UNIT ROOT TESTING ON BUFFERED AUTOREGRESSIVE MODEL

DI WANG AND WAI KEUNG LI

University of Hong Kong

Supplementary Material

S1 Auxiliary Lemmas

The proofs of the theorems rely on some auxiliary lemmas. Therefore, we state and prove these lemmas first.

Denote the value at time t of the time series with length T as X_{Tt}, and let $X(s)$ for $s \in [0,1]$ be a stochastic process. We say that the time series X_{Tt} converges weakly to the stochastic process $X(\cdot)$, denoted by $X_{Tt} \Rightarrow X(\cdot)$, if for any $s \in [0,1]$ the time series $X_{T[Ts]} \Rightarrow X(s)$ as $T \to \infty$ where $[x]$ denotes the largest integer smaller or equal to x.

Lemma A1. If the array $X_{Ts} \Rightarrow X(s)$ as $T \to \infty$ and $X(s)$ is continuous almost surely, then

$$\frac{1}{T}\sum_{t=1}^{T} X_{Tt}R_t(\gamma) \Rightarrow R(\gamma)\int_0^1 X(s)ds$$

where $R(\gamma) := \mathbb{E}[R_t(\gamma)]$ is the stationary probability of Z_t staying in regime 1.

Proof of Lemma A1

We generally follow the techniques in Caner and Hansen (2001). Let $d_t(\gamma) = R_t(\gamma) - R(\gamma)$

so that $E[d_t(\gamma)] = 0$.

$$\frac{1}{T}\sum_{t=1}^{T} X_{Tt}R_t(\gamma) = \frac{1}{T}\sum_{t=1}^{T} X_{Tt}d_t(\gamma) + R(\gamma)\frac{1}{T}\sum_{t=1}^{T} X_{Tt}$$

and

$$\frac{1}{T}\sum_{t=1}^{T} X_{Tt} \Rightarrow \int_0^1 X(s)ds$$

Hence it is sufficient to show that

$$\sup_{\gamma\in\Gamma}\left|\frac{1}{T}\sum_{t=1}^{T} X_{Tt}d_t(\gamma)\right| \xrightarrow{p} 0$$

For a fixed $\varepsilon > 0$, since $X(s)$ is continuous almost surely, there exists $\delta > 0$ such that

$$P\left(2\sup_{|s-s'|\le\delta}|X(s) - X(s')| \le \varepsilon\right) \ge 1 - \varepsilon$$

Set $N = [1/\delta]$, for $k = 0, \ldots, N-1$, set $t_k = [kT\delta] + 1$, then

$$E\sum_{k=0}^{N-1}\sup_{\gamma\in\Gamma}\left|\frac{1}{T}\sum_{t=t_k}^{t_{k+1}-1} d_t(\gamma)\right| = E\sup_{\gamma\in\Gamma}\left|\frac{1}{T\delta}\sum_{t=t_k}^{t_{k+1}-1} d_t(\gamma)\right| \to 0$$

by the uniform weak law of large numbers. Hence,

$$\begin{aligned}
\sup_{\gamma\in\Gamma}\left|\frac{1}{T}\sum_{t=1}^{T} X_{Tt}d_t(\gamma)\right| &= \sup_{\gamma\in\Gamma}\left|\frac{1}{T}\sum_{k=0}^{N-1}\sum_{t=t_k}^{t_{k+1}-1} X_{Tt}d_t(\gamma)\right| \\
&\le \frac{1}{T}\sum_{k=0}^{N-1}|X_{Tt_k}|\sup_{\gamma\in\Gamma}\left|\sum_{t=t_k}^{t_{k+1}-1} d_t(\gamma)\right| + \frac{1}{T}\sum_{k=0}^{N-1}\sum_{t=t_k}^{t_{k+1}-1}|X_{Tt} - X_{Tt_k}|\sup_{\gamma\in\Gamma}|d_t(\gamma)| \\
&\le \sup_{1\le t\le T}|X_{Tt}|\sum_{k=0}^{N-1}\sup_{\gamma\in\Gamma}\left|\frac{1}{T}\sum_{t=t_k}^{t_{k+1}-1} d_t(\gamma)\right| + 2\sup_{t-t'\le T\delta}|X_{Tt} - X_{Tt'}| \\
&\Rightarrow 2\sup_{|s-s'|\le\delta}|X(s) - X(s')| \\
&\le \varepsilon
\end{aligned}$$

