Testing for the buffered autoregressive processes

Zhu, Ke and Yu, Philip L.H. and Li, Wai Keung

Institute of Applied Mathematics, Chinese Academy of Sciences, Department of Statistics and Actuarial Science, University of Hong Kong, Department of Statistics and Actuarial Science, University of Hong Kong

25 November 2013

TESTING FOR THE BUFFERED AUTOREGRESSIVE PROCESSES

By Ke Zhu, Philip L.H. Yu and Wai Keung Li

Chinese Academy of Sciences and University of Hong Kong

This paper investigates a quasi-likelihood ratio (LR) test for the thresholds in buffered autoregressive processes. Under the null hypothesis of no threshold, the LR test statistic converges to a function of a centered Gaussian process. Under local alternatives, this LR test has nontrivial asymptotic power. Furthermore, a bootstrap method is proposed to obtain the critical value for our LR test. Simulation studies and one real example are given to assess the performance of this LR test. The proof in this paper is not standard and can be used in other non-linear time series models.

1. Introduction. After the seminal work of Tong (1978), threshold autoregressive (TAR) models have achieved a great success in practice; see, e.g., Tong (1990) for earlier works and Tong (2011) and the references therein for more recent ones. Generally speaking, the TAR model says that the structure of an AR model shifts among different regimes, i.e.,

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \left(\psi_0 + \sum_{i=1}^{p} \psi_i y_{t-i}\right) R_t + \varepsilon_t, \tag{1.1}$$

where $R_t = I(y_{t-d} \leq r)$ is the regime indicator of y_t, r is the threshold parameter, $d(\geq 1)$ is the delay parameter, and ε_t is an uncorrelated error sequence with zero mean and variance $\sigma^2(> 0)$. There have been a lot of interests to detect the threshold in TAR models since 1990s. Chan (1990, 1993) and Chan and Tong (1990) first accomplished this task by considering a likelihood ratio (LR) test for TAR models. Moreover, Tsay (1989) gave some novel methods in this context; Hansen (1996) studied the Wald test and Lagrange multiplier (LM) test for TAR models; Wong and Li (1997, 2000) studied LM test for TAR-ARCH models; Li and Ling (2013) investigated the portmanteau test for threshold double AR models; see also Tsay (1998), Hansen (1999), Caner and Hansen (2001), Ling and Tong (2005), Li and Li (2008, 2011), and Zhu and Ling (2012).

Keywords and phrases: AR(p) model, Bootstrap method, Buffered AR(p) model, Likelihood ratio test, Marked empirical process, Threshold AR(p) model.

Under model (1.1), the regime of y_t shifts when the state of y_{t-d} changes. In practice, the regime of y_t may not shift immediately, and there could be a buffering region in which the regime of y_t depends on the regime of y_{t-d}. Li, Guan, Li, and Yu (2012) first formulated this situation by assuming that R_t in model (1.1) satisfies

$$R_t = \begin{cases} 1 & \text{if } y_{t-d} \leq r_L \\ 0 & \text{if } y_{t-d} > r_U \\ R_{t-1} & \text{otherwise} \end{cases}, \tag{1.2}$$

where r_L and r_U are two threshold parameters such that $r_L \leq r_U$. They called model (1.1)-(1.2) the buffered AR (BAR) model, and the region in which y_{t-d} lies between r_L and r_U is called the buffering region. Also, they found that the BAR model is the best selected model for the sunspot series in Tong (1990) and GNP series in Tiao and Tsay (1994), and hence it may provide us with a new way to understand the non-linear time series. However, how to test for BAR models is still unknown, and it is more challenging than testing for TAR models because the regime of y_t in this case depends on past observations infinitely far away.

In this paper, we investigate a quasi-LR test for the thresholds in BAR models. Under the null hypothesis of no threshold, the LR test statistic converges to a function of a centered Gaussian process. Under local alternatives, this LR test has nontrivial asymptotic power. Our result nests the one in Chan (1990) as a special case, but its proof is not standard and different from the proof in that paper. Furthermore, a bootstrap method is proposed to obtain the critical value for our LR test. Simulation studies and one real example are given to assess the performance of this LR test.

This paper is organized as follows. Section 2 states our main result on the LR test. Section 3 proposes a bootstrap procedure. The simulation results and one real example are given in Section 4. The proofs are provided in the Appendix, which can be found in Zhu, Yu, and Li (2013). Throughout the paper, some symbols are conventional. $|A| = (tr(A'A))^{1/2}$ is the Euclidean norm of a matrix A. $\|A\|_s = (E|A|^s)^{1/s}$ is the L^s-norm ($s \geq 1$) of a random matrix. A' is the transpose of matrix A. $o_p(1)$ ($O_p(1)$) denotes a sequence of random numbers converging to zero (bounded) in probability. $\to_d$ denotes convergence in distribution and $\Rightarrow$ denotes weak convergence. $I(\cdot)$ is an indicator function.

2. Likelihood ratio test. Let $\phi = (\phi_0, \cdots, \phi_p)'$, $\psi = (\psi_0, \cdots, \psi_p)'$, $\lambda = (\phi', \psi')'$, $\gamma = (r_L, r_U)$, and $x_t = (1, y_{t-1}, \cdots, y_{t-p})'$. Then, model (1.1)-(1.2) becomes

$$y_t = x_t(\gamma)'\lambda + \varepsilon_t, \tag{2.1}$$

where $x_t(\gamma) = (x_t', h_t(\gamma)')'$, $h_t(\gamma) = x_t R_t(\gamma)$, and $R_t(\gamma)$ is defined as in (1.2). Here, we assume that all the roots of the characteristic equation $\phi(x) = x^p - \phi_1 x^{p-1} - \cdots - \phi_p$ lie inside the unit circle, and both p and d are known. Without loss of generality, we further assume that $d \le p$ if $p \ge 1$, because we can set $p = d$ with $\phi_{p+1} = \cdots = \phi_d = 0$ and $\psi_{p+1} = \cdots = \psi_d = 0$ in (2.1) when $d > p \ge 1$.

Suppose that $\{y_0, \cdots, y_N\}$ are $N+1$ consecutive observations from model (2.1) with the true parameters λ_0 and γ_0, where $\lambda_0 = (\phi_0', \psi_0')'$, $\phi_0 = (\phi_{00}, \cdots, \phi_{p0})'$, $\psi_0 = (\psi_{00}, \cdots, \psi_{p0})'$, and $\gamma_0 = (r_{L0}, r_{U0})$. We consider the following hypotheses:

$$\begin{cases} H_0 : \psi_0 = 0, \\ H_1 : \psi_0 \neq 0 \text{ for some } \gamma. \end{cases} \tag{2.2}$$

Model (2.1) is an AR(p) model under H_0 and it is a buffered AR(p) (BAR(p)) model under H_1. When $r_L = r_U$ (i.e., the buffering region is absent), (2.2) is for testing the threshold in the threshold AR(p) (TAR(p)) model, for which the likelihood ratio (LR) test was studied by Chan (1990, 1991) provided that $\varepsilon_t \sim N(0, 1)$ is a sequence of i.i.d. random variables. When $r_L \neq r_U$, since

$$\begin{aligned} R_t(\gamma) = {} & I(y_{t-d} \le r_L) \\ & + \sum_{j=1}^{\infty} I(y_{t-j-d} \le r_L) \prod_{i=1}^{j} I(r_L < y_{t-i+1-d} \le r_U) \text{ a.s.,} \end{aligned} \tag{2.3}$$

we can see that $R_t(\gamma)$ depends on all past observations infinitely far away. Note that $R_t(\gamma)$ in Chan (1990) only depends on y_{t-d}. Thus, the test in Chan (1990) is not a LR test any more and may be less powerful in this case. Motivated by this, we consider an alternative LR test for (2.2).

Denote $Y = (y_p, \cdots, y_N)'$ and $Z_\gamma = (X, X_\gamma) = \big(x_p(\gamma), x_{p+1}(\gamma), \cdots, x_N(\gamma)\big)'$, where

$$\begin{aligned} X &= (x_p, x_{p+1}, \cdots, x_N)', \\ X_\gamma &= \big(h_p(\gamma), h_{p+1}(\gamma), \cdots, h_N(\gamma)\big)'. \end{aligned}$$

Let $n = N - p + 1$ be the effective number of observations. Following Chan (1990), we know that for any fixed value of γ, the LR test statistic is

$$LR_n(\gamma) = \frac{n\,[\sigma_n^2 - \sigma_n^2(\gamma)]}{\sigma_n^2},$$

where

$$\sigma_n^2 = \frac{1}{n}\{Y'Y - (Y'X)(X'X)^{-1}(X'Y)\}, \tag{2.4}$$

$$\sigma_n^2(\gamma) = \frac{1}{n}\{Y'Y - (Y'Z_\gamma)(Z'_\gamma Z_\gamma)^{-1}(Z'_\gamma Y)\}. \tag{2.5}$$

Since the exact value of γ is unknown under H_0, it is natural to construct the LR test by using the maximum of $LR_n(\gamma)$ over the range of γ; see Davis (1977, 1987). Thus, our LR test statistic is defined as

$$LR_n = \sup_{\gamma \in \Gamma} LR_n(\gamma),$$

where $\Gamma \equiv \{(r_L, r_U); a \leq r_L \leq r_U \leq b\}$ and $[a, b]$ is a predetermined interval. Here, we truncate the full range of γ, since LR_n may diverge to infinity in probability as $n \to \infty$; see Andrews (1993a).

