Supplementary Material for "Bootstrap-Assisted Unit Root Testing With Piecewise Locally Stationary Errors"

Yeonwoo Rho[1] and Xiaofeng Shao[2]

Michigan Technological University[1]

University of Illinois at Urbana-Champaign[2]

February 9, 2018

This supplementary material contains all technical proofs for results (Section A), implementation details for the DWB and RDWB methods (Section B), and full power curves for all the models (Section C).

A Technical Appendix

The symbols $O_p(1)$ and $o_p(1)$ signify being bounded in probability and convergence to zero in probability, respectively. Denote by P^*, E^*, and var* the probability, expectation, and variance, respectively, conditional on data $\mathcal{X}_n = (X_{1,n}, \dots, X_{n,n})$. For notational simplicity, the dependence of $X_{t,n}$, $u_{t,n}$, and $W_{t,n}$ on n are often suppressed, and these quantities are written as X_t, u_t, and W_t, respectively. For a sequence of random variables $\{Y_n\}$, $Y_n = o_p^*(1)$ in probability is used if for any $\epsilon > 0$, $P^*\{|Y_n| > \epsilon\} \to 0$ in probability, as defined in Chang and Park (2003, p.386). We define $S_t = S_{t,n} = \sum_{i=1}^t u_{i,n}$. The positive constant C is generic and may vary from place to place. The symbol $\mathcal{I}_j$ is used in different places to indicate different objects. For notational simplicity, we often write $G(s, \mathcal{F}_t) := G_{\zeta_s}(s, \mathcal{F}_t)$ and $c(s;h) := c_{\zeta_s}(s;h)$, omitting the subscript ζ_s, where $\zeta_s = j$ such that $s \in [b_j, b_{j+1})$ and $\zeta_s = \tau$ if $s = 1$. Let $\gamma_h(r) = \int_0^r c(s;h)ds$. Notice that by definition, $\gamma_0(1) = \sigma_u^2$, and these symbols are interchangeably used in the proofs.

Recall that $\mathcal{F}_t = (\dots, \varepsilon_{t-1}, \varepsilon_t)$ with ε_t i.i.d. (0,1), and $\{\varepsilon_t'\}$ is an i.i.d. copy of $\{\varepsilon_t\}$. Following Wu (2005), for $I \subset \mathbb{Z}$, define $\mathcal{F}_{t,I}$ be the same as $\mathcal{F}_t$ except that ε_j is replaced by ε_j for $j \in I$. In particular, for $i \le t$, $\mathcal{F}_{t,\{i\}} = (\dots, \varepsilon_{i-1}, \varepsilon_i', \varepsilon_{i+1}, \dots, \varepsilon_t)$. Denote by $\mathcal{F}_{t,i}^* = \mathcal{F}_{t,\{k \in \mathbb{Z}: k \le i\}}$.

To keep the proofs concise, the case with no deterministic trend functions, i.e., $\beta \equiv 0$, is presented. The statements in Theorems 2.1, 3.1, and 3.2 hold by replacing $B_\sigma(r)$ with $B_{\sigma|Z}(r)$ and X_t with $\widehat{X}_t$. The following four lemmas prove some basic properties of $\{u_t\}$ and $\{X_t\}$ that are useful in the subsequent proofs.

Lemma **A.1**. *Assume (A1)-(A4). Fix* $j \in \{0, 1, \ldots, \tau\}$.

(i) For any $t, t' \in [b_j n, b_{j+1} n)$, $\big|cov(u_t, u_{t'}) - c_j(t/n; |t-t'|)\big| \le C(|t-t'|/n)$.

(ii) For any $s \ne s' \in [b_j, b_{j+1}]$, $\big|c_j(s;h) - c_j(s';h)\big| \le C|s-s'|$ *uniformly over* $h \in \mathbb{N}$.

(iii) For any $\rho > 0$, $\sup_{s\in[b_j,b_{j+1}]} \sum_{h=0}^{\infty} \big|h^\rho c_j(s,h)\big| \le C \sum_{h=0}^{\infty} h^\rho \chi^h < \infty$.

(iv) $\displaystyle\sup_{b_j \le s \ne s' < b_{j+1}} \frac{|\sigma(s) - \sigma(s')|}{|s-s'|(-\log|s-s'|+1)} \le C$.

In addition, if $j = \tau$, *(i) and (iv) also hold for all* $t, t' \in [b_\tau n, n]$ *or for supremum over* $\{b_\tau \le s \ne s' \le 1\}$.

Proof of Lemma A.1. (i) For all $t, t' \in [b_j n, b_{j+1} n)$,

$$\text{cov}(u_t, u_{t'}) = c_j(t/n; |t-t'|) - \text{cov}\big\{G_j(t/n, \mathcal{F}_t), G_j(t/n, \mathcal{F}_{t'}) - G_j(t'/n, \mathcal{F}_{t'})\big\}.$$

From the Cauchy-Schwartz inequality and (A1), $\big|\text{cov}\big\{G_j(t/n, \mathcal{F}_t), G_j(j/n, \mathcal{F}_{t'}) - G_j(t'/n, \mathcal{F}_{t'})\big\}\big| \le ||G_j(t/n, \mathcal{F}_t)||_2 \; ||G_j(t/n, \mathcal{F}_{t'}) - G_j(t'/n, \mathcal{F}_{t'})||_2 \le C(|t-t'|/n)$, which completes the proof. If $j = \tau$, the same argument holds for all $t, t' \in [b_\tau n, n]$.