Lemma A2. Define the partial sum process $Y_t(\gamma) = \sum_{i=1}^{t} e_i R_i(\gamma)$, and the scaled array with $s \in [0,1]$, $Y_T(s,\gamma) = \frac{1}{\sqrt{T}} Y_{[Ts]}(\gamma) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} e_t R_t(\gamma)$. Under Assumptions 1 and 2, as $T \to \infty$, we have

(A2.1.)

$$Y_T(s,\gamma) \Rightarrow \sigma W(s, R(\gamma))$$

(A2.2.)

$$\int_0^1 X_T(s) dY_T(s,\gamma) \Rightarrow \sigma \int_0^1 X(s) dW(s, R(\gamma))$$

Proof of Lemma A2

Note that $\{e_t R_t(\gamma)\}$ is a strictly stationary and ergodic Martingale Difference Sequence (MDS) with variance $E(e_t^2 R_t(\gamma)) = \sigma^2 R(\gamma)$. Thus, by the MDS Central Limit Theorem (Hansen, 2017), we have

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} e_t R_t(\gamma) \Rightarrow \sigma W(R(\gamma))$$

Thus, for the partial sum, we have

$$Y_T(s,\gamma) := \frac{1}{\sqrt{T}} \sum_{t=1}^{[Ts]} e_t R_t \Rightarrow \sigma W(s, R(\gamma))$$

Then, if we apply the weak convergence of stochastic integrals with respect to two-parameter process in Caner and Hansen (2001), we can directly obtain Lemma A2.2.

S2 Proof of Theorems

Proof of Lemmas 1 and 2

Note that we assume that $\sum_{i=1}^{p-1}|\phi_i|<1$ and $\sum_{i=1}^{p-1}|\psi_i|<1$, Li et al. (2015) showed that Δy_t is strictly stationary and ergodic. Then for Lemma 1.1 and Lemma 2.1, namely the stationary part, we can directly follow the convergence results for the stationary BAR (the proofs of Lemma 1 and Lemma 2 in Zhu, Yu and Li (2014)).

For Lemma 1.2, note that

$$\begin{aligned}&\frac{1}{T}\sum u_{t-1}\frac{y_{t-1}}{\sqrt{T}}R_t(\gamma)\\ =&\frac{1}{T}\sum u_{t-1}\frac{y_{t-2}}{\sqrt{T}}R_{t-1}(\gamma)+\frac{1}{T}\sum u_{t-1}\left[\frac{y_{t-1}}{\sqrt{T}}R_t(\gamma)-\frac{y_{t-2}}{\sqrt{T}}R_{t-1}(\gamma)\right]\end{aligned}$$

where

$$\frac{1}{T}\sum u_{t-1}\frac{y_{t-2}}{\sqrt{T}}R_{t-1}(\gamma)\xrightarrow{p}0$$

as u_{t-1} and y_{t-2} are independent with mean zero and

$$\begin{aligned}&\frac{1}{T}\sum u_t\left[\frac{y_t}{\sqrt{T}}R_t(\gamma)-\frac{y_{t-1}}{\sqrt{T}}R_{t-1}(\gamma)\right]\\ =&\frac{1}{T}\sum\frac{u_t^2}{\sqrt{T}}R_t(\gamma)+\frac{1}{T}\sum\frac{u_ty_{t-1}}{\sqrt{T}}[R_t(\gamma)-R_{t-1}(\gamma)]\xrightarrow{p}0.\end{aligned}$$

where the first term goes to zero obviously, and the second term goes to zero because of the independence of u_t and $y_{t-1}/\sqrt{T}$ and the boundedness of $R_t(\gamma)-R_{t-1}(\gamma)$. And the asymptotic orthogonality of u_{t-i} and y_{t-1} can be applied similarly, for $i=2,\ldots,p$.

Lemmas 1.3 and 2.3 can be implied directly by Lemma A1.