Let $K_{\gamma\delta} = E[x_t(\gamma)x_t(\delta)']$. To study the asymptotic theory of LR_n, we need the following three technical assumptions:

Assumption 2.1. *y_t is strictly stationary, ergodic and absolutely regular with mixing coefficients $\beta(m) = O(m^{-A})$ for some $A > v/(v-1)$ and $r > v > 1$; $E|y_t|^{4r} < \infty$, $E|\varepsilon_t|^{4r} < \infty$; and $K_{\gamma\gamma}$ is positive definite.*

Assumption 2.2. *y_t has a bounded and continuous density function.*

Assumption 2.3. *There exists an $A_0 > 1$ such that $2A_0 rv/(r-v) < A$.*

Assumptions 2.1-2.2 are from Hansen (1996), in which the weak convergence of empirical process is derived by using the method in Doukhan, Massart, and Rio (1995). When $\sum_{i=1}^{p} |\phi_i| < 1$ and $\sum_{i=1}^{p} |\phi_i + \psi_i| < 1$, Li, Guan, Li, and Yu (2012) showed that model (2.1) is strictly stationary and ergodic. Assumption 2.3 is needed to prove Lemma A.1 in the Appendix. When $A > v/(v-1)$, a sufficient condition for Assumption 2.3 is that $v < 3r/(2r+1)$, which is stronger than $v < r$ as required in Assumption 2.1. Particularly, when ε_t is a sequence of i.i.d. random variables with a bounded and continuous density function, $\beta(m)$ decays exponentially under H_0 as shown in Pham and Tran (1985). Thus, the mixing condition of y_t in Assumption 2.1 and also Assumptions 2.2-2.3 hold in this case.

Furthermore, we state two key lemmas, under which a uniform expansion of $LR_n(\gamma)$ can be derived.

LEMMA 2.1. *If Assumptions 2.1-2.3 hold, then (i) it follows that*

$$\sup_{\gamma\in\Gamma}\left|\left\{\frac{X_\gamma' X_\gamma}{n}-\frac{X_\gamma' X}{n}\left(\frac{X'X}{n}\right)^{-1}\frac{X'X_\gamma}{n}\right\}^{-1}\right.$$
$$\left. -(\Sigma_\gamma-\Sigma_\gamma\Sigma^{-1}\Sigma_\gamma)^{-1}\right| = o_p(1);$$

(ii) furthermore, under H_0 it follows that

$$\sup_{\gamma\in\Gamma}\left|T_\gamma-\Big(-\Sigma_\gamma\Sigma^{-1},I\Big)\frac{1}{\sqrt{n}}Z_\gamma'\varepsilon\right| = o_p(1),$$

where $\varepsilon=(\varepsilon_p,\cdots,\varepsilon_N)'$, $T_\gamma = n^{-1/2}\left\{X_\gamma' - X_\gamma' X(X'X)^{-1}X'\right\}Y$, $\Sigma = E(x_t x_t')$, and $\Sigma_\gamma = E[x_t x_t' R_t(\gamma)]$.

PROOF. See the Appendix in Zhu, Yu, and Li (2013). □

LEMMA 2.2. *If Assumptions 2.1-2.3 hold, then it follows that*

$$\frac{1}{\sqrt{n}}Z_\gamma'\varepsilon \Rightarrow \sigma G_\gamma$$

as $n\to\infty$, where G_γ is a Gaussian process with zero mean function and covariance kernel $K_{\gamma\delta}$.

PROOF. See the Appendix in Zhu, Yu, and Li (2013). □

Note that

$$\frac{1}{\sqrt{n}}Z_\gamma'\varepsilon = \frac{1}{\sqrt{n}}\sum_{t=p}^{N}(x_t', x_t' R_t(\gamma))'\varepsilon_t.$$

We call $\{n^{-1/2}Z_\gamma'\varepsilon\}$ a marked empirical process as in Stute (1997), where each y_{t-i-d} in $R_t(\gamma)$ is a marker. In view of (2.3), we know that $\{n^{-1/2}Z_\gamma'\varepsilon\}$ involves infinitely many markers, and this is also the case when Ling and Tong (2005) studied the LR test for TMA models. However, their method seems hard to be implemented in our case. Compared with the proof of Lemma 2.1 in Chan (1990) or Ling and Tong (2005), the proofs of Lemmas 2.1-2.2 in the Appendix are not standard and can be used in other non-linear time series models.

We are now ready to present our main result as follows:

THEOREM 2.1. *If Assumptions 2.1-2.3 hold, then under H_0 it follows that*

$$LR_n \to_d \sup_{\gamma\in\Gamma} G_\gamma'\Omega_\gamma G_\gamma$$

as $n\to\infty$, where $\Omega_\gamma = (-\Sigma_\gamma\Sigma^{-1},I)'\left(\Sigma_\gamma-\Sigma_\gamma\Sigma^{-1}\Sigma_\gamma\right)^{-1}(-\Sigma_\gamma\Sigma^{-1},I)$.

PROOF. By (2.4)-(2.5) and a direct calculation, we have

$$n\left[\sigma_n^2 - \sigma_n^2(\gamma)\right]$$

$$= T'_\gamma \left\{ \frac{X'_\gamma X_\gamma}{n} - \frac{X'_\gamma X}{n} \left(\frac{X'X}{n}\right)^{-1} \frac{X'X_\gamma}{n} \right\}^{-1} T_\gamma. \tag{2.6}$$

By Lemmas 2.1-2.2, the conclusion follows directly from the same argument as for Theorem 2.3 in Chan (1990). □

REMARK 2.1. *Note that*

$$G'_\gamma \Omega_\gamma G_\gamma = \xi'_\gamma \left(\Sigma_\gamma - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma\right)^{-1} \xi_\gamma,$$

where $\xi_\gamma = (-\Sigma_\gamma \Sigma^{-1}, I)\, G_\gamma$. *Then, by a direct calculation, we can easily show that for each* $\gamma \in \Gamma$, $G'_\gamma \Omega_\gamma G_\gamma$ *follows a* χ^2 *distribution, namely, for fixed* γ, *the test statistic* $LR_n(\gamma)$ *is asymptotically pivotal under* H_0.

REMARK 2.2. *Although the result in Theorem 2.1 nests the one in Theorem 2.3(ii) of Chan (1990) as a special case, it is necessary to mention some difference between our LR test and that in Chan (1990). First, the denominator of* $LR_n(\gamma)$ *in our case is different from that in Chan (1990), but we can easily show that these two denominators are asymptotically equivalent; see also Ling and Tong (2005). Second, since the region of* Γ *is larger than that in Chan (1990), our LR test needs more computational efforts than that in Chan (1990).*

REMARK 2.3. *As Chan (1990), we only obtained the result under the condition that* $Var(\varepsilon_t) = \sigma^2$. *The case that the threshold effect happens in the variance of* ε_t *needs a further study in the future.*

Next, we study the asymptotical local power of LR_n by considering the following local alternative hypothesis:

$$H_{1n} : \psi_0 = \frac{h}{\sqrt{n}} \quad \text{for a constant vector } h \in \mathcal{R}^{p+1}.$$

THEOREM 2.2. *If Assumptions 2.1-2.3 hold, then under* H_{1n} *it follows that*

$$LR_n \to_d \sup_{\gamma \in \Gamma} \left\{ G'_\gamma \Omega_\gamma G_\gamma + h' \mu_{\gamma\gamma_0} h \right\},$$

as $n \to \infty$, *where* $M_{\gamma\gamma_0} = E[x_t x'_t R_t(\gamma) R_t(\gamma_0)]$ *and*

$$\mu_{\gamma\gamma_0} = \frac{1}{\sigma^2} \left(M_{\gamma\gamma_0} - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma\right)' \left(\Sigma_\gamma - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma\right)^{-1} \left(M_{\gamma\gamma_0} - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma\right).$$

PROOF. Note that $Y = X\phi_0 + X_{\gamma_0}h/\sqrt{n} + \varepsilon$ under H_{1n}. Thus,

$$\begin{aligned} T_\gamma &= \frac{1}{\sqrt{n}}\left\{X'_\gamma - X'_\gamma X(X'X)^{-1}X'\right\}\varepsilon + \frac{1}{n}\left\{X'_\gamma - X'_\gamma X(X'X)^{-1}X'\right\}X_{\gamma_0}h \\ &= \frac{1}{\sqrt{n}}\Big(-(X'_\gamma X)(X'X)^{-1}, I\Big)Z'_\gamma\varepsilon + \frac{1}{n}\left\{X'_\gamma - X'_\gamma X(X'X)^{-1}X'\right\}X_{\gamma_0}h. \end{aligned}$$

By (2.6) and Lemmas 2.1-2.2, the conclusion follows directly from the same argument as for Theorem 2.3 in Chan (1990). □

In practice, the values of a and b can be set to empirical quantiles of $\{y_t\}_{t=0}^N$ as in Chan (1991) and Andrews (1993b), although so far how to choose the optimal a, b remains unclear in theory. In this case, we can always find a smallest $n_0 \geq p$ such that y_{n_0-d} stays outside the region $[a, b]$, where the integer n_0 depends on data sample $\{y_0, \cdots, y_N\}$. This means that we can observe $R_{n_0}(\gamma)$, and then further calculate $\{R_t(\gamma)\}_{t=n_0+1}^N$ iteratively by

$$R_t(\gamma) = I(y_{t-d} \leq r_L) + R_{t-1}I(r_L < y_{t-d} \leq r_U).$$

For the remaining observations $\{y_t\}_{t=0}^{n_0-1}$ whose regions are not well identified, we then set their regions to be 0. Thus, we can only use $\tilde{R}_t(\gamma)$ rather than $R_t(\gamma)$ in practice, where

$$\tilde{R}_t(\gamma) = \begin{cases} 0 & \text{for } t = 0, \cdots, n_0 - 1, \\ R_t(\gamma) & \text{for } t = n_0, \cdots, N. \end{cases} \tag{2.7}$$

Let $\tilde{LR}_n$ be defined in the same way as LR_n with $R_t(\gamma)$ being replaced by $\tilde{R}_t(\gamma)$. The following corollary shows that $\tilde{LR}_n$ and LR_n have the same asymptotic property.

