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 \( \varepsilon_t \) 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

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

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. $|A| = (\text{tr}(A^tA))^{1/2}$ 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. $A^t$ is the transpose of matrix $A$. $o_p(1)$ $(O_p(1))$ denotes a sequence of random numbers converging to zero (bounded) in probability. $\rightarrow_d$ denotes convergence in distribution and $\Rightarrow$ denotes weak convergence. $I(\cdot)$ is an indicator function.
2. Likelihood ratio test. Let \( \phi = (\phi_0, \ldots, \phi_p)', \psi = (\psi_0, \ldots, \psi_p)', \lambda = (\phi', \psi')', \gamma = (r_L, r_U) \), and \( x_t = (1, y_{t-1}, \ldots, y_{t-p})' \). Then, model (1.1)-(1.2) becomes

\[
y_t = x_t(\gamma)'\lambda + \epsilon_t,
\]

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 \( H_p \) lie inside the unit circle, and both \( p \) and \( d \) are known. Without loss of generality, we further assume that \( d \leq p \) if \( p \geq 1 \), 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 \( d > p \geq 1 \).

Suppose that \( \{y_0, \ldots, y_N\} \) are \( N + 1 \) consecutive observations from model (2.1) with the true parameters \( \lambda_0 \) and \( \gamma_0 \), where \( \lambda_0 = (\phi_0', \psi_0')', \phi_0 = (\phi_00, \ldots, \phi_0p)', \psi_0 = (\psi_00, \ldots, \psi_0p)' \), and \( \gamma_0 = (r_{L0}, r_{U0}) \). We consider the following hypotheses:

\[
\begin{aligned}
H_0 : \psi_0 &= 0, \\
H_1 : \psi_0 &\neq 0 \text{ for some } \gamma.
\end{aligned}
\]

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 \( \epsilon_t \sim N(0, 1) \) is a sequence of i.i.d. random variables. When \( r_L \neq r_U \), since

\[
R_t(\gamma) = I(y_{t-d} \leq r_L)
\]

\[
+ \sum_{j=1}^{\infty} I(y_{t-j-d} \leq r_L) \prod_{i=1}^{j} I(r_L < y_{t-i+1-d} \leq r_U) \quad \text{a.s.},
\]

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

\[
X = (x_p, x_{p+1}, \cdots, x_N)',
\]

\[
X_\gamma = (h_p(\gamma), h_{p+1}(\gamma), \cdots, h_N(\gamma))'.
\]

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

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

where \( \sigma_n^2 \) is the maximum likelihood estimate of \( \sigma^2 \) under the null hypothesis.
where

\begin{align}
\sigma_n^2 &= n^{-1} \{Y'Y - (Y'X)(X'X)^{-1}(X'Y)\}, \\
(2.4) \\
\sigma_n^2(\gamma) &= n^{-1} \{Y'Y - (Y'Z_\gamma)(Z_\gamma'Z_\gamma)^{-1}(Z_\gamma'Y)\}. \tag{2.5}
\end{align}

Since the exact value of \( \gamma \) 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 \( \gamma \); see Davis (1977, 1987). Thus, our LR test statistic is defined as

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

where \( \Gamma = \{ (r_L, r_U); a \leq r_L \leq r_U \leq b \} \) and \([a, b]\) is a predetermined interval. Here, we truncate the full range of \( \gamma \), 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 \( 2A_0rv/(r-v) < A \).