Proof of Theorem 1

Denote $M(\gamma)=X'(\gamma)X/T$ and note that

$$M(\gamma)=\begin{bmatrix}M_{11}(\gamma) & M_{21}(\gamma)'\\ M_{21}(\gamma) & M_{22}(\gamma)\end{bmatrix}=\frac{1}{T}\sum_{t=p}^{T}\begin{bmatrix}y_{t-1}^2 & y_{t-1}w_{t-1}'\\ w_{t-1}y_{t-1}' & w_{t-1}w_{t-1}'\end{bmatrix}R_t(\gamma)$$

where $w_{t-1} = (u_{t-1}, \ldots, u_{t-p+1})'$.

Note that $X'X$ is a special case of $X'(\gamma)X$ and denote $M = X'X/T$.

From Lemma 1.2, it can be obtained that $M(\gamma)$ is asymptotically block diagonal, then $LR_T(\gamma)$ is

$$
\begin{aligned}
LR_T(\gamma) &= S_1'(\gamma)[M_{11}(\gamma) - M_{11}(\gamma)M_{11}^{-1}M_{11}(\gamma)]^{-1}S_1(\gamma) \\
&\quad + S_2'(\gamma)[M_{22}(\gamma) - M_{22}(\gamma)M_{22}^{-1}M_{22}(\gamma)]^{-1}S_2(\gamma) + o_p(1) \\
&= \frac{S_1'(\gamma)}{\sqrt{T}} \left[\frac{M_{11}(\gamma)}{T} - \frac{M_{11}(\gamma)}{T} \frac{M_{11}}{T}^{-1} \frac{M_{11}(\gamma)}{T} \right]^{-1} \frac{S'}{\sqrt{T}} \\
&\quad + S_2'(\gamma)[M_{22}(\gamma) - M_{22}(\gamma)M_{22}^{-1}M_{22}(\gamma)]^{-1}S_2(\gamma) + o_p(1)
\end{aligned}
$$

where

$$
\begin{aligned}
\frac{1}{\sqrt{T}} S_1(\gamma) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \frac{y_{t-1}}{\sqrt{T}} e_t R_t(\gamma) - \frac{1}{\sqrt{T}} M_{11}(\gamma) M_{11}^{-1} \sum_{t=1}^{T} \frac{y_{t-1}}{\sqrt{T}} e_t \\
&\Rightarrow \int_0^1 W(s) dW(s, R(\gamma)) - R(\gamma) \int_0^1 W(s) dW(s),
\end{aligned}
$$

$$
\begin{aligned}
S_2(\gamma) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{T} w_t e_t R_t(\gamma) - \frac{1}{\sqrt{T}} M_{22}(\gamma) M_{22}^{-1} \sum_{t=1}^{T} w_t e_t \\
&\Rightarrow G(\gamma) - \Sigma_\gamma \Sigma^{-1} G,
\end{aligned}
$$

$$
\frac{M_{11}(\gamma)}{T} - \frac{M_{11}(\gamma)}{T} \frac{M_{11}^{-1}}{T} \frac{M_{11}(\gamma)}{T} \Rightarrow R(\gamma)(1 - R(\gamma)) \int_0^1 W^2(s) ds,
$$

and

$$
M_{22}(\gamma) - M_{22}(\gamma) M_{22}^{-1} M_{22}(\gamma) \to \Sigma_\gamma - \Sigma_\gamma \Sigma^{-1} \Sigma_\gamma
$$

according to Lemmas 1 and 2.

Proof of Proposition 1

Since the bootstrap sample and statistics fully depend on the realization of the original sample, we denote $\overset{p}{\Rightarrow}$ as weak convergence in probability (Giné and Zinn, 1990). First, the bootstrap residuals $\tilde{e}_t = \hat{e}_t v_t$ is a martingale difference sequence associated with $\mathcal{F}(t) = \{\hat{e}_1, \ldots, \hat{e}_t\}$, then by the MDS central limit theorem (Hansen, 2017) and Assumption 1, we have the invariance principle for the bootstrap residual $T^{-1/2}\sum_{t=1}^{[Ts]} \tilde{e} \overset{p}{\Rightarrow} \sigma W(s)$.