COROLLARY 2.1. *If Assumptions 2.1-2.3 hold, then (i) under H_0 it follows that*

$$\tilde{LR}_n \to_d \sup_{\gamma\in\Gamma} G'_\gamma \Omega_\gamma G_\gamma \quad as \quad n \to \infty;$$

(ii) under H_{1n} it follows that

$$\tilde{LR}_n \to_d \sup_{\gamma\in\Gamma}\left\{G'_\gamma \Omega_\gamma G_\gamma + h'\mu_{\gamma\gamma_0}h\right\} \quad as \quad n \to \infty.$$

PROOF. See the Appendix in Zhu, Yu, and Li (2013). □

3. Bootstrapped critical value. In this section, we use a bootstrap method to obtain the critical value for our LR test; see also Hansen (1996) and Li and Li (2011). First, we let

$$\hat{\varepsilon}_t = y_t - x_t(\gamma)'\lambda_n(\gamma) \tag{3.1}$$

with

$$\lambda_n(\gamma) \equiv \arg\min_{\lambda\in\Lambda} \sum_{t=p}^{N} \varepsilon_t^2(\lambda,\gamma) = \left[Z_\gamma' Z_\gamma\right]^{-1}\left[Z_\gamma' Y\right],$$

where Λ is a compact parametric space of λ, and $\varepsilon_t(\lambda,\gamma) = y_t - x_t(\gamma)'\lambda$. Next, we set

$$\hat{LR}_n(\gamma) = \frac{\hat{Z}_n'(\gamma)(X_{1n}(\gamma), I)'[X_{2n}(\gamma)]^{-1}(X_{1n}(\gamma), I)\hat{Z}_n(\gamma)}{\sigma_n^2}, \tag{3.2}$$

where $\hat{\varepsilon} = (\hat{\varepsilon}_p v_p, \cdots, \hat{\varepsilon}_N v_N)'$, $\{v_t\}_{t=p}^N$ is a sequence of i.i.d. $N(0,1)$ random variables, and

$$\hat{Z}_n(\gamma) = \frac{1}{\sqrt{n}} Z_\gamma'\hat{\varepsilon}, \quad X_{1n}(\gamma) = -\frac{X_\gamma' X}{n}\left(\frac{X'X}{n}\right)^{-1},$$
$$\text{and} \quad X_{2n}(\gamma) = \frac{X_\gamma' X_\gamma}{n} - \frac{X_\gamma' X}{n}\left(\frac{X'X}{n}\right)^{-1}\frac{X'X_\gamma}{n}.$$

Define

$$\hat{LR}_n \equiv \sup_{\gamma\in\Gamma} \hat{LR}_n(\gamma). \tag{3.3}$$

The asymptotic theory of $\hat{LR}_n$ is stated in the following theorem:

THEOREM 3.1. *If Assumptions 2.1-2.3 hold, then under H_0 or H_{1n}, it follows that*

$$\hat{LR}_n | y_0, \cdots, y_N \to_d \sup_{\gamma\in\Gamma} G_\gamma'\Omega_\gamma G_\gamma \quad \textit{in probablity as } n\to\infty.$$

PROOF. See the Appendix in Zhu, Yu, and Li (2013). □

REMARK 3.1. *In practice, $\hat{LR}_n$ is calculated with $R_t(\gamma)$ being replaced by $\tilde{R}_t(\gamma)$. However, by using the same argument as for Corollary 2.1, we can show that it does not affect the asymptotic property of $\hat{LR}_n$.*

Note that the conditional limiting distribution in Theorem 3.1 is the same as the null distribution in Theorem 2.1. Then, conditional on the data sample $\{y_0, \cdots, y_N\}$, for any given significance level α, we use the following bootstrap procedure to obtain our critical value:

(i) generate i.i.d. N(0,1) samples $\{v_t\}_{t=p}^{N}$, and then calculate $\hat{LR}_n$ via (3.1)-(3.3);

(ii) repeat step (i) for J times to get $\{\hat{LR}_n^{(1)}, \cdots, \hat{LR}_n^{(J)}\}$;

(iii) choose $c_{n,\alpha}^J$ be the α-th upper percentile of $\{\hat{LR}_n^{(1)}, \cdots, \hat{LR}_n^{(J)}\}$.

From now on, we choose $c_{n,\alpha}^J$ as the critical value for our LR test, i.e., at the significance level α, if $LR_n \geq c_{n,\alpha}^J$, we reject H_0; otherwise, we accept it. In Section 4, we shorten $c_{n,\alpha}^J$ as c_n for brevity.

In the end, we give a critical corollary as follows:

Corollary 3.1. *If Assumptions 2.1-2.3 hold, then (i) under H_0 it follows that*

$$\lim_{n\to\infty} \lim_{J\to\infty} P\left(LR_n \geq c_{n,\alpha}^J\right) = \alpha;$$

(ii) under H_{1n} it follows that

$$\lim_{h\to\infty} \lim_{n\to\infty} \lim_{J\to\infty} P\left(LR_n \geq c_{n,\alpha}^J\right) = 1.$$

Proof. See the Appendix in Zhu, Yu, and Li (2013). □

Corollary 3.1 guarantees that our bootstrapped critical value $c_{n,\alpha}^J$ is asymptotically valid, and our LR test has power to detect H_{1n}. This method is also feasible to obtain the critical value for the LR test in Chan (1990) by setting $\gamma_L \equiv \gamma_U$. Moreover, since $\hat{LR}_n(\gamma)$ is a step-function, the amount of computation on $c_{n,\alpha}^J$ depends only on the effective sample size n and the bootstrapped sample size J. Hence, this will reduce our computational burden significantly in application.

4. Simulation and one real example. In this section, we first compare the performance of our LR test (LR_n) and Chan's (1990) LR test (LR_n^*) in the finite sample. We generate 1000 replications of sample size $n = 200$ from the following BAR model:

$$y_t = y_{t-1} - 0.09 y_{t-2} + (\psi_1 y_{t-1} + \psi_2 y_{t-2}) R_t(\gamma) + \varepsilon_t, \tag{4.1}$$

where $R_t(\gamma)$ is defined as in (1.2) with $d = 1$, ε_t has $N(0,1)$ distribution, and the initial values $y_0 = y_1 = R_1(\gamma) = 0$. We choose $\gamma = (0,0)$, $(0,0.5)$, $(0,1.5)$ or $(0,2)$, and use the significance level $\alpha = 0.05$. Since the pair of characteristic roots is $(0.1, 0.9)$ in the regime of $R_t(\gamma) = 0$, we choose $(\psi_1, \psi_2) = (0, 0)$, $(0.1, -0.09)$, $(0.3, -0.27)$, $(0.5, -0.45)$ or $(0.7, -0.63)$ such that the pair of characteristic roots in the regime of $R_t(\gamma) = 1$ is $(0.1, 0, 9)$, $(0.2, 0.9)$, $(0.4, 0.9)$, $(0.6, 0.9)$ or $(0.8, 0.9)$, respectively. For each replication, the value of a and b for the interval $[a, b]$ are set to be the empirical 10th and 90th quantiles of data sample, the critical value for LR_n is calculated by the bootstrap method in Section 3 with $J = 1000$, and the critical value for LR_n^* is either calculated in the same way as the one for LR_n or taken as 15.18 according to Table 2 in Chan (1991).

TABLE 1
Rejection rates

ψ		γ			LR_n^*	
ψ_1	ψ_2	r_L	r_U	LR_n	LR_{1n}^*	LR_{2n}^*
0.0	0.0	—	—	4.9	4.9	3.4
0.1	-0.09	0.0	0.0	7.7	7.7	3.8
		0.0	0.5	7.5	7.4	3.7
		0.0	1.5	7.6	6.5	3.2
		0.0	2.0	7.5	7.0	5.4
0.3	-0.27	0.0	0.0	31.9	34.2	14.3
		0.0	0.5	30.6	30.3	16.5
		0.0	1.5	33.4	29.6	15.4
		0.0	2.0	32.0	27.1	15.6
0.5	-0.45	0.0	0.0	64.7	69.1	54.0
		0.0	0.5	76.0	79.6	55.2
		0.0	1.5	76.1	75.5	56.0
		0.0	2.0	75.2	72.6	53.9
0.7	-0.63	0.0	0.0	95.8	97.1	86.4
		0.0	0.5	89.4	90.1	89.5
		0.0	1.5	96.0	96.0	87.8
		0.0	2.0	95.9	95.9	89.9

Table 1 lists the rejection rates of LR_n and LR_n^* with different values of ψ and γ. The results for LR_n^* based on the bootstrapped critical value and Chan's (1991) critical value are denoted by LR_{1n}^* and LR_{2n}^*, respectively. The sizes of these tests correspond to the case when $(\psi_1, \psi_2) = (0, 0)$. From Table 1, we find that the sizes of LR_n and LR_{1n}^* are close to their nominal ones, but the size of LR_{2n}^* is very conservative. Although the power of all tests becomes larger as the two regimes for $R_t(\gamma) = 0$ and $R_t(\gamma) = 1$ are more distinguishing, the power of LR_{2n}^* is less than that of LR_n or LR_{1n}^* in all cases. This suggests that the bootstrapped critical values may be more precise than the critical values in Chan (1991) for LR_n^* test. When the distance between r_L and r_U is small, LR_n is less powerful than LR_{1n}^*, and its power is greater than the power of LR_{1n}^* as the distance between r_L and r_U becomes large. As we expected, this is because LR_n (or LR_n^*) is the LR test when r_L and r_U are far from (or closed to) each other. Overall, the simulation results show that LR_n has a good performance especially when the buffering region is wide.