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 \( \varepsilon_t \) 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
\[
\sup_{\gamma \in \Gamma} \left\{ \frac{X'_\gamma X_{\gamma}}{n} - \frac{X'_\gamma X}{n} \left( \frac{X'X}{n} \right)^{-1} \frac{X'X_{\gamma}}{n} \right\}^{-1} - (\Sigma_{\gamma} - \Sigma \Sigma^{-1} \Sigma_{\gamma})^{-1} = o_p(1);
\]
(ii) furthermore, under $H_0$ it follows that
\[
\sup_{\gamma \in \Gamma} \left| T_{\gamma} - \left( - \Sigma \Sigma^{-1}, I \right) \frac{1}{\sqrt{n}} Z'_{\gamma} \varepsilon \right| = o_p(1),
\]
where $\varepsilon = (\varepsilon_p, \ldots, \varepsilon_N)'$, $T_{\gamma} = n^{-1/2} \left\{ X'_{\gamma} - X'_{\gamma} X (X'X)^{-1} X' \right\} Y$, $\Sigma = E(x_t x_t')$, 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 $n \rightarrow \infty$, where $G_{\gamma}$ 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
\[
LR_n \rightarrow_d \sup_{\gamma \in \Gamma} G'_{\gamma} \Omega_{\gamma} G_{\gamma}
\]
as $n \rightarrow \infty$, where $\Omega_{\gamma} = (-\Sigma_{\gamma} \Sigma^{-1}, I)' (\Sigma_{\gamma} - \Sigma \Sigma^{-1} \Sigma_{\gamma})^{-1} (-\Sigma_{\gamma} \Sigma^{-1}, I)$.
Proof. By (2.4)-(2.5) and a direct calculation, we have

\[
\begin{align*}
n \left[ \sigma_n^2 - \sigma_n^2(\gamma) \right] & = T'_n \left\{ \frac{X'_n X_n}{n} - \frac{X'_n X_n}{n} \left( \frac{X'X_n}{n} \right)^{-1} \frac{X'X_n}{n} \right\}^{-1} T_n.
\end{align*}
\] (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 \( \chi^2 \) distribution, namely, for fixed \( \gamma \), 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 \( \Gamma \) 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 \( \text{Var}(\varepsilon_t) = \sigma^2 \). The case that the threshold effect happens in the variance of \( \varepsilon_t \) needs a further study in the future.

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

\[ H_{1n} : \psi_0 = \frac{h}{\sqrt{n}} \text{ for a constant vector } h \in \mathbb{R}^{p+1}. \]

Theorem 2.2. If Assumptions 2.1-2.3 hold, then under \( H_{1n} \) it follows that

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

as \( n \rightarrow \infty \), where \( M_{\gamma_0} = E[x_t x_t' R_t(\gamma) R_t(\gamma_0)] \) and

\[ \mu_{\gamma_0} = \frac{1}{\sigma^2} \left( M_{\gamma_0} - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma \right)' \left( \Sigma_\gamma - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma \right)^{-1} \left( M_{\gamma_0} - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma \right). \]
Proof. Note that \( Y = X \phi_0 + X_{\gamma_0} h / \sqrt{n} + \varepsilon \) under \( H_{1n} \). Thus,

\[
T_{\gamma} = \frac{1}{\sqrt{n}} \left\{ X_\gamma' - X_\gamma' X (X'X)^{-1} X' \right\} \varepsilon + \frac{1}{n} \left\{ X_\gamma' - X_\gamma' X (X'X)^{-1} X' \right\} X_{\gamma_0} h
\]

\[
= \frac{1}{\sqrt{n}} \left( - (X_\gamma' X)(X'X)^{-1}, I \right) Z'_\varepsilon + \frac{1}{n} \left\{ X_\gamma' - X_\gamma' X (X'X)^{-1} X' \right\} X_{\gamma_0} h.
\]

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

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

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

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

Let \( \tilde{L}R_n \) 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 \( \tilde{L}R_n \) 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

\[
\tilde{L}R_n \rightarrow_d \sup_{\gamma \in \Gamma} G'_\gamma \Omega_\gamma G_\gamma \quad \text{as} \quad n \rightarrow \infty;
\]

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

\[
\tilde{L}R_n \rightarrow_d \sup_{\gamma \in \Gamma} \left\{ G'_\gamma \Omega_\gamma G_\gamma + h' \mu_{\gamma_0} h \right\} \quad \text{as} \quad n \rightarrow \infty.
\]

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

\( \hat{\varepsilon}_t = y_t - \mathbf{x}_t(\gamma)'\lambda_n(\gamma) \)  

with

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

where \( \Lambda \) is a compact parametric space of \( \lambda \), and \( \varepsilon_t(\lambda, \gamma) = y_t - \mathbf{x}_t(\gamma)'\lambda \). Next, we set

\( \hat{\text{LR}}_n(\gamma) = \frac{\hat{\mathbf{Z}}'_n(\gamma)(\mathbf{X}_1n(\gamma), I)[\mathbf{X}_2n(\gamma)]^{-1}(\mathbf{X}_1n(\gamma), I)\hat{\mathbf{Z}}_n(\gamma)}{\sigma_n^2} \)

