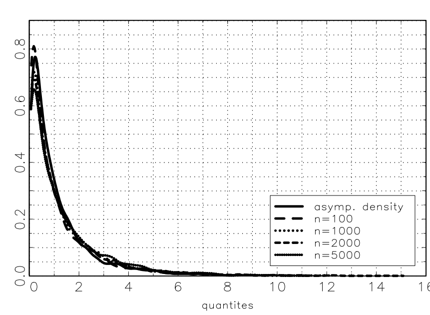
$\Gamma_2 = [0, 3]$



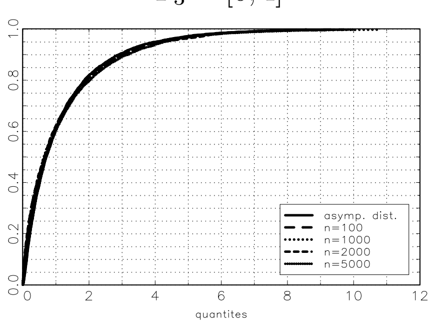
$\Gamma_2 = [0, 3]$



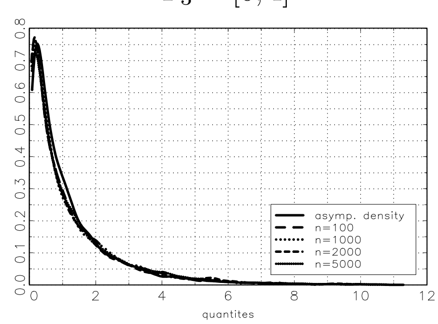
$\Gamma_3 = [0, 4]$



$\Gamma_3 = [0, 4]$

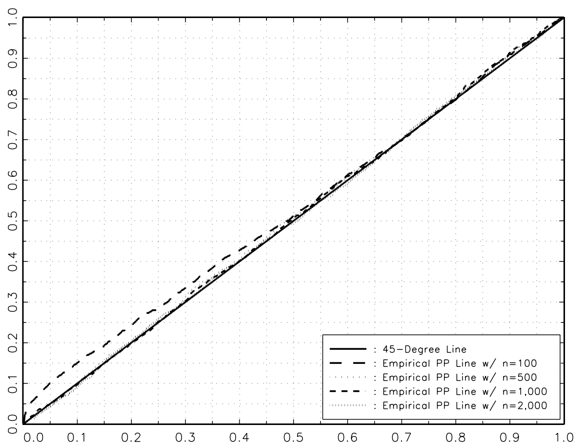


$\Gamma_4 = [0, 5]$

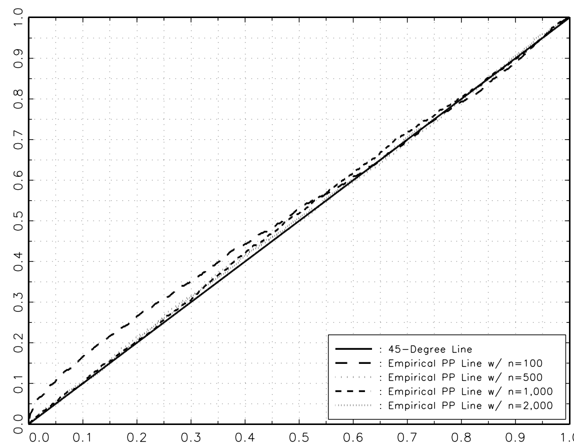


$\Gamma_4 = [0, 5]$

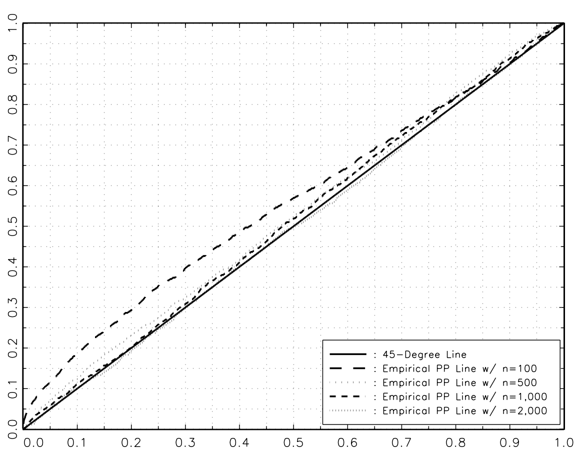
Figure A.2: EMPIRICAL NULL DISTRIBUTIONS OF THE QLR STATISTIC AND ITS NULL LIMIT DISTRIBUTION (LSTAR MODEL CASE). Notes: (*i*) Number of Iterations: 5,000; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t = i_t\ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.5^2$ and $\{\ell_t\} \sim \text{Laplace}(0, 2)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{(1 + \exp(-\gamma y_{t-1}))^{-1}\} + u_t$ and $u_t = i_t\ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.5^2$ and $\{\ell_t\} \sim \text{Laplace}(0, 2)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.