Next, we study the quarterly U.S. real GNP (in 1982 dollars) from the first quarter of 1947 to the first quarter of 1991. Its 100 times log-return, denoted by $\{y_t\}$, has a total of 176 observations; see Figure 1. We apply our test LR_n and the LR test LR_n^* in Chan (1990) to this data set. The results with different values of p and d are reported in Table 2. From Table 2, we find that a marginal threshold effect can be detected at the 5% significance level in either BAR or TAR model with $p = d = 2$. Our finding is consistent to the ones in Potter (1995) and Hansen (1996), in which they also detected a marginal threshold effect in the TAR model by using the sup-LM test. Hence, we fit $\{y_t\}$ by the following two specifications:

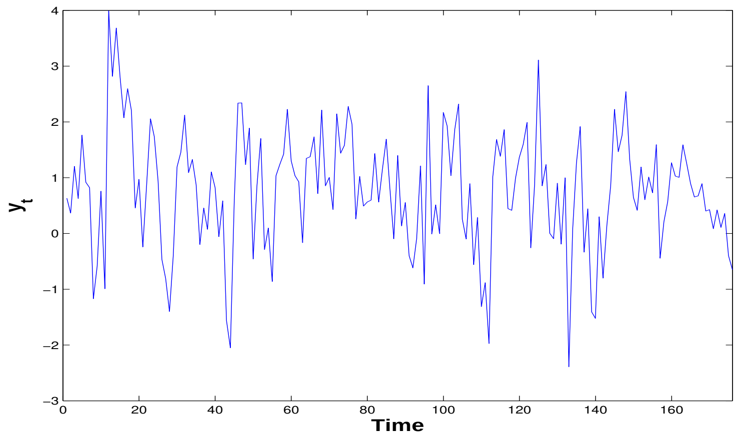

FIG 1. *100 times log-return of quarterly U.S. real GNP (in 1982 dollars) from the first quarter of 1947 to the first quarter of 1991.*

TABLE 2
Results of tests applied to data set $\{y_t\}^\dagger$

		BAR model				TAR model			
p	d	LR_n	$c_{0.1}$	$c_{0.05}$	$c_{0.01}{}^\S$	LR_n^*	$c^*_{0.1}$	$c^*_{0.05}$	$c^*_{0.01}{}^\S$
1	1	4.29	13.66	16.51	23.29	4.29	9.69	11.79	18.58
2	1	9.08	17.97	22.07	30.76	5.83	14.57	17.75	24.92
2	2	21.08‡	18.53	21.36	29.58	13.69‡	12.47	14.52	18.82
3	1	7.18	20.88	23.93	31.63	6.46	15.60	19.10	26.02
3	2	18.15	21.34	24.62	31.70	13.84	14.59	16.70	21.92
3	3	14.38	20.07	23.67	32.50	8.16	17.02	20.83	30.15

† The value of a and b are set to be the 10th and 90th quantiles of $\{y_t\}$.
‡ The p-values for LR_n and LR_n^* are 0.053 and 0.064, respectively.
§ c_α (or c^*_α) is obtained by the bootstrap method in Section 3 with $J = 1000$.

$$
(4.2) \qquad \begin{cases} y_t = \begin{cases} 1.2211 + 0.1597y_{t-1} + 0.4017y_{t-2} + \varepsilon_t & \text{if } R_t = 1 \\ (0.1979) \ \ (0.1236) \qquad (0.1656) & \\ 0.0704 + 0.3754y_{t-1} + 0.3031y_{t-2} + \varepsilon_t & \text{if } R_t = 0 \\ (0.1245) \ \ (0.0856) \qquad (0.0954) & \end{cases}, \\ \text{where} \\ R_t = \begin{cases} 1 & \text{if } y_{t-2} \le -0.617 \\ 0 & \text{if } y_{t-2} > 1.237 \\ R_{t-1} & \text{otherwise} \end{cases} \end{cases}
$$

and

$$
(4.3) \qquad \begin{cases} y_t = \begin{cases} -0.4515 + 0.3924y_{t-1} - 0.8379y_{t-2} + \varepsilon_t & \text{if } R_t = 1 \\ (0.2620) \ \ (0.1400) \qquad (0.2628) & \\ 0.3971 + 0.3241y_{t-1} + 0.1822y_{t-2} + \varepsilon_t & \text{if } R_t = 0 \\ (0.1503) \ \ (0.0845) \qquad (0.1129) & \end{cases}, \\ \text{where} \\ R_t = \begin{cases} 1 & \text{if } y_{t-2} \le -0.008 \\ 0 & \text{otherwise} \end{cases}, \end{cases}
$$

where models (4.2) and (4.3) are estimated by the least squares method with the standard errors in parentheses, and their estimated values of σ_ε^2 are 0.85 and 0.90, respectively. For model (4.2), the first 20 autocorrelations or partial autocorrelations of the residuals $\{\hat{\varepsilon}_t\}$ or $\{\hat{\varepsilon}_t^2\}$ are not significant at the 5% level; see Figure 2. Similar results hold for model (4.3), and hence they are not reported here. Thus, it may imply that both models are adequate to fit $\{y_t\}$. Moreover, the values of log-likelihood for models (4.2) and (4.3) are -233.1 and -237.3, respectively, and hence a BAR(2) model is more suitable than TAR(2) model to fit $\{y_t\}$.

It is interesting to see that models (4.2) and (4.3) basically tell us different stories. Following Tiao and Tsay (1994), if we treat a negative growth in GNP as 'contraction' and a positive growth as 'expansion', model (4.2) shows that the region of y_t does not shift unless we have experienced a big 'contraction' or 'expansion' two years before, while model (4.3) indicates that the region of y_t almost fully relies on the kind of economic status that we have at that time. To our best knowledge, the society or government may not have a big or quick response to a moderate growth in

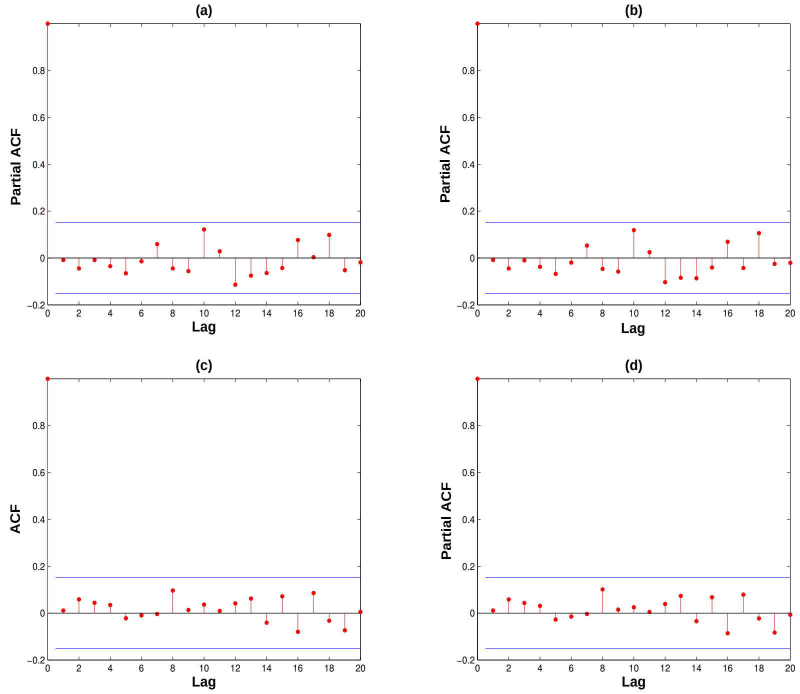

FIG 2. *(a) the autocorrelations for* $\{\hat{\varepsilon}_t\}$*; (b) the partial autocorrelations for* $\{\hat{\varepsilon}_t\}$*; (c) the autocorrelations for* $\{\hat{\varepsilon}_t^2\}$*; and (d) the partial autocorrelations for* $\{\hat{\varepsilon}_t^2\}$*.*

GNP, and hence the region of y_t is most likely unchanged in this case. Thus, based on these facts, it is fair to conclude that a BAR(2) model is more reasonable than TAR(2) model to fit $\{y_t\}$.

In the end, it is also of interest to fit $\{y_t\}$ by a three-regime TAR model as follows:

$$(4.4)\quad y_t = \begin{cases} -0.4969 + 0.3735 y_{t-1} - 0.8500 y_{t-2} + \varepsilon_t \text{ if } y_{t-2} \leq -0.288 \\ \ (0.3649)\ \ (0.1399) \qquad\ (0.3193) \\ -3.3614 + 1.1691 y_{t-1} - 15.872 y_{t-2} + \varepsilon_t \text{ if } -0.288 < y_{t-2} \leq -0.058 \\ \ (1.2807)\ \ (1.0193) \qquad\ (4.3454) \\ 0.3837 + 0.3233 y_{t-1} + 0.1908 y_{t-2} + \varepsilon_t \quad \text{if } y_{t-2} > -0.058 \\ (0.1439)\ \ (0.0818) \qquad\ (0.1083) \end{cases},$$

where model (4.4) is estimated by the least squares method with the standard errors in parentheses, and the estimated value of σ_ε^2 is 0.84. As model (4.2), model (4.4) may also be adequate to fit $\{y_t\}$ by looking at the first 20 autocorrelations and partial autocorrelations of the residuals $\{\hat{\varepsilon}_t\}$ and $\{\hat{\varepsilon}_t^2\}$. However, the number of effective observations for these regimes from lower to upper are 25, 10 and 139, respectively. Thus, although the value of log-likelihood for model (4.4) is -231.6 greater than that for model (4.2), a model with two regimes for $\{y_t\}$ seems more likely. Therefore, compared to model (4.4), we prefer to fit $\{y_t\}$ by a BAR(2) model in view of this point.

ACKNOWLEDGEMENTS

The authors greatly appreciate the helpful comments of Dr. G.D. Li, two anonymous referees, the Associate Editor, and the Editor Qiwei Yao. The authors would to thank the Research Grants Council of the Hong Kong SAR Government, GRF grant HKU703711P, for partial support. The first author's research is also supported by NSFC(No.11201459) and the National Center for Mathematics and Interdisciplinary Sciences, CAS.

REFERENCES

[1] Andrews, D.W.K. (1993a) Tests for parameter instability and structural change with unknow change point. *Econometrica* **61**, 821-856.

[2] Andrews, D.W.K. (1993b) An introduction to econometric applications of functional limit theory for dependent random variables. *Econometric Reviews* **12**, 183-216.

[3] Caner, M. and Hansen, B.E. (2001) Threshold autoregression with a unit root. *Econometrica* **69**, 1555-1596.

[4] Chan, K.S. (1990) Testing for threshold autoregression. *Annals of Statistics* **18**, 1886-1894.

[5] Chan, K.S. (1991) Percentage points of likelihood ratio tests for threshold autoregression. *Journal of the Royal Statistical Society Series B* **53**, 691-696.

[6] Chan, K.S. and Tong, H. (1990) On likelihood ratio tests for threshold autoregression. *Journal of the Royal Statistical Society Series B* **52**, 469-476.

[7] Davies, R.B. (1977) Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrica* **64**, 247-254.

[8] Davies, R.B. (1987) Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrica* **74**, 33-43.

[9] Doukhan, P., Massart, P., and Rio, E. (1995) Invariance principles for absolutely regular empirical processes. *Annales de I'Institut H. Poincare* **31** 393-427.