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

\[
\hat{\mathbf{Z}}_n(\gamma) = \frac{1}{\sqrt{n}} \mathbf{Z}_\gamma' \hat{\varepsilon}, \quad \mathbf{X}_1n(\gamma) = -\frac{\mathbf{X}_\gamma' \mathbf{X}}{n} \left( \frac{\mathbf{X}' \mathbf{X}}{n} \right)^{-1},
\]

and \( \mathbf{X}_2n(\gamma) = \frac{\mathbf{X}_\gamma' \mathbf{X}}{n} - \frac{\mathbf{X}_\gamma' \mathbf{X}}{n} \left( \frac{\mathbf{X}' \mathbf{X}}{n} \right)^{-1} \frac{\mathbf{X}' \mathbf{X}_\gamma}{n} \).

Define

\( \hat{\text{LR}}_n \equiv \sup_{\gamma \in \Gamma} \hat{\text{LR}}_n(\gamma) \).

The asymptotic theory of \( \hat{\text{LR}}_n \) 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

\[
\hat{\text{LR}}_n | y_0, \cdots, y_N \rightarrow_d \sup_{\gamma \in \Gamma} G'_\gamma \Omega_\gamma G_\gamma \quad \text{in probability as } n \rightarrow \infty.
\]

**Proof.** See the Appendix in Zhu, Yu, and Li (2013). \( \square \)

**Remark 3.1.** In practice, \( \hat{\text{LR}}_n \) 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 \( \hat{\text{LR}}_n \).
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 \( \alpha \), 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 \( \alpha \)-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 \( \alpha \), if \( \hat{LR}_n \geq c_{n,\alpha}^J \), 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

\[
\lim_{n \to \infty} \lim_{J \to \infty} P(\hat{LR}_n \geq c_{n,\alpha}^J) = \alpha;
\]

(ii) under \( H_1 \) it follows that

\[
\lim_{h \to \infty} \lim_{n \to \infty} \lim_{J \to \infty} P(\hat{LR}_n \geq c_{n,\alpha}^J) = 1.
\]

**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 \( H_1 \). This method is also feasible to obtain the critical value for the LR test in Chan (1990) by setting \( \gamma_L \equiv \gamma_V \). 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:

\[
y_t = y_{t-1} - 0.09y_{t-2} + (\psi_1y_{t-1} + \psi_2y_{t-2})R_t(\gamma) + \varepsilon_t,
\]

where \( R_t(\gamma) \) is defined as in (1.2) with \( d = 1 \), \( \varepsilon_t \) 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>
<thead>
<tr>
<th>Table 1</th>
<th>Rejection rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\psi_1)</td>
<td>(\psi_2)</td>
</tr>
<tr>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.1</td>
<td>-0.09</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0.3</td>
<td>-0.27</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>-0.45</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0.7</td>
<td>-0.63</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1 lists the rejection rates of \(LR_n\) and \(LR^*_n\) with different values of \(\psi\) and \(\gamma\). 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:

![Figure 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.](image)

**Table 2**

Results of tests applied to data set $\{y_t\}$

<table>
<thead>
<tr>
<th>$p$</th>
<th>$d$</th>
<th>$LR_n$</th>
<th>$c_{0.1}$</th>
<th>$c_{0.05}$</th>
<th>$c_{0.01}$</th>
<th>$LR^*_n$</th>
<th>$c^*_{0.1}$</th>
<th>$c^*_{0.05}$</th>
<th>$c^*_{0.01}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4.29</td>
<td>13.66</td>
<td>16.51</td>
<td>23.29</td>
<td>4.29</td>
<td>9.69</td>
<td>11.79</td>
<td>18.58</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>9.08</td>
<td>17.97</td>
<td>22.07</td>
<td>30.76</td>
<td>5.83</td>
<td>14.57</td>
<td>17.75</td>
<td>24.92</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>21.08</td>
<td>18.53</td>
<td>21.36</td>
<td>29.58</td>
<td>13.69</td>
<td>12.47</td>
<td>14.52</td>
<td>18.82</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>7.18</td>
<td>20.88</td>
<td>23.93</td>
<td>31.63</td>
<td>6.46</td>
<td>15.60</td>
<td>19.10</td>
<td>26.02</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>14.38</td>
<td>20.07</td>
<td>23.67</td>
<td>32.50</td>
<td>8.16</td>
<td>17.02</td>
<td>20.83</td>
<td>30.15</td>
</tr>
</tbody>
</table>

$\dagger$ The value of $a$ and $b$ are set to be the 10th and 90th quantiles of $\{y_t\}$.