(ii) It follows from the triangular inequality, Cauchy-Schwartz inequality, and (A1) that for any $s \ne s' \in [b_j, b_{j+1}]$, $|c_j(s;h) - c_j(s';h)| \le ||G_j(s, \mathcal{F}_0)||_2 ||G_j(s, \mathcal{F}_h) - G_j(s', \mathcal{F}_h)||_2 + ||G_j(s', \mathcal{F}_h)||_2 ||G_j(s, \mathcal{F}_0) - G_j(s', \mathcal{F}_0)\}||_2 \le C|s-s'|$ holds uniformly over $h \in \mathbb{N}$.

(iii) This is a straightforward consequence of Lemma A.1 in Shao and Wu (2007), Theorem 1 in Wu (2005), and the assumption (A3).

(iv) It follows from (A3) that $|c_j(s;h) - c_j(s';h)| \le 2C\chi^h$ for all $h \in \mathbb{N}$ and $s, s' \in [b_j, b_{j+1})$. Let m be the smallest positive integer such that $\chi^m \le |s-s'|$. Then using (ii), $|\sigma(s) - \sigma(s')| \le \sum_{h=-\infty}^{\infty} |c_j(s;h) - c_j(s';h)| \le C(\sum_{|h|\le m-1} |s-s'| + \sum_{|h|\ge m} \chi^h) \le C\{m|s-s'| + \chi^m(1-\chi)^{-1}\} \le C|s-s'|(-\log\chi)^{-1}(-\log|s-s'|) + C(1-\chi)^{-1}|s-s'| \le C|s-s'|(-\log|s-s'|+1)$. Notice that constant C's do not depend on s or s'. Thus the proof is complete. If $j = \tau$, the same argument holds for $s, s' \in [b_\tau, 1]$. ♢

Lemma **A.2**. *Under the conditions (A2)-(A3), for any* $i = 1, \ldots, n-h$, $h = 0, \ldots, n-i$,

$$|E(u_i u_{i+h})| \le C\chi^h,$$

where C *is a constant that does not depend on* h, i, *or* n *and* χ *is from (A3).*

Proof of Lemma A.2. By definition, $\mathcal{F}_i$ and $\mathcal{F}^*_{i+h,i}$ are independent. Therefore, $E\{G(i/n, \mathcal{F}_i)G((i+h)/n, \mathcal{F}^*_{i+h,i})\} = 0$, and

$$\begin{aligned} E(u_i u_{i+h}) &= E\left[G(i/n, \mathcal{F}_i)\{G((i+h)/n, \mathcal{F}_{i+h}) - G((i+h)/n, \mathcal{F}_{i+h,\{i\}})\}\right] \\ &\quad + E\left[G(i/n, \mathcal{F}_i)\{G((i+h)/n, \mathcal{F}_{i+h,\{i\}}) - G((i+h)/n, \mathcal{F}^*_{i+h,i})\}\right] \end{aligned}$$

Then, by the Cauchy-Schwartz inequality,

$$\begin{aligned} |E(u_i u_{i+h})| &\leq ||G(i/n, \mathcal{F}_i)||_2 ||G((i+h)/n, \mathcal{F}_{i+h}) - G((i+h)/n, \mathcal{F}_{i+h,\{i\}})||_2 \\ &\quad + ||G(i/n, \mathcal{F}_i)||_2 ||G((i+h)/n, \mathcal{F}_{i+h,\{i\}}) - G((i+h)/n, \mathcal{F}^*_{i+h,i})||_2 \end{aligned}$$

By (A2), $||G(i/n, \mathcal{F}_i)||_2 < C < \infty$, and by (A3) $||G((i+h)/n, \mathcal{F}_{i+h}) - G((i+h)/n, \mathcal{F}_{i+h,\{i\}})||_2 < ||G((i+h)/n, \mathcal{F}_{i+h}) - G((i+h)/n, \mathcal{F}_{i+h,\{i\}})||_4 \leq C\chi^h$. Thus the first term is bounded by $C\chi^h$, where C does not depend on h, i, or n.

Now the proof is complete if the following statement is shown:

$$||G((i+h)/n, \mathcal{F}_{i+h,\{i\}}) - G((i+h)/n, \mathcal{F}^*_{i+h,i})||_4 \leq C\chi^h.$$

Define $\mathcal{F}^*_{i+h,\{i\},m} = \mathcal{F}_{i+h,A}$, where $A = \{k \in \mathbb{Z} : k \leq i-m-1\} \bigcup \{i\}$. In particular, if $m = 0$, $\mathcal{F}^*_{i+h,\{i\},0} = \mathcal{F}^*_{i+h,i}$. Then $||G((i+h)/n, \mathcal{F}_{i+h,\{i\}}) - G((i+h)/n, \mathcal{F}^*_{i+h,i})||_4 = ||\sum_{m=0}^{\infty} G((i+h)/n, \mathcal{F}^*_{i+h,\{i\},m}) - G((i+h)/n, \mathcal{F}^*_{i+h,\{i\},m+1})||_4 \leq \sum_{m=0}^{\infty} ||G((i+h)/n, \mathcal{F}^*_{i+h,\{i\},m}) - G((i+h)/n, \mathcal{F}^*_{i+h,\{i\},m+1})||_4 \leq C\sum_{m=0}^{\infty} \chi^{h+m+1} = C\chi^{h+1}/(1-\chi) \leq C\chi^h$, where the last C does not depend on h, i, or n. Thus the proof is complete. $\diamondsuit$

Let $\mathrm{cum}(Y_0, Y_1, Y_2, Y_3)$ denote the fourth-order cumulant. When $E(Y_i) = 0$, $i = 0, 1, 2, 3$, the following relation (see page 36 in Rosenblatt (1985), for example) is often used in the subsequent proofs:

$$\mathrm{cov}(Y_0Y_1, Y_2Y_3) = E(Y_0Y_2)E(Y_1Y_3) + E(Y_0Y_3)E(Y_1Y_2) + \mathrm{cum}(Y_0, Y_1, Y_2, Y_3). \quad