[10] Hansen, B.E. (1996) Inference when a nuisance parameter is not indentified under the null hypothesis. *Econometrica* **64**, 413-430.

[11] Hansen, B.E. (1999) Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics* **93**, 345-368.

[12] Li, G.D., Guan, B., Li, W.K., and Yu, P.L.H. (2012) Buffered threshold autoregressive time series models. Working paper. University of Hong Kong.

[13] Li, G.D. and Li, W.K. (2008) Testing for threshold moving average with conditional heteroscedasticity. *Statistica Sinica* **18**, 647-665.

[14] Li, G.D. and Li, W.K. (2011) Testing a linear time series models against its threshold extension. *Biometrika* **98**, 243-250.

[15] Li, D. and Ling, S. (2013) On a threshold double autoregressive model. Working paper. Hong Kong University of Science and Technology.

[16] Ling, S. and Tong, H. (2005) Testing a linear MA model against threshold MA models. *Annals of Statistics* **33**, 2529-2552.

[17] Pham, T.D. and Tran, L.T. (1985) Some mixing properties of time series models. *Stochastic Processes and their applications* **19**, 297-303.

[18] Potter, S.M. (1995) A nonlinear approach to U.S. GNP. *Journal of Applied Econometrics* **10**, 109-125.

[19] Stute, W. (1997) Nonparametric model checks for regression. *Annals of Statistics* **25**, 613-641.

[20] Tiao, G.C. and Tsay, R.S. (1994) Some advances in non-linear and adaptive modelling in time-series. *Journal of Forecasting* **13**, 109-131.

[21] Tong, H. (1978) On a threshold model. In Pattern Recognition and Signal Processing (C.H. Chen, ed.) 101-141. Sijthoff and Noordhoff, Amsterdam.

[22] Tong, H. (1990) Non-linear Time Series. A Dynamical System Approach. Clarendon Press, Oxford.

[23] Tong, H. (2011) Threshold models in time series analysis–30 years on (with discussions). *Statistics and Its Interface* **4**, 107-135.

[24] Tsay, R.S. (1989) Testing and modeling threshold autoregressive processes. *Journal of American Statistical Association* **84**, 231-240.

[25] Tsay, R.S. (1998) Testing and modeling multivariate threshold models. *Journal of American Statistical Association* **93**, 1188-1202.

[26] Wong, C.S. and Li, W.K. (1997) Testing for threshold autoregression with conditional heteroscedasticity. *Biometrika* **84**, 407-418.

[27] Wong, C.S. and Li, W.K. (2000) Testing for double threshold autoregressive conditional heteroscedastic model. *Statistica Sinica* **10**, 173-189.

[28] Zhu, K. and Ling, S. (2012) Likelihood ratio tests for the structural change of an AR(p) model to a threshold AR(p) model. *Journal of Time Series Analysis* **33**, 223-232.

[29] Zhu, K., Yu, P.L.H. and Li, W.K. (2013) Testing for the buffered autoregressive processes (Supplementary Material).

Chinese Academy of Sciences
Institute of Applied Mathematics
HaiDian District, Zhongguancun
Bei Jing, China
E-mail: kzhu@amss.ac.cn

Department of Statistics and Actuarial Science
University of Hong Kong
Pokfulam Road, Hong Kong
E-mail: plhyu@hku.hk
hrntlwk@hku.hk

TESTING FOR THE BUFFERED AUTOREGRESSIVE PROCESSES (SUPPLEMENTARY MATERIAL)

By Ke Zhu, Philip L.H. Yu and Wai Keung Li

Chinese Academy of Sciences and University of Hong Kong

APPENDIX: PROOFS

In this appendix, we first give the proofs of Lemmas 2.1-2.2. Denote C as a generic constant which may vary from place to place in the rest of this paper. The proofs of Lemmas 2.1-2.2 rely on the following three basic lemmas:

Lemma A.1. *Suppose that y_t is strictly stationary, ergodic and absolutely regular with mixing coefficients $\beta(m) = O(m^{-A})$ for some $A > v/(v-1)$ and $r > v > 1$; and there exists an $A_0 > 1$ such that $2A_0rv/(r-v) < A$. Then, for any $\gamma = (r_L, r_U) \in \Gamma$, we have*

$$\sum_{j=1}^{\infty} \left\{ E\left[\prod_{i=1}^{j} I(r_L < y_{t-i} \leq r_U) \right] \right\}^{(r-v)/2A_0rv} < \infty.$$

Proof. First, denote $\xi_i = I(r_L < y_{t-i} \leq r_U)$. Then, ξ_i is strictly stationary, ergodic and α-mixing with mixing coefficients $\alpha(m) = O(m^{-A})$. Next, take $\iota \in \Big([2A_0rv/(r-v)+1]/(A+1), 1\Big)$, and let $p = \lfloor j^\iota \rfloor$ and $s = \lfloor j/j^\iota \rfloor$, where $\lfloor x \rfloor$ is the largest integer not greater than x. When $j \geq j_0$ is large enough, we can always find $\{\xi_{kp+1}\}_{k=0}^{s-1}$, a subsequence of $\{\xi_i\}_{i=1}^{j}$.

Furthermore, let $\mathcal{F}_m^n = \sigma(\xi_i, m \leq i \leq n)$. Then, $\xi_{kp+1} \in \mathcal{F}_{kp+1}^{kp+2}$. Note that $E[\xi_{kp+1}] < P(a \leq y_t \leq b) \triangleq \rho \in (0,1)$. Hence, by Proposition 2.6 in Fan and Yao (2003, p.72), we have that for $j \geq j_0$,

$$\begin{aligned}
E\left[\prod_{i=1}^{j} \xi_i\right] &\leq E\left[\prod_{k=0}^{s-1} \xi_{kp+1}\right] \\
&= \left\{ E\left[\prod_{k=0}^{s-1} \xi_{kp+1}\right] - \prod_{k=0}^{s-1} E[\xi_{kp+1}] \right\} + \prod_{k=0}^{s-1} E[\xi_{kp+1}] \\
&\leq 16(s-1)\alpha(p) + \rho^s \\
&\leq C\lfloor j/j^\iota \rfloor \lfloor j^\iota \rfloor^{-A} + \rho^{\lfloor j/j^\iota \rfloor}.
\end{aligned}$$

Therefore, since $(r-v)/2A_0rv > 0$, by using the inequality $(x+y)^k \leq C(x^k+y^k)$ for any $x, y, k > 0$, it follows that

$$\begin{aligned} \sum_{j=1}^{\infty} \left\{ E\left[\prod_{i=1}^{j} \xi_i\right] \right\}^{(r-v)/2A_0rv} &\leq (j_0-1) + \sum_{j=j_0}^{\infty} \left\{ E\left[\prod_{i=1}^{j} \xi_i\right] \right\}^{(r-v)/2A_0rv} \\ &\leq (j_0-1) + C\sum_{j=j_0}^{\infty} \left[\lfloor j/j^{\iota}\rfloor \lfloor j^{\iota}\rfloor^{-A}\right]^{(r-v)/2A_0rv} \\ &\quad + C\sum_{j=j_0}^{\infty} \rho^{\lfloor j/j^{\iota}\rfloor (r-v)/2A_0rv}. \end{aligned} \tag{A.1}$$

Since $\iota > [2A_0rv/(r-v)+1]/(A+1)$, we have $(\iota A + \iota - 1)(r-v)/2A_0rv > 1$, and hence $\sum_{j=1}^{\infty} j^{-(\iota A+\iota-1)(r-v)/2A_0rv} < \infty$, which implies that

$$\sum_{j=j_0}^{\infty} \left[\frac{\lfloor j/j^{\iota}\rfloor}{\lfloor j^{\iota}\rfloor^{A}}\right]^{(r-v)/2A_0rv} \leq \sum_{j=1}^{\infty} \left[\frac{j}{j^{\iota}(j^{\iota}-1)^{A}}\right]^{(r-v)/2A_0rv} < \infty. \tag{A.2}$$

On the other hand, since $\left(\rho^{\lfloor j/j^{\iota}\rfloor (r-v)/2A_0rv}\right)^{1/j} < 1$, by Cauchy's root test, we have

$$\sum_{j=j_0}^{\infty} \rho^{\lfloor j/j^{\iota}\rfloor (r-v)/2A_0rv} < \sum_{j=1}^{\infty} \rho^{\lfloor j/j^{\iota}\rfloor (r-v)/2A_0rv} < \infty. \tag{A.3}$$

Now, the conclusion follows directly from (A.1)-(A.3). This completes the proof. $\square$

LEMMA A.2. *Suppose that the conditions in Lemma A.1 hold, and y_t has a bounded and continuous density function. Then, there exists a $B_0 > 1$ such that for any $\gamma_1, \gamma_2 \in \Gamma$, we have*

$$\|R_t(\gamma_1) - R_t(\gamma_2)\|_{2rv/(r-v)} \leq C|\gamma_1 - \gamma_2|^{(r-v)/2B_0rv}.$$

PROOF. Let $\gamma_1 = (r_{1L}, r_{1U})$ and $\gamma_2 = (r_{2L}, r_{2U})$. Since $R_t(\gamma) = I(y_{t-d} \leq r_L) + R_{t-1}(\gamma)I(r_L < y_{t-d} \leq r_U)$, we have

$$R_t(\gamma_1) - R_t(\gamma_2) = \Delta_t(\gamma_1, \gamma_2) + I(r_{1L} < y_{t-d} \leq r_{1U})\left[R_{t-1}(\gamma_1) - R_{t-1}(\gamma_2)\right],$$

where

$$\begin{aligned} \Delta_t(\gamma_1, \gamma_2) &= I(r_{2L} < y_{t-d} \leq r_{1L}) \\ &\quad + R_{t-1}(\gamma_2)\left[I(r_{1L} < y_{t-d} \leq r_{1U}) - I(r_{2L} < y_{t-d} \leq r_{2U})\right]. \end{aligned}$$

Thus, by iteration we can show that

$$
\begin{aligned}
&R_t(\gamma_1) - R_t(\gamma_2) \\
&= \Delta_t(\gamma_1,\gamma_2) + \sum_{j=1}^{\infty} \Delta_{t-j}(\gamma_1,\gamma_2) \prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \le r_{1U}).
\end{aligned} \tag{A.4}
$$