$\dagger$ The p-values for $LR_n$ and $LR^*_n$ are 0.053 and 0.064, respectively.

$\S$ $c_\alpha$ (or $c^*_\alpha$) is obtained by the bootstrap method in Section 3 with $J = 1000$. 

$\|$ $c_{0.01}$ is the bootstrap critical value at the 0.01 significance level.
\[
y_t = \begin{cases} 
   1.2211 + 0.1597y_{t-1} + 0.4017y_{t-2} + \varepsilon_t & \text{if } R_t = 1 \\
   (0.1979) \quad (0.1236) \quad (0.1656) \\
   0.0704 + 0.3754y_{t-1} + 0.3031y_{t-2} + \varepsilon_t & \text{if } R_t = 0 \\
   (0.1245) \quad (0.0856) \quad (0.0954) 
\end{cases}
\]  
\text{(4.2)}

where

\[
R_t = \begin{cases} 
   1 & \text{if } y_{t-2} \leq -0.617 \\
   0 & \text{if } y_{t-2} > 1.237 \\
   R_{t-1} & \text{otherwise}
\end{cases}
\]

and

\[
y_t = \begin{cases} 
   -0.4515 + 0.3924y_{t-1} - 0.8379y_{t-2} + \varepsilon_t & \text{if } R_t = 1 \\
   (0.2620) \quad (0.1400) \quad (0.2628) \\
   0.3971 + 0.3241y_{t-1} + 0.1822y_{t-2} + \varepsilon_t & \text{if } R_t = 0 \\
   (0.1503) \quad (0.0845) \quad (0.1129) 
\end{cases}
\]  
\text{(4.3)}

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 \( \sigma^2 \) 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 $\{\hat{\varepsilon}_2^2\}$; and (d) the partial autocorrelations for $\{\hat{\varepsilon}_1^2\}$.

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:

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

where model (4.4) is estimated by the least squares method with the standard errors in parentheses, and the estimated value of $\sigma^2_\varepsilon$ 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 $\{\hat{\varepsilon}_1^2\}$. 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

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


**Chinese Academy of Sciences**
**Institute of Applied Mathematics**
Haidian District, Zhongguancun
Beijing, 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 hrntlk@hku.hk
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, $\xi_i$ is strictly stationary, ergodic and $\alpha$-mixing with mixing coefficients $\alpha(m) = O(m^{-A})$. Next, take $\iota \in \left( \lfloor 2A_0rv/(r - v) + 1 \rfloor/(A + 1), 1 \right)$, and let $p = \lfloor j^* \rfloor$ and $s = \lfloor j/j^* \rfloor$, 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}^{\infty}.

Furthermore, let $\mathcal{F}_m = \sigma(\xi_i, m \leq i \leq n)$. Then, $\xi_{kp+1} \in \mathcal{F}_m$. 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 $j \geq j_0$,

$$E \left[ \prod_{i=1}^{j} \xi_i \right] \leq E \left[ \prod_{k=0}^{s-1} \xi_{kp+1} \right]$$

$$= \left\{ E \left[ \prod_{k=0}^{s-1} \xi_{kp+1} \right] - \prod_{k=0}^{s-1} E \left[ \xi_{kp+1} \right] \right\} + \prod_{k=0}^{s-1} E \left[ \xi_{kp+1} \right]$$

$$\leq 16(s-1)\alpha(p) + \rho^s$$

$$\leq C\lfloor j/j^* \rfloor \lfloor j^* \rfloor^{-A} + \rho^{\lfloor j/j^* \rfloor}.$$
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

\[
\sum_{j=1}^{\infty} \left\{ E \left[ \prod_{i=1}^{j} \xi_i \right] \right\}^{(r-v)/2A_0rv} \leq (j_0 - 1) + C \sum_{j=j_0}^{\infty} \left[ \left\lfloor j/j' \right\rfloor \left| j' - 1 \right\rfloor^{-A} \right]^{(r-v)/2A_0rv} \]