Following the proof of Lemma 1, by Lemma A.1, we can easily prove the following statements for $\tilde{y}_t$ and $\tilde{w}_t = (\tilde{u}_{t-1}, \ldots, \tilde{u}_{t-p+1})'$

$$n^{-1}\sum_{t=s}^{T} \tilde{w}_t \tilde{w}_t' \overset{a.s.}{\to} \hat{\Sigma} \text{ and } n^{-1}\sum_{t=s}^{T} \tilde{w}_t \tilde{w}_t' \tilde{R}_t(\gamma) \overset{a.s.}{\to} \hat{\Sigma}_\gamma$$

$$n^{-3/2}\sum_{t=s}^{T} \tilde{y}_{t-1}\tilde{u}_{t-1}\tilde{R}_t(\gamma) \overset{p}{\to} 0 \text{ and } n^{-2}\sum_{t=s}^{T} \tilde{y}_{t-1}^2 \tilde{R}_t(\gamma) \Rightarrow \hat{R}(\gamma)\int_0^1 W(s)^2 ds$$

where $\hat{\Sigma} = \mathbb{E}(\tilde{w}_t \tilde{w}_t')$, $\hat{\Sigma}_\gamma = \mathbb{E}(\tilde{w}_t \tilde{w}_t' \tilde{R}_t(\gamma))$ and $\hat{R}(\gamma) = \mathbb{E}(\tilde{R}_t(\gamma))$ depend on each realization and the corresponding estimator $\hat{\lambda}_0$. Similarly, we can also prove the bootstrap version of Lemma 2, and it is noteworthy that the corresponding limiting distributions depend on the estimator $\hat{\lambda}_0$.

With the unit root constraint in the bootstrap data generating process, the bootstrap LR estimator with a given γ is

$$\widetilde{LR}_n(\gamma) = \frac{\tilde{S}'(\gamma)\left[\tilde{X}'(\gamma)\tilde{X}(\gamma)/n - (\tilde{X}'(\gamma)\tilde{X}/n)(\tilde{X}'\tilde{X}/n)^{-1}(\tilde{X}'\tilde{X}(\gamma)/n)\right]^{-1}\tilde{S}(\gamma)}{\tilde{\sigma}^2(\gamma)/\sigma^2}$$

where

$$\tilde{S}(\gamma) = \frac{1}{\sqrt{n}}\left[\tilde{X}'(\gamma) - \tilde{X}'(\gamma)\tilde{X}(\tilde{X}'\tilde{X})^{-1}X'\right]\tilde{\varepsilon}.$$

Conditioning on each realization and estimator, the limiting distribution is $\hat{Q}_1(\gamma) + \hat{Q}_2(\gamma)$

where

$$\hat{Q}_1(\gamma) = J_1(\hat{R}(\gamma))' \left[\hat{R}(\gamma)(1-\hat{R}(\gamma)) \int_0^1 W^2(s)ds\right]^{-1} J_1(\hat{R}(\gamma)) \text{ and}$$

$$\hat{Q}_2(\gamma) = [\hat{G}(\gamma) - \hat{\Sigma}_\gamma \hat{\Sigma}^{-1}\hat{G}]'[\hat{\Sigma}_\gamma - \hat{\Sigma}_\gamma \hat{\Sigma}^{-1}\hat{\Sigma}_\gamma]^{-1}[\hat{G}(\gamma) - \hat{\Sigma}_\gamma \hat{\Sigma}^{-1}\hat{G}].$$

Finally, since under the null hypothesis $\hat{\lambda}_0$ is a consistent estimator and the limiting distribution $\hat{Q}_1(\gamma)$ and $\hat{Q}_2(\gamma)$ are continuous with respect to the estimator $\hat{\lambda}_0$. By the continuous mapping theorem, we can derive that the weak convergence of the bootstrap estimator in probability

$$\widetilde{LR}_n | y_1, \ldots, y_T \overset{p}{\Rightarrow} \sup_{\gamma \in \Gamma} LR(\gamma).$$

Next, we prove that the bootstrap size is correct.