Next, for brevity, we assume that $r_{2L} \le r_{1L} \le r_{2U} \le r_{1U}$, because the proofs for other cases are similar. Note that for any $j \ge 0$, $R_{t-j-1}(\gamma_2) \le 1$ and

$$
\begin{aligned}
&I(r_{1L} < y_{t-j-d} \le r_{1U}) - I(r_{2L} < y_{t-j-d} \le r_{2U}) \\
&\quad = I(r_{2U} < y_{t-j-d} \le r_{1U}) - I(r_{2L} < y_{t-j-d} \le r_{1L}).
\end{aligned}
$$

Let $f(x)$ be the density function of y_t. Since $\sup_x f(x) < \infty$ and $|\Delta_{t-j}(\gamma_1,\gamma_2)| \le 2$, by Hölder's inequality and Taylor's expansion, it follows that for any $s \ge 1$,

$$
\begin{aligned}
E|\Delta_{t-j}(\gamma_1,\gamma_2)|^s &\le 2^{s-1} E|\Delta_{t-j}(\gamma_1,\gamma_2)| \\
&\le 2^{s-1}\left[2\sup_x f(x)|r_{1L} - r_{2L}| + \sup_x f(x)|r_{1U} - r_{2U}|\right] \\
&\le C|\gamma_1 - \gamma_2|.
\end{aligned} \tag{A.5}
$$

Let $A_0 > 1$ be specified in Lemma A.1, and choose B_0 such that $1/A_0 + 1/B_0 = 1$. By Hölder's inequality and (A.5), we can show that

$$
\begin{aligned}
&E\left|\Delta_{t-j}(\gamma_1,\gamma_2)\prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \le r_{1U})\right|^{2rv/(r-v)} \\
&\le \left\{E[\Delta_{t-j}(\gamma_1,\gamma_2)]^{2B_0 rv/(r-v)}\right\}^{1/B_0} \\
&\quad \times \left\{E\left[\prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \le r_{1U})\right]\right\}^{1/A_0} \\
&\le 2^{[2B_0 rv/(r-v)]-1}\left\{E|\Delta_{t-j}(\gamma_1,\gamma_2)|\right\}^{1/B_0} \\
&\quad \times \left\{E\left[\prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \le r_{1U})\right]\right\}^{1/A_0} \\
&\le C|\gamma_1-\gamma_2|^{1/B_0}\left\{E\left[\prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \le r_{1U})\right]\right\}^{1/A_0}.
\end{aligned} \tag{A.6}
$$

By (A.4)-(A.6), Minkowski's inequality, Lemma A.1 and the compactness of Γ, we have

$$
\begin{aligned}
&\|R_t(\gamma_1) - R_t(\gamma_2)\|_{2rv/(r-v)} \\
&\quad \le C|\gamma_1 - \gamma_2|^{(r-v)/2rv} + C|\gamma_1 - \gamma_2|^{(r-v)/2B_0 rv} \\
&\qquad \times \sum_{j=1}^{\infty} \left\{ E\left[\prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \le r_{1U}) \right] \right\}^{(r-v)/2A_0 rv} \\
&\quad \le C|\gamma_1 - \gamma_2|^{(r-v)/2B_0 rv}.
\end{aligned}
$$

This completes the proof. □

LEMMA A.3. *Suppose that the conditions in Lemma A.2 hold and $Ey_t^4 < \infty$. Then,*

$$
\sup_{\gamma \in \Gamma} \left| \frac{X_\gamma' X}{n} - \Sigma_\gamma \right| \to 0 \quad a.s. \ \ as\ n \to \infty.
$$

PROOF. For brevity, we only prove the uniform convergence for $n^{-1}\sum_{t=1}^{n} \phi_t(\gamma)$, the last component of $n^{-1}X_\gamma' X$, where

$$
\phi_t(\gamma) = y_{t-p}^2 R_t(\gamma).
$$

First, for fix $\varepsilon > 0$, we partition Γ by $\{B_1, \cdots, B_{K_\varepsilon}\}$, where $B_k = \{(r_L, r_U); \omega_k < r_L \le \omega_{k+1}, \nu_k < r_U \le \nu_{k+1}\} \cap \Gamma$. Here, $\{\omega_k\}$ and $\{\nu_k\}$ are chosen such that

$$
(\omega_{k+1} - \omega_k)^{(r-v)/2B_0 rv} < C_1\varepsilon \ \text{ and } \ (\nu_{k+1} - \nu_k)^{(r-v)/2B_0 rv} < C_1\varepsilon, \tag{A.7}
$$

where $B_0 > 1$ is specified as in Lemma A.2, and $C_1 > 0$ will be selected later.

Next, we set

$$
f_t^u(\varepsilon) = y_{t-p}^2 R_t(\omega_{k+1}, \nu_{k+1}) \quad \text{and} \quad f_t^l(\varepsilon) = y_{t-p}^2 R_t(\omega_k, \nu_k).
$$

By construction, since $R_t(\gamma)$ is a nondecreasing function with respect to r_L and r_U, for any $\gamma \in \Gamma$, there is some k such that $\gamma \in B_k$ and $f_t^l(\varepsilon) \le \phi_t(\gamma) \le f_t^u(\varepsilon)$.

Furthermore, since $rv/(2rv - r + v) < 1$, we have

$$
\left\| y_{t-p}^2 \right\|_{2rv/(2rv-r+v)} < \left\| y_{t-p}^2 \right\|_2 < \infty.
$$

Thus, by Hölder's inequality, Lemma A.2 and (A.7), we have

$$
\begin{aligned}
E&\left[f_t^u(\varepsilon) - f_t^l(\varepsilon)\right] \\
&\le \left\|y_{t-p}^2\right\|_{2rv/(2rv-r+v)} \left\|R_t(\omega_{k+1}, \nu_{k+1}) - R_t(\omega_k, \nu_k)\right\|_{2rv/(r-v)} \\
&\le C\left[(\omega_{k+1} - \omega_k)^{(r-v)/2B_0rv} + (\nu_{k+1} - \nu_k)^{(r-v)/2B_0rv}\right] \\
&\le 2CC_1\varepsilon.
\end{aligned}
$$

By setting $C_1 = (2C)^{-1}$, we have $E\left[f_t^u(\varepsilon) - f_t^l(\varepsilon)\right] \le \varepsilon$. Thus, the conclusion holds according to Theorem 2 in Pollard (1984, p.8). This completes the proof. □

Proof of Lemma 2.1. First, since $K_{\gamma\gamma}$ is positive definite by Assumption 2.1, we know that both Σ and Σ_γ are positive definite. By using the same argument as for Lemma 2.1(iv) in Chan (1990), it is not hard to show that for every $\gamma \in \Gamma$, $\Sigma_\gamma - \Sigma_\gamma\Sigma^{-1}\Sigma_\gamma'$ is positive definite. Second, by the ergodic theorem, it is easy to see that

$$
\frac{X'X}{n} \to \Sigma \quad \text{a.s. as } n \to \infty. \tag{A.8}
$$

Third, by Lemma A.3 we have

$$
\sup_{\gamma\in\Gamma}\left|\frac{X_\gamma'X}{n} - \Sigma_\gamma\right| \to 0 \quad \text{and} \quad \sup_{\gamma\in\Gamma}\left|\frac{X_\gamma'X_\gamma}{n} - \Sigma_\gamma\right| \to 0 \quad \text{a.s.} \tag{A.9}
$$

as $n \to \infty$. Note that if H_0 holds, we have

$$
\begin{aligned}
T_\gamma &= \frac{1}{\sqrt{n}}\left\{X_\gamma' - X_\gamma'X(X'X)^{-1}X'\right\}\varepsilon \\
&= \frac{1}{\sqrt{n}}\Big(-(X_\gamma'X)(X'X)^{-1}, I\Big)Z_\gamma'\varepsilon.
\end{aligned}
$$

Then, (i) and (ii) follow readily from (A.8)-(A.9). This completes the proof. □

Proof of Lemma 2.2. Denote

$$
G_n(\gamma) \equiv \frac{1}{\sqrt{n}}Z_\gamma'\varepsilon = \frac{1}{\sqrt{n}}\sum_{t=p}^{N} x_t(\gamma)\varepsilon_t.
$$

It is straightforward to show that the finite dimensional distribution of $\{G_n(\gamma)\}$ converges to that of $\{\sigma G_\gamma\}$. By Pollard (1990, Sec.10), we only need to verify the stochastic equicontinuity of $\{G_n(\gamma)\}$. To establish it, we use Theorem 1, Application

4 in Doukhan, Massart, and Rio (1995, p.405); see also Andrews (1993) and Hansen (1996).

First, the envelop function is $\sup_\gamma |x_t(\gamma)\varepsilon_t| = \bar{x}_t|\varepsilon_t|$, where $\bar{x}_t = \sup_\gamma |x_t(\gamma)|$. By Hölder's inequality and Assumption 2.1, we know that the envelop function is L^{2v} bounded. Next, for any $\gamma_1, \gamma_2 \in \Gamma$, by Assumptions 2.1-2.3, Lemma A.2 and Hölder's inequality, we have

$$\begin{aligned}
\|x_t(\gamma_1)\varepsilon_t - x_t(\gamma_2)\varepsilon_t\|_{2v} &= \|h_t(\gamma_1)\varepsilon_t - h_t(\gamma_2)\varepsilon_t\|_{2v} \\
&\leq \|x_t\varepsilon_t\|_{2r}\|R_t(\gamma_1) - R_t(\gamma_2)\|_{2rv/(r-v)} \\
&\leq C\|x_t\|_{4r}\|\varepsilon_t\|_{4r}\,|\gamma_1 - \gamma_2|^{(r-v)/2B_0rv} \\
&\leq C|\gamma_1 - \gamma_2|^{(r-v)/2B_0rv}
\end{aligned}$$

for some $B_0 > 1$, where the last inequality holds since $\|x_t\|_{4r}\|\varepsilon_t\|_{4r} < \infty$.