(A.1)

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

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

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

\[
\sum_{j=j_0}^{\infty} \rho^{\left\lfloor j/j' \right\rfloor (r-v)/2A_0rv} < \sum_{j=1}^{\infty} \rho^{\left\lfloor j/j' \right\rfloor (r-v)/2A_0rv} < \infty.
\]

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

\[
\|R_t(\gamma_1) - R_t(\gamma_2)\|_{2r/(r-v)} \leq C|\gamma_1 - \gamma_2|^{r-v}/2B_0rv.
\]
Thus, by iteration we can show that

\[ R_t(\gamma_1) - R_t(\gamma_2) = \Delta_t(\gamma_1, \gamma_2) + \sum_{j=1}^{\infty} \Delta_{t-j}(\gamma_1, \gamma_2) \prod_{i=1}^{j} I(r_{1L} < y_{t-i-\delta} \leq r_{1U}). \]  

(A.4)

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

\[ I(r_{1L} < y_{t-j-d} \leq r_{1U}) - I(r_{2L} < y_{t-j-d} \leq r_{2U}) = I(r_{2U} < y_{t-j-d} \leq r_{1U}) - I(r_{2L} < y_{t-j-d} \leq r_{1L}). \]

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

\[ E|\Delta_{t-j}(\gamma_1, \gamma_2)|^s \leq 2^{s-1} E|\Delta_{t-j}(\gamma_1, \gamma_2)| \]

\[ \leq 2^{s-1} \left[ 2 \sup_x f(x)|r_{1L} - r_{2L}| + \sup_x f(x)|r_{1U} - r_{2U}| \right] \]

(A.5)

\[ \leq C|\gamma_1 - \gamma_2|. \]

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

\[ E \left| \Delta_{t-j}(\gamma_1, \gamma_2) \prod_{i=1}^{j} I(r_{1L} < y_{t-i-\delta} \leq r_{1U}) \right|^{2rv/(r-v)} \]

\[ \leq \left\{ E[\Delta_{t-j}(\gamma_1, \gamma_2)]^{2B_0rv/(r-v)} \right\}^{1/B_0} \]

\[ \times \left\{ E \left[ \prod_{i=1}^{j} I(r_{1L} < y_{t-i-\delta} \leq r_{1U}) \right] \right\}^{1/A_0} \]

\[ \leq 2^{(2B_0rv/(r-v)-1)} \left\{ E[\Delta_{t-j}(\gamma_1, \gamma_2)] \right\}^{1/B_0} \]

\[ \times \left\{ E \left[ \prod_{i=1}^{j} I(r_{1L} < y_{t-i-\delta} \leq r_{1U}) \right] \right\}^{1/A_0} \]

(A.6)

\[ \leq C|\gamma_1 - \gamma_2|^{1/B_0} \left\{ E \left[ \prod_{i=1}^{j} I(r_{1L} < y_{t-i-\delta} \leq r_{1U}) \right] \right\}^{1/A_0}. \]

By (A.4)-(A.6), Minkowski’s inequality, Lemma A.1 and the compactness of \( \Gamma \), we
have
\[
\|R_t(\gamma_1) - R_t(\gamma_2)\|_{2v/(r-v)} \leq C|\gamma_1 - \gamma_2|^{(r-v)/2rv} + C|\gamma_1 - \gamma_2|^{(r-v)/2B_0rv} \\
\times \sum_{j=1}^{\infty} \left\{ E \left[ \prod_{i=1}^{j} I(r_{1L} < y_{t-i-d} \leq r_{1U}) \right] \right\}^{(r-v)/2B_0rv} \leq C|\gamma_1 - \gamma_2|^{(r-v)/2B_0rv}.
\]
This completes the proof.

\[ \]

**Lemma A.3.** Suppose that the conditions in Lemma A.2 hold and \( Ey_t^4 < \infty \). Then,
\[
\sup_{\gamma \in \Gamma} \left| \frac{X'_{\gamma}X}{n} - \Sigma_\gamma \right| \to 0 \text{ a.s. as } n \to \infty.
\]