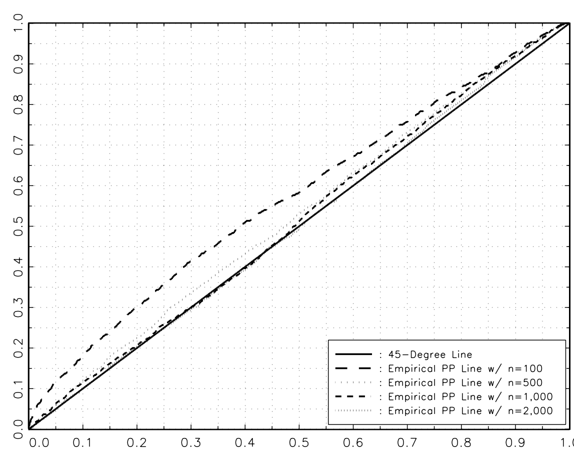
$\Gamma_1 = [0, 2]$



$\Gamma_2 = [0, 3]$

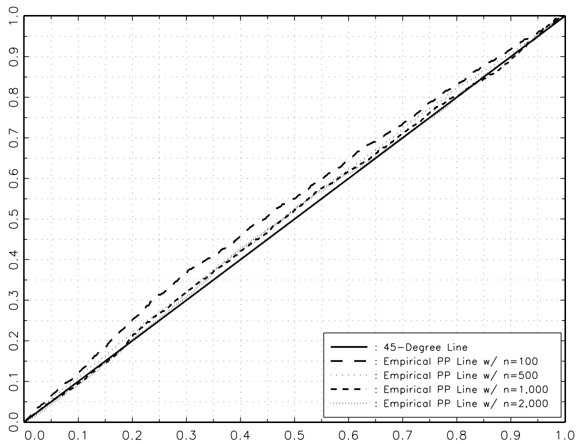


$\Gamma_3 = [0, 4]$

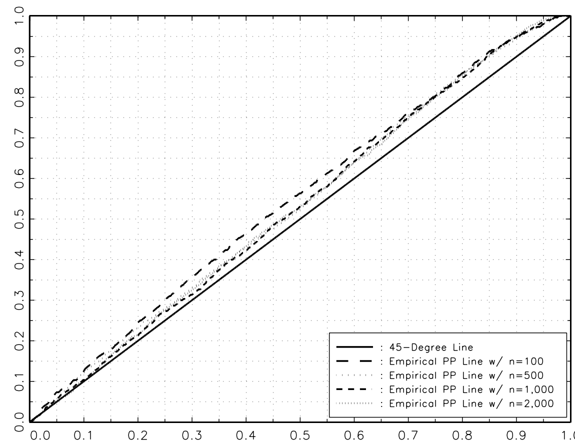


$\Gamma_4 = [0, 5]$

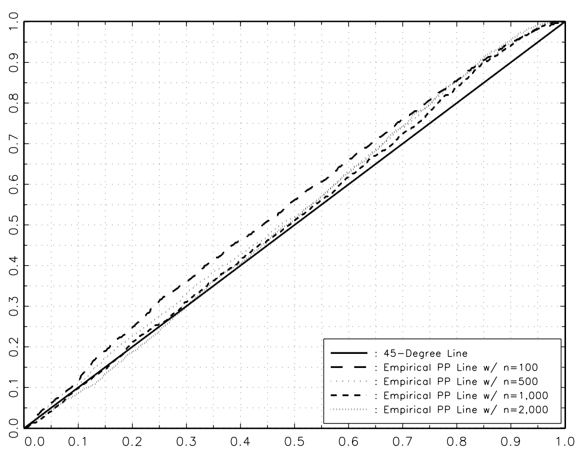
Figure A.3: PP PLOTS OF THE QLR STATISTIC USING THE WEIGHTED BOOTSTRAP (ESTAR MODEL CASE). Notes: (*i*) Number of Iterations: 2,000, Bootstrap Iterations: 300; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t \sim$ IID $N(0, 1)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{1 - \exp(-\gamma y_{t-1}^2)\} + u_t$ and $u_t \sim$ IID $N(0, 1)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.



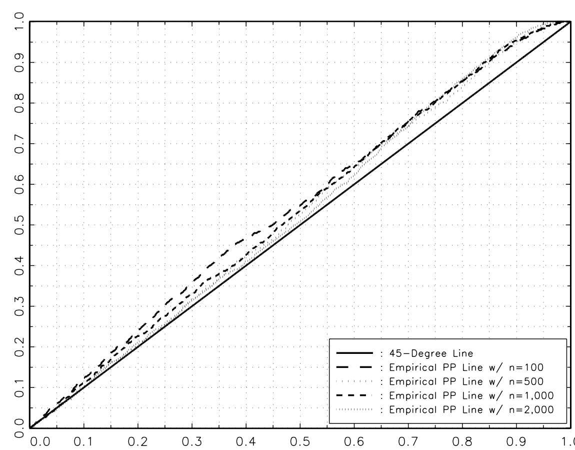
$\Gamma_1 = [0, 2]$



$\Gamma_2 = [0, 3]$



$\Gamma_3 = [0, 4]$



$\Gamma_4 = [0, 5]$

Figure A.4: PP Plots of the QLR Statistic Using the Weighted Bootstrap (LSTAR Model Case). Notes: (*i*) Number of Iterations: 2,000, Bootstrap Iterations: 300; (*ii*) DGP: $y_t = 0.5y_{t-1} + u_t$ and $u_t = i_t\ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.25$ and $\{\ell_t\} \sim$ Laplace$(0, 2)$; (*iii*) Model: $y_t = \pi y_{t-1} + \theta y_{t-1}\{(1 + \exp(-\gamma y_{t-1}))^{-1} - 1/2\} + u_t$ and $u_t = i_t\ell_t$, where $\{i_t\}$ is an IID sequence in which $\mathbf{P}\{i_t = 1\} = 1 - 0.5^2$ and $\{\ell_t\} \sim$ Laplace$(0, 2)$; and (*iv*) $\Gamma_1 = [0, 2]$, $\Gamma_2 = [0, 3]$, $\Gamma_3 = [0, 4]$, $\Gamma_4 = [0, 5]$.