Now, following the argument in Hansen (1996, p.426), we know that $G_n(\gamma)$ is stochastically equicontinuous. This completes the proof. $\square$

Next, we give Lemmas A.4-A.6, in which Lemma A.4 is crucial for proving Lemma A.5, and Lemmas A.5 and A.6 are needed to prove Corollary 2.1 and Theorem 3.1, respectively.

LEMMA A.4. *Suppose that y_t is strictly stationary and ergodic. Then, (i) $n_0 = O_p(1)$; (ii) furthermore, if $E|y_t|^2 < \infty$ and $E|\varepsilon_t|^2 < \infty$, for any $a_n = o(1)$, we have*

$$\sup_{\gamma\in\Gamma}\left|a_n \sum_{t=p}^{n_0-1} x_t x_t' R_t(\gamma)\right| = O_p(1) \tag{A.10}$$

and

$$\sup_{\gamma\in\Gamma}\left|a_n \sum_{t=p}^{n_0-1} h_t(\gamma)\varepsilon_t\right| = O_p(1). \tag{A.11}$$

PROOF. First, by the ergodic theory, we have that

$$\frac{1}{M}\sum_{t=p}^{M} I(a \leq y_{t-d} \leq b) = P(a \leq y_{t-d} \leq b) \triangleq \kappa > 0 \quad \text{a.s.}$$

as $M \to \infty$. Thus, $\forall \eta > 0$, there exists an integer $M(\eta) > 0$ such that

$$P\left(\frac{1}{M}\sum_{t=p}^{M} I(a \leq y_{t-d} \leq b) < \frac{\kappa}{2}\right) < \eta.$$

By the definition of n_0, it follows that

$$
\begin{aligned}
P\left(n_0 > M\right) &= P\left(\sum_{t=p}^{M} I(a \le y_{t-d} \le b) = 0\right) \\
&= P\left(\frac{1}{M}\sum_{t=p}^{M} I(a \le y_{t-d} \le b) = 0\right) \\
&\le P\left(\frac{1}{M}\sum_{t=p}^{M} I(a \le y_{t-d} \le b) < \frac{\kappa}{2}\right) \\
&< \eta, \qquad \text{(A.12)}
\end{aligned}
$$

i.e., (i) holds. Furthermore, by taking $\tilde{M} = M^2$, from (A.12) and Markov's inequality, it follows that $\forall \eta > 0$,

$$
\begin{aligned}
&P\left(\sup_{\gamma\in\Gamma}\left|a_n \sum_{t=p}^{n_0-1} x_t x_t' R_t(\gamma)\right| > \tilde{M}\right) \\
&= P\left(\sup_{\gamma\in\Gamma}\left|a_n \sum_{t=p}^{n_0-1} x_t x_t' R_t(\gamma)\right| > \tilde{M}, n_0 \le M\right) \\
&\le P\left(\max_{p\le k\le M}\sup_{\gamma\in\Gamma}\left|a_n \sum_{t=p}^{k-1} x_t x_t' R_t(\gamma)\right| > \tilde{M}\right) \\
&\le \sum_{k=p}^{M} P\left(a_n \sum_{t=p}^{k-1} |x_t|^2 > \tilde{M}\right) \\
&\le a_n \sum_{k=p}^{M}\sum_{t=p}^{k-1} \frac{E|x_t|^2}{\tilde{M}} \\
&= O\left(\frac{a_n M^2}{\tilde{M}}\right) = O\left(a_n\right) < \eta \qquad \text{(A.13)}
\end{aligned}
$$

as n is large enough. Thus, we know that equation (A.10) holds. Next, by Hölder's inequality and a similar argument as for (A.13), it is not hard to show that $\forall \eta > 0$,

$$
P\left(\sup_{\gamma\in\Gamma}\left|a_n \sum_{t=p}^{n_0-1} h_t(\gamma)\varepsilon_t\right| > \tilde{M}\right) \le O\left(a_n\right) < \eta
$$

as n is large enough, i.e., (A.11) holds. This completes the proof. □

LEMMA A.5. *If Assumptions 2.1-2.3 hold, then it follows that under H_0 or H_{1n},*

$$(i)\ \sup_{\gamma\in\Gamma}\left|\frac{1}{n}\left(X_\gamma-\tilde{X}_\gamma\right)'X\right|=o_p(1),$$
$$(ii)\ \sup_{\gamma\in\Gamma}\left|\frac{1}{n}\left(X_\gamma'X_\gamma-\tilde{X}_\gamma'\tilde{X}_\gamma\right)\right|=o_p(1),$$
$$(iii)\ \sup_{\gamma\in\Gamma}\left|T_\gamma-\tilde{T}_\gamma\right|=o_p(1),$$

where $\tilde{X}_\gamma$ and $\tilde{T}_\gamma$ are defined in the same way as X_γ and T_γ, respectively, with $R_t(\gamma)$ being replaced by $\tilde{R}_t(\gamma)$.

PROOF. (i) Note that

$$\frac{1}{n}\left(X_\gamma-\tilde{X}_\gamma\right)'X=\frac{1}{\sqrt{n}}\left[\frac{1}{\sqrt{n}}\sum_{t=p}^{n_0-1}x_tx_t'R_t(\gamma)\right].$$

Hence, we know that (i) holds by taking $a_n=n^{-1/2}$ in equation (A.10).

(ii) By a similar argument as for (i), we can show that (ii) holds.

(iii) Note that when $\lambda_0=(\phi_0',h'/\sqrt{n})'$, we have

$$\begin{aligned}T_\gamma-\tilde{T}_\gamma=&\frac{1}{\sqrt{n}}\left(X_\gamma-\tilde{X}_\gamma\right)'\varepsilon-\frac{1}{\sqrt{n}}\left(X_\gamma-\tilde{X}_\gamma\right)'X(X'X)^{-1}X'\varepsilon\\&-\frac{1}{n}\left(X_\gamma-\tilde{X}_\gamma\right)'X(X'X)^{-1}X'X_{\gamma_0}h\\&-\frac{1}{n}\tilde{X}_\gamma'X(X'X)^{-1}X'\left(X_{\gamma_0}-\tilde{X}_{\gamma_0}\right)h\\&+\frac{1}{n}\left(X_\gamma'X_{\gamma_0}-\tilde{X}_\gamma'\tilde{X}_{\gamma_0}\right)h\\&\triangleq I_{1n}(\gamma)-I_{2n}(\gamma)-I_{3n}(\gamma)-I_{4n}(\gamma)+I_{5n}(\gamma)\ \text{ say.}\end{aligned}$$

First, since

$$I_{1n}(\gamma)=\frac{1}{n^{1/4}}\left[\frac{1}{n^{1/4}}\sum_{t=p}^{n_0-1}h_t(\gamma)\varepsilon_t\right],$$

it follows that $\sup_\gamma|I_{1n}(\gamma)|=o_p(1)$ by taking $a_n=n^{-1/4}$ in equation (A.11). Next, since

$$I_{2n}(\gamma)=\left[\frac{1}{n}\sum_{t=p}^{n_0-1}x_tx_t'R_t(\gamma)\right]\left(\frac{X'X}{n}\right)^{-1}\frac{X'\varepsilon}{\sqrt{n}},$$

we have that $\sup_\gamma|I_{2n}(\gamma)|=o_p(1)$ from (i). Similarly, we can show that $\sup_\gamma|I_{in}(\gamma)|=o_p(1)$ for $i=3,4,5$. Hence, under H_0 (i.e., $h\equiv0$) or H_{1n}, we know that (iii) holds. This completes the proof. □

LEMMA A.6. *If Assumptions 2.1-2.3 hold, then it follows that under H_0 or H_{1n},*

$$\sup_{\gamma\in\Gamma}\sqrt{n}|\lambda_n(\gamma)-\lambda_0| = O_p(1).$$

PROOF. First, for any $\gamma\in\Gamma$, by Taylor's expansion we have

$$\begin{aligned}
&\sum_{t=p}^{N}\left[\varepsilon_t^2(\lambda_n(\gamma),\gamma)-\varepsilon_t^2(\lambda_0,\gamma)\right]\\
&\quad = -(\lambda_n(\gamma)-\lambda_0)'\left(\sum_{t=p}^{N}2\varepsilon_t(\lambda_0,\gamma)x_t(\gamma)\right)\\
&\qquad + (\lambda_n(\gamma)-\lambda_0)'\left(\sum_{t=p}^{N}x_t(\gamma)x_t(\gamma)'\right)(\lambda_n(\gamma)-\lambda_0).
\end{aligned}\tag{A.14}$$

Next, when $\lambda_0 = (\phi_0', h'/\sqrt{n})'$, we can show that

$$\begin{aligned}
\frac{1}{\sqrt{n}}\sum_{t=p}^{N}\varepsilon_t(\lambda_0,\gamma)x_t(\gamma) &= \frac{1}{\sqrt{n}}Z_\gamma'\varepsilon + \frac{1}{\sqrt{n}}\sum_{t=p}^{N}x_t(\gamma)[x_t(\gamma_0)-x_t(\gamma)]'\lambda_0\\
&= \frac{1}{\sqrt{n}}Z_\gamma'\varepsilon + \frac{1}{n}\sum_{t=p}^{N}\begin{pmatrix} x_tx_t'[R_t(\gamma_0)-R_t(\gamma)] \\ x_tx_t'[R_t(\gamma)R_t(\gamma_0)-R_t(\gamma)]\end{pmatrix}h\\
&\triangleq G_n^*(\gamma).
\end{aligned}\tag{A.15}$$