**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 fixed \( \varepsilon > 0 \), we partition \( \Gamma \) by \( \{B_1, \ldots, B_{K_{\varepsilon}}\} \), where \( B_k = \{(r_L, r_U); \omega_k < r_L \leq \omega_{k+1}, \nu_k < r_U \leq \nu_{k+1}\} \cap \Gamma \). Here, \( \{\omega_k\} \) and \( \{\nu_k\} \) are chosen such that
\[
(\omega_{k+1} - \omega_k)^{(r-v)/2B_0rv} < C_1 \varepsilon \text{ and } (\nu_{k+1} - \nu_k)^{(r-v)/2B_0rv} < C_1 \varepsilon,
\]
where \( B_0 > 1 \) is specified as in Lemma A.2, and \( C_1 > 0 \) will be selected later.

Next, we set
\[
f_t^u(\varepsilon) = y_{t-p}^2 R_t(\omega_{k+1}, \nu_{k+1}) \text{ and } f_t^l(\varepsilon) = y_{t-p}^2 R_t(\omega_k, \nu_k).
\]
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 \( f_t^l(\varepsilon) \leq \phi_t(\gamma) \leq f_t^u(\varepsilon) \).

Furthermore, since \( rv/(2rv - r + v) < 1 \), we have
\[
\|y_{t-p}^2\|_{2v/(2rv-r+v)} < \|y_{t-p}^2\|_2 < \infty.
\]
Thus, by Hölder’s inequality, Lemma A.2 and (A.7), we have

\[
E \left[ f_n^u(\varepsilon) - f_l^t(\varepsilon) \right] \\
\leq \left\| y_{t-r}^2 \right\|_{2r^v/(2r^v - r - v)} \| R_t(\omega_{k+1}, \nu_{k+1}) - R_t(\omega_k, \nu_k) \|_{2r^v/(r - v)} \\
\leq C \left[ (\omega_{k+1} - \omega_k)^{(r-v)/2B_0r^v} + (\nu_{k+1} - \nu_k)^{(r-v)/2B_0r^v} \right] \\
\leq 2CC_1\varepsilon.
\]

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

**Proof of Lemma 2.1.** First, since \( K_{\gamma\gamma} \) is positive definite by Assumption 2.1, we know that both \( \Sigma \) and \( \Sigma_{\gamma} \) 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 \( \gamma \in \Gamma \), \( \Sigma_{\gamma} - \Sigma_{\gamma}^{-1}\Sigma'_{\gamma} \) is positive definite. Second, by the ergodic theorem, it is easy to see that

\[
(X'X)/n \rightarrow \Sigma \quad \text{a.s. as } n \rightarrow \infty.
\]

Third, by Lemma A.3 we have

\[
\sup_{\gamma \in \Gamma} \left| \frac{X'X}{n} - \Sigma_{\gamma} \right| \rightarrow 0 \quad \text{and} \quad \sup_{\gamma \in \Gamma} \left| \frac{X'X_{\gamma}}{n} - \Sigma_{\gamma} \right| \rightarrow 0 \quad \text{a.s.}
\]

as \( n \rightarrow \infty \). Note that if \( H_0 \) holds, we have

\[
T_{\gamma} = \frac{1}{\sqrt{n}} \left\{ X'_{\gamma} - X'_{\gamma}X(X'X)^{-1}X' \right\} \varepsilon \\
= \frac{1}{\sqrt{n}} \left( - (X'_{\gamma}X)(X'X)^{-1}, I \right)Z'_{\gamma}\varepsilon.
\]

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

**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
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 \( L^2 \) 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
\[
\|x_t(\gamma_1)\varepsilon_t - x_t(\gamma_2)\varepsilon_t\|_{2v}^2 
\leq \|x_t\|_{2}^2 \|R_t(\gamma_1) - R_t(\gamma_2)\|_{2rv/(r-v)}^2 
\leq C\|x_t\|_{4r}^4 \|\varepsilon_t\|_{4r}^4 |\gamma_1 - \gamma_2|^{(r-v)/2B_0rv}
\]
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.