$$\begin{aligned}
&P(LR_n \geq c^B_{n,\alpha}) \\
=&\mathbb{E}[P(LR_n \geq c^B_{n,\alpha} | y_0, \ldots, y_N)] \\
=&\mathbb{E}[P(\hat{F}_{n,B}(LR_n) \leq 1-\alpha | y_0, \ldots, y_N)]
\end{aligned}$$

By the Glivenko-Cantelli Theorem and Proposition 1, it follows that under H_0,

$$\begin{aligned}
&\lim_{n\to\infty} \lim_{B\to\infty} P(LR_n \geq c^B_{n,\alpha}) \\
=&\lim_{n\to\infty} \mathbb{E}[P(\hat{F}_n(LR_n) \geq 1-\alpha | y_0, \ldots, y_N)] \\
=&\lim_{n\to\infty} \mathbb{E}[P(F_0(LR_n) \geq 1-\alpha | y_0, \ldots, y_N)] \\
=&\lim_{n\to\infty} P(F_0(LR_n) \geq 1-\alpha) = \alpha
\end{aligned}$$

where F_0 is the c.d.f. of $\sup_{\gamma\in\Gamma} LR(\gamma)$.

Proof of Theorem 2

The OLS $\hat{\phi}$ in the BAR(p) is

$$\hat{\phi} = \left(\sum_{t=s}^{T} \begin{bmatrix} y_{t-1}^2 & y_{t-1}u_{t-1} & \dots & y_{t-1}u_{t-p+1} \\ u_{t-1}y_{t-1} & u_{t-1}^2 & \dots & u_{t-1}u_{t-p+1} \\ \dots & \dots & \dots & \dots \\ u_{t-p+1}y_{t-1} & u_{t-p+1}u_{t-1} & \dots & u_{t-p+1}^2 \end{bmatrix} R_t(\gamma) \right)^{-1} \left(\sum_{t=s}^{T} \begin{bmatrix} y_{t-1}e_t \\ u_{t-1}e_t \\ \dots \\ u_{t-p+1}e_t \end{bmatrix} \right)$$

Thus, by the convergence property in Lemmas 1 and 2, we have

$$\begin{bmatrix} n & 0 & \dots & 0 \\ 0 & \sqrt{n} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \sqrt{n} \end{bmatrix} \hat{\phi} \Rightarrow \begin{bmatrix} R(\gamma)\int_0^1 W(s)^2 ds & 0' \\ 0 & \Sigma_\gamma \end{bmatrix}^{-1} \begin{bmatrix} \sigma \int_0^1 W(s) dW(s, R(\gamma)) \\ G(\gamma) \end{bmatrix}$$

Therefore, we can obtain that

$$t_1(\gamma) \Rightarrow \frac{\int_0^1 W(s) dW(s, R(\gamma))}{\sqrt{R(\gamma)\int_0^1 W(s)^2 ds}}.$$

Similarly, we can obtain that

$$t_2(\gamma) \Rightarrow \frac{\int_0^1 W(s) d[W(s,1) - W(s, R(\gamma))]}{\sqrt{(1 - R(\gamma))\int_0^1 W(s)^2 ds}}.$$

Proof of Theorem 3

Under Assumptions 1, 2 and the null hypothesis: $\phi_0 = \psi_0 = 0$, we have

$$\begin{aligned} \Delta y_t &= x_t'\phi R_t(\gamma_0) + x_t'\psi(1 - R_t(\gamma_0)) + e_t \\ &= x_t'\phi R_t(\gamma) + x_t'\psi(1 - R_t(\gamma)) + x_t'(\psi - \phi)[R_t(\gamma_0) - R_t(\gamma)] + e_t \\ &= x_t'\phi R_t(\gamma) + x_t'\psi(1 - R_t(\gamma)) + w_t'(\psi_{-0} - \phi_{-0})[R_t(\gamma_0) - R_t(\gamma)] + e_t \end{aligned}$$

where $\phi_{-0} = (\phi_1, \ldots, \phi_{p-1})'$ and $\psi_{-0} = (\psi_1, \ldots, \psi_{p-1})'$.