Let $\lambda_{min}(\gamma) > 0$ be the minimum eigenvalue of $K_{\gamma\gamma}$. Then, by equations (A.14)-(A.15), $\forall\eta>0$, there exists a $M(\eta)>0$ such that

$$\begin{aligned}
&P\left(\sup_{\gamma\in\Gamma}\sqrt{n}|\lambda_n(\gamma)-\lambda_0| > M\right)\\
&= P\Bigg(\sqrt{n}|\lambda_n(\gamma)-\lambda_0| > M,\ \sum_{t=p}^{N}\left[\varepsilon_t^2(\lambda_n(\gamma),\gamma)-\varepsilon_t^2(\lambda_0,\gamma)\right]\le 0\\
&\qquad\quad \text{for some } \gamma\in\Gamma)\\
&\le P\Big(\sqrt{n}|\lambda_n(\gamma)-\lambda_0| > M,\ -2\sqrt{n}|\lambda_n(\gamma)-\lambda_0||G_n^*(\gamma)|\\
&\qquad\quad +n|\lambda_n(\gamma)-\lambda_0|^2[\lambda_{min}(\gamma)+o_p(1)]\le 0 \text{ for some } \gamma\in\Gamma\Big)\\
&\le P\Big(M<\sqrt{n}|\lambda_n(\gamma)-\lambda_0|\le 2[\lambda_{min}(\gamma)+o_p(1)]^{-1}|G_n^*(\gamma)|\\
&\qquad\quad \text{for some } \gamma\in\Gamma)\\
&\le P\left(|G_n^*(\gamma)| > M[\lambda_{min}(\gamma)+o_p(1)]/2 \text{ for some } \gamma\in\Gamma\right)\\
&\le \eta,
\end{aligned}$$

where the last inequality holds because $G_n^*(\gamma) = O_p(1)$ by Lemma 2.2 and Lemma A.3. Hence, under H_0 (i.e., $h \equiv 0$) or H_{1n}, our conclusion holds. This completes the proof. $\square$

Proof of Corollary 2.1. The conclusion follows directly from Theorems 2.1-2.2 and Lemma A.5. $\square$

Proof of Theorem 3.1. We use the method in the proof of Theorem 2 in Hansen (1996). Let W denote the set of samples ω for which

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=p}^{N} \sup_{\gamma\in\Gamma} |x_t(\gamma)|\varepsilon_t^2 < \infty \text{ a.s.,} \tag{A.16}$$

$$\lim_{n\to\infty} \sup_{\gamma,\delta\in\Gamma} \left| \frac{1}{n} \sum_{t=p}^{N} x_t(\gamma)x_t(\delta)'\varepsilon_t^2 - \sigma^2 K_{\gamma\delta} \right| \to 0 \;\text{ a.s.} \tag{A.17}$$

Since $\sup_{\gamma\in\Gamma} |x_t(\gamma)| \le \sqrt{2}|x_t|$ and $E|x_t|\varepsilon_t^2 < \infty$ due to Assumption 2.1, by the ergodic theorem we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{t=p}^{N} \sup_{\gamma\in\Gamma} |x_t(\gamma)|\varepsilon_t^2 \le \lim_{n\to\infty} \frac{\sqrt{2}}{n} \sum_{t=p}^{N} |x_t|\varepsilon_t^2 < \infty \text{ a.s.,}$$

i.e., (A.16) holds. Furthermore, by Assumptions 2.1-2.3 and a similar argument as for Lemma A.3, it is not hard to see that

$$\lim_{n\to\infty} \sup_{\gamma,\delta\in\Gamma} \left| \frac{1}{n} \sum_{t=p}^{N} x_t(\gamma)x_t(\delta)'\varepsilon_t^2 - \sigma^2 K_{\gamma\delta} \right| \to 0 \;\text{ a.s.,}$$

i.e., (A.17) holds. Thus, $P(W) = 1$. Take any $\omega \in W$. For the remainder of the proof, all operations are conditionally on ω, and hence all of the randomness appears in the i.i.d. N(0, 1) variables $\{v_t\}$.

Define

$$Z_n^*(\gamma) = \frac{1}{\sqrt{n}} \sum_{t=p}^{N} x_t(\gamma)\varepsilon_t v_t.$$

By using the same argument as in Hansen (1996, p.426-427), we have

$$Z_n^*(\gamma) \Rightarrow \sigma G_\gamma \;\text{ a.s. as } n \to \infty. \tag{A.18}$$

Note that

$$\sup_{\gamma\in\Gamma} |\hat{Z}_n(\gamma) - Z_n^*(\gamma)| \le \sup_{\gamma\in\Gamma} \left| \frac{1}{n} \sum_{t=p}^{N} x_t(\gamma)x_t(\gamma)'v_t \right| \sup_{\gamma\in\Gamma} \left| \sqrt{n}(\lambda_n(\gamma) - \lambda_0) \right|.$$

Using the same argument as for (A.18) (see, e.g., Hansen (1996, p.427)), we have

$$\frac{1}{n}\sum_{t=p}^{N} x_t(\gamma)x_t(\gamma)'v_t \Rightarrow 0 \text{ a.s. as } n \to \infty. \tag{A.19}$$

Now, by Lemma A.6 and (A.19), it follows that under H_0 or H_{1n},

$$\hat{Z}_n(\gamma) - Z_n^*(\gamma) \Rightarrow 0 \ \text{ in probability as } n \to \infty. \tag{A.20}$$

Thus, by (A.18) and (A.20), we know that under H_0 or H_{1n},

$$\hat{Z}_n(\gamma) \Rightarrow \sigma G_\gamma \text{ in probability as } n \to \infty. \tag{A.21}$$

Next, we consider the functional

$$L : x(\cdot) \in D_{2p+2}(\Gamma) \to \frac{1}{\sigma^2}\sup_{\gamma\in\Gamma} x(\gamma)'\Omega_\gamma x(\gamma),$$

where $D_{2p+2}(\Gamma)$ denotes the function spaces of all functions, mapping $\mathcal{R}^2(\Gamma)$ into $\mathcal{R}^{2p+2}$, that are right continuous and have right-hand limits. Clearly, $L(\cdot)$ is a continuous functional; see e.g., Chan (1990, p.1891). By the continuous mapping theory and (A.21), it follows that under H_0 or H_{1n},

$$L(\hat{Z}_n(\gamma)) \Rightarrow L(\sigma G_\gamma) \text{ in probability as } n \to \infty. \tag{A.22}$$

Furthermore, since $\sigma_n^2 \to \sigma^2$ a.s. and $(X_{1n}(\gamma), I)'[X_{2n}(\gamma)]^{-1}(X_{1n}(\gamma), I) \to \Omega_\gamma$ uniformly in γ by Lemma A.3, we have that

$$\sup_{\gamma\in\Gamma} \hat{LR}_n(\gamma) = L(\hat{Z}_n(\gamma)) + o_p(1). \tag{A.23}$$

Finally, the conclusion follows from (A.22)-(A.23). This completes the proof. □

Proof of Corollary 3.1. Conditional on the sample $\{y_0, \cdots, y_N\}$, let $\hat{F}_{n,J}$ and $\hat{F}_n$ be the conditional empirical c.d.f. and c.d.f. of $\hat{LR}_n$, respectively. Then,

$$\begin{aligned} &P\left(LR_n \geq c_{n,\alpha}^J\right) \\ &= E\left[P\left(LR_n \geq c_{n,\alpha}^J | y_0, \cdots, y_N\right)\right] \\ &= E\left[P\left(\hat{F}_{n,J}(LR_n) \geq 1-\alpha | y_0, \cdots, y_N\right)\right]. \end{aligned}$$

By the Glivenko-Cantelli Theorem and Theorem 3.1, it follows that under H_0 or H_{1n},

$$\begin{aligned}
&\lim_{n\to\infty}\lim_{J\to\infty} P\left(LR_n \geq c_{n,\alpha}^J\right)\\
&= \lim_{n\to\infty} E\left[P\left(\hat{F}_n(LR_n) \geq 1-\alpha | y_0,\cdots,y_N\right)\right]\\
&= \lim_{n\to\infty} E\left[P\left(F_0(LR_n) \geq 1-\alpha | y_0,\cdots,y_N\right)\right]\\
&= \lim_{n\to\infty} P\left(F_0(LR_n) \geq 1-\alpha\right), \qquad \text{(A.24)}
\end{aligned}$$

where F_0 is the c.d.f. of $\sup_{\gamma\in\Gamma} G_\gamma'\Omega_\gamma G_\gamma$. Thus, by (A.24) and Theorem 2.1, under H_0 we have

$$\lim_{n\to\infty}\lim_{J\to\infty} P\left(LR_n \geq c_{n,\alpha}^J\right) = P\left(\sup_{\gamma\in\Gamma} G_\gamma'\Omega_\gamma G_\gamma \geq F_0^{-1}(1-\alpha)\right) = \alpha,$$

i.e., (i) holds. Furthermore, by (A.24) and Theorem 2.2, under H_{1n} we have

$$\lim_{h\to\infty}\lim_{n\to\infty}\lim_{J\to\infty} P\left(LR_n \geq c_{n,\alpha}^J\right) = \lim_{h\to\infty} P\left(B_h \geq F_0^{-1}(1-\alpha)\right) = 1,$$

where $B_h \triangleq \sup_{\gamma\in\Gamma}\left\{G_\gamma'\Omega_\gamma G_\gamma + h'\mu_{\gamma\gamma_0}h\right\}$ and the last equation holds since $B_h \to \infty$ in probability as $h\to\infty$. Thus, (ii) holds. This completes the proof. $\square$

REFERENCES

[1] Andrews, D.W.K. (1993) An introduction to econometric applications of functional limit theory for dependent random variables. *Econometric Reviews* **12**, 183-216.

[2] Chan, K.S. (1990) Testing for threshold autoregression. *Annals of Statistics* **18**, 1886-1894.

[3] Doukhan, P., Massart, P., and Rio, E. (1995) Invariance principles for absolutely regular empirical processes. *Annales de I'Institut H. Poincare* **31** 393-427.

[4] Fan, J. and Yao, Q. (2003) Nonlinear time series: Nonparametric and parametric methods. Springer, New York.

[5] Hansen, B.E. (1996) Inference when a nuisance parameter is not indentified under the null hypothesis. *Econometrica* **64**, 413-430.

[6] Pollard, D. (1984) Convergence of Stochastic Processes. New York: Springer.

[7] Pollard, D. (1990) Empirical processes: Theory and applications. NSF-CMBS Regional Conference Series in Probability and Statistics, Vol 2. Hayward, CA: Institute of Mathematical Statistics.

Chinese Academy of Sciences
Institute of Applied Mathematics
HaiDian District, Zhongguancun
Bei Jing, China
E-mail: zkxaa@ust.hk

Department of Statistics and Actuarial Science
University of Hong Kong
Pokfulam Road, Hong Kong
E-mail: plhyu@hku.hk
hrntlwk@hku.hk