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
\[(A.10) \quad \sup_{\gamma \in \Gamma} \left| a_n \sum_{t=p}^{n_0-1} x_t x'_t R_t(\gamma) \right| = O_p(1) \]
and
\[(A.11) \quad \sup_{\gamma \in \Gamma} \left| a_n \sum_{t=p}^{n_0-1} h_t(\gamma)\varepsilon_t \right| = O_p(1). \]

**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

$$P(n_0 > M) = P\left(\sum_{t=p}^{M} I(a \leq y_{t-d} \leq b) = 0\right)$$

$$= P\left(\frac{1}{M} \sum_{t=p}^{M} I(a \leq y_{t-d} \leq b) = 0\right)$$

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

$$< \eta,$$

(A.12)

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

$$P\left(\sup_{\gamma \in \Gamma} \left| a_n \sum_{t=p}^{n_0-1} x_t x_t' R_t(\gamma) \right| > \tilde{M}\right)$$

$$= P\left(\sup_{\gamma \in \Gamma} \left| a_n \sum_{t=p}^{n_0-1} x_t x_t' R_t(\gamma) \right| > \tilde{M}, n_0 \leq M\right)$$

$$\leq P\left(\max_{0 \leq k \leq M} \sup_{\gamma \in \Gamma} \left| a_n \sum_{t=p}^{k-1} x_t x_t' R_t(\gamma) \right| > \tilde{M}\right)$$

$$\leq \sum_{k=p}^{M} P\left(a_n \sum_{t=p}^{k-1} |x_t|^2 > \tilde{M}\right)$$

$$\leq a_n \sum_{k=p}^{M} \sum_{t=p}^{k-1} \frac{E|x_t|^2}{M}$$

(A.13)

$$= O\left(\frac{a_n M^2}{M}\right) = O\left(a_n\right) < \eta$$

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 $\forall \eta > 0$,

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

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

\[\square\]
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 $X_\gamma$ and $T_\gamma$, respectively, with $R_t(\gamma)$ being replaced by $\tilde{R}_t(\gamma)$.

Proof. (i) Note that

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

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

\[ T_\gamma - \tilde{T}_\gamma = \frac{1}{\sqrt{n}} \left( X_\gamma - \tilde{X}_\gamma \right)' \varepsilon - \frac{1}{\sqrt{n}} \left( X_\gamma - \tilde{X}_\gamma \right)' X (X'X)^{-1} X' \varepsilon \]

\[ - \frac{1}{n} \left( X_\gamma - \tilde{X}_\gamma \right)' X (X'X)^{-1} X' X_\gamma h \]

\[ - \frac{1}{n} \tilde{X}'_\gamma X (X'X)^{-1} X' \left( X_\gamma - \tilde{X}_\gamma \right) h \]

\[ + \frac{1}{n} \left( X'_\gamma X_\gamma - \tilde{X}'_\gamma \tilde{X}_\gamma \right) h \]

\[ \triangleq I_{1n}(\gamma) - I_{2n}(\gamma) - I_{3n}(\gamma) - I_{4n}(\gamma) + I_{5n}(\gamma) \text{ say.} \]

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

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

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., $h \equiv 0$) or $H_{1n}$, we know that (iii) holds. This completes the proof. \qed
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 
\[ \sum_{t=p}^{N} \left[ \varepsilon_t^2(\lambda_n(\gamma), \gamma) - \varepsilon_t^2(\lambda_0, \gamma) \right] = \cdots \]
\[ (A.14) \]
\[ + (\lambda_n(\gamma) - \lambda_0)' \left( \sum_{t=p}^{N} x_t(\gamma) x_t(\gamma)' \right) (\lambda_n(\gamma) - \lambda_0). \]

Next, when $\lambda_0 = (\phi'_0, h'/\sqrt{n})'$, we can show that 
\[ \frac{1}{\sqrt{n}} \sum_{t=p}^{N} \varepsilon_t(\lambda_0, \gamma) x_t(\gamma) = \frac{1}{\sqrt{n}} Z_n' \varepsilon + \frac{1}{\sqrt{n}} \sum_{t=p}^{N} x_t(\gamma) x_t(\gamma)' \lambda_0 \]
\[ = \frac{1}{\sqrt{n}} Z_n' \varepsilon + \frac{1}{n} \sum_{t=p}^{N} \left( x_t x_t' [R_t(\gamma_0) - R_t(\gamma)] \right) h \]
\[ (A.15) \]
\[ = G_n^*(\gamma). \]