Therefore, we have

$$t_1(\gamma) = \frac{N_T(\gamma) + A_T(\gamma)}{\sqrt{D_T(\gamma)\hat{\sigma}^2(\gamma)}}$$

where

$$N_T(\gamma) = \frac{1}{T}\sum_{t=1}^{T} y_{t-1}e_t R_t(\gamma) - \frac{1}{T}\sum_{t=1}^{T} y_{t-1}w_{t-1}R_t(\gamma)\left(\sum_{t=1}^{T} w_{t-1}^2 R_t(\gamma)\right)^{-1}\sum_{t=1}^{T} w_{t-1}e_t R_t(\gamma)$$

$$A_T(\gamma) = \frac{1}{T}\sum_{t=1}^{T} y_{t-1}R_t(\gamma)w_t'(\phi_{-0} - \psi_{-0})[R_t(\gamma_0) - R_t(\gamma)]$$

$$D_T(\gamma) = \frac{1}{T^2}\sum_{t=1}^{T} y_{t-1}^2 R_t(\gamma) - \frac{1}{T}\sum_{t=1}^{T} y_{t-1}w_{t-1}R_t(\gamma)\left(\sum_{t=1}^{T} w_{t-1}^2 R_t(\gamma)\right)^{-1}\sum_{t=1}^{T} w_{t-1}e_t R_t(\gamma)$$

We first show that $A_T(\hat{\gamma}) \xrightarrow{p} 0$. Note that

$$A_T(\gamma) \leq \frac{1}{\sqrt{T}}\max_{t\leq T}|y_{t-1}R_t(\gamma)|\frac{1}{\sqrt{T}}\sum_{t=1}^{T} w_t'(\phi_{-0} - \psi_{-0})[R_t(\gamma_0) - R_t(\gamma)]$$

In Li et al. (2015), the convergence of the estimated thresholds was shown to be $T(\hat{r}_L - r_{L0}) = O_p(1)$ and $T(\hat{r}_U - r_{U0}) = O_p(1)$. Also, in Zhu, Yu and Li (2014), under the α-mixing assumption, for $r > v > 1$ there exists a $B > 1$ such that for any $\gamma_1, \gamma_2 \in \Gamma$, we have

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

Therefore, we can prove $A_T(\hat{\gamma}) \xrightarrow{p} 0$, and further

$$t_1(\hat{\gamma}) = \frac{N_T(\hat{\gamma})}{\sqrt{D_T(\hat{\gamma})\hat{\sigma}^2(\hat{\gamma})}} + o_p(1).$$

Denote the long-term variance $\sum_{k=\infty}^{\infty}\mathbb{E}(\Delta y_t \Delta y_{t+k})$ as σ_y^2. Then, we have

$$\begin{pmatrix} \frac{1}{\sqrt{R(\gamma_0)T}}\sum_{t=1}^{[Ts]} e_t R_t(\gamma) \\ \frac{1}{\sqrt{(1-R(\gamma_0))T}}\sum_{t=1}^{[Ts]} e_t(1-R_t(\gamma)) \\ \frac{1}{\sigma_y\sqrt{T}}\sum_{t=1}^{[Ts]}\Delta y_t \end{pmatrix} \Rightarrow \begin{pmatrix} W_1(s) \\ W_2(s) \\ W_3(s) \end{pmatrix} \equiv W(s)$$

where

$$W_1(s) = \frac{W(s, R(\gamma_0))}{\sqrt{R(\gamma_0)}}$$
$$W_2(s) = \frac{W(s,1) - W(s, R(\gamma_0))}{\sqrt{1-R(\gamma_0)}}$$

Then, the vector Brownian motion $W(s)$ has the covariance matrix

$$\mathbb{E}W(1)W(1)' = \begin{pmatrix} 1 & 0 & \delta_1 \\ 0 & 1 & \delta_2 \\ \delta_1 & \delta_2 & 1 \end{pmatrix}$$

where δ_1 is the long run correlation between $e_t R_t(\gamma_0)$ and Δy_{t+k} and δ_2 of $e_t[1-R_t(\gamma_0)]$ and Δy_{t+k}.