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 
\[ P \left( \sup_{\gamma \in \Gamma} \sqrt{n}|\lambda_n(\gamma) - \lambda_0| > M \right) \]
\[ = P \left( \sqrt{n}|\lambda_n(\gamma) - \lambda_0| > M, \sum_{t=p}^{N} \left[ \varepsilon_t^2(\lambda_n(\gamma), \gamma) - \varepsilon_t^2(\lambda_0, \gamma) \right] \leq 0 \right) \]
for some $\gamma \in \Gamma$ 
\[ \leq P \left( \sqrt{n}|\lambda_n(\gamma) - \lambda_0| > M, \cdots \right) \]
\[ \leq P \left( M < \sqrt{n}|\lambda_n(\gamma) - \lambda_0| \leq 2[\lambda_{\min}(\gamma) + o_p(1)]^{-1}|G_n^*(\gamma)| \right) \]
\[ \leq \eta, \]
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. \hfill \Box

**Proof of Corollary 2.1.** The conclusion follows directly from Theorems 2.1-2.2 and Lemma A.5. \hfill \Box

**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 $\omega$ for which

\begin{equation}
\lim_{n \to \infty} \frac{1}{n} \sum_{t=p}^{N} |x_t(\gamma)| \varepsilon_t^2 < \infty \text{ a.s.,} \quad (A.16)
\end{equation}

\begin{equation}
\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.,} (A.17)
\end{equation}

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

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{t=p}^{N} |x_t(\gamma)| \varepsilon_t^2 \leq \sqrt{2} \sum_{t=p}^{N} |x_t| \varepsilon_t^2 < \infty \text{ a.s.,} \]

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 $\omega$, and hence all of the randomness appears in the i.i.d. $N(0, 1)$ variables \{vt\}.

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

\begin{equation}
Z_n^*(\gamma) \Rightarrow \sigma G_{\gamma} \text{ a.s. as } n \to \infty. \quad (A.18)
\end{equation}

Note that

\[ \sup_{\gamma \in \Gamma} |\hat{Z}_n(\gamma) - Z_n^*(\gamma)| \leq \sup_{\gamma \in \Gamma} \left| \frac{1}{n} \sum_{t=p}^{N} x_t(\gamma)x_t(\gamma)' v_t \right| \sup_{\gamma \in \Gamma} \left| \sqrt{n}(\lambda_n(\gamma) - \lambda_0) \right|. \]
Using the same argument as for (A.18) (see, e.g., Hansen (1996, p.427)), we have

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

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

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

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

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

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 $H_{1n}$,

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

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 $\gamma$ by Lemma A.3, we have that

\[(A.23) \quad \sup_{\gamma \in \Gamma} \hat{L}R_n(\gamma) = L(\hat{Z}_n(\gamma)) + o_p(1).\]

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

**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 $\hat{L}R_n$, respectively. Then,

\[
P \left( LR_n \geq c_{n,\alpha}' \right) \\
= E \left[ P \left( LR_n \geq c_{n,\alpha}' | y_0, \cdots, y_N \right) \right] \\
= E \left[ P \left( \hat{F}_{n,J}(LR_n) \geq 1 - \alpha | y_0, \cdots, y_N \right) \right].\]
By the Glivenko-Cantelli Theorem and Theorem 3.1, it follows that under $H_0$ or $H_{1n}$,

$$
\begin{align*}
\lim_{n \to \infty} \lim_{J \to \infty} P \left( LR_n \geq c_{n,\alpha}^{J} \right) &= \lim_{n \to \infty} E \left[ P \left( \hat{F}_n(LR_n) \geq 1 - \alpha \mid y_0, \cdots, y_N \right) \right] \\
&= \lim_{n \to \infty} E \left[ P \left( F_0(LR_n) \geq 1 - \alpha \mid y_0, \cdots, y_N \right) \right] \\
&= \lim_{n \to \infty} P \left( F_0(LR_n) \geq 1 - \alpha \right),
\end{align*}
$$

(A.24)

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

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

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

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

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

REFERENCES