From the covariance matrix, we can obtain

$$\begin{pmatrix} W_1(s) \\ W_2(s) \end{pmatrix} = \begin{pmatrix} \sqrt{1-\delta_1^2}W_{1-3}(s) \\ \sqrt{1-\delta_1^2}W_{2-3}(s) \end{pmatrix} + \begin{pmatrix} \delta_1 \\ \delta_2 \end{pmatrix} W_3(s)$$

where (W_{1-3}, W_{2-3}) is independent of W_3.

Proof of Proposition 2

Similar to proposition 1, we denote $\overset{p}{\Rightarrow}$ as weak convergence in probability. As the invariance principle for the bootstrap residuals has been introduced in the proof of Proposition 1, we can easily prove it by following the proof of Theorem 3.

Denote the estimated threshold values as $\hat{\gamma}$ and the fitted estimates as $\hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_{p-1})'$ and $\hat{\psi} = (\hat{\psi}_1, \ldots, \hat{\psi}_{p-1})'$. Under the unit root hypothesis, for any γ we have

$$\tilde{t}_1(\gamma) = \frac{\tilde{N}_T(\gamma) + \tilde{A}_T(\gamma)}{\sqrt{\tilde{D}_T(\gamma)\hat{\sigma}^2}}$$

where

$$\tilde{N}_T(\gamma) = \frac{1}{T}\sum_{t=1}^{T}\tilde{y}_{t-1}\tilde{e}_t\tilde{R}_t(\gamma) - \frac{1}{T}\sum_{t=1}^{T}\tilde{y}_{t-1}\tilde{w}_{t-1}\tilde{R}_t(\gamma)\left(\sum_{t=1}^{T}\tilde{w}_{t-1}^2\tilde{R}_t(\gamma)\right)^{-1}\sum_{t=1}^{T}\tilde{w}_{t-1}\tilde{e}_t\tilde{R}_t(\gamma),$$

$$\tilde{A}_T(\gamma) = \frac{1}{T}\sum_{t=1}^{T}\tilde{y}_{t-1}\tilde{R}_t(\gamma)\tilde{w}_t'(\hat{\phi} - \hat{\psi})[\tilde{R}_t(\hat{\gamma}) - \tilde{R}_t(\gamma)],$$

$$\tilde{D}_T(\gamma) = \frac{1}{T^2}\sum_{t=1}^{T}\tilde{y}_{t-1}^2\tilde{R}_t(\gamma) - \frac{1}{T}\sum_{t=1}^{T}\tilde{y}_{t-1}\tilde{w}_{t-1}\tilde{R}_t(\gamma)\left(\sum_{t=1}^{T}\tilde{w}_{t-1}^2\tilde{R}_t(\gamma)\right)^{-1}\sum_{t=1}^{T}\tilde{w}_{t-1}\tilde{e}_t\tilde{R}_t(\gamma)$$

Following the proof of Theorem 3, we can show that $\tilde{A}_T(\gamma) \xrightarrow{p} 0$. Moreover, by the bootstrap version of Lemma 1 and 2, we can obtain that

$$\tilde{t}_1 \Rightarrow \sqrt{1 - \delta^2(\hat{\gamma})Z_1} + \delta^2(\hat{\gamma})DF.$$

The limiting distribution of $\tilde{t}_2$ can be obtained similarly. Finally, since $\delta^2(\hat{\gamma})$ and the distribution of Z_1 and Z_2 are all continuous with respect to the consistent estimator $\hat{\gamma}$, $\hat{\phi}$ and $\hat{\psi}$, by the continuous mapping theorem, we can proof the Proposition 2.

References

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

Giné, E. & J. Zinn (1990). Bootstrapping general empirical measures. *Annals of Probability* **18**, 85169.

Li, G., Guan, B., Li, W. K., and Yu. P. L. H. (2015). Hysteretic autoregressive time series models. *Biometrika* **102**, 717-723.

Hansen, B. (2017). *Econometrics*. Retrieved from `http://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf`.

Zhu, K., Yu, P. L. H. and Li, W. K. (2014). Testing for the buffered autoregressive processes. *Statistica Sinica* **24**, 971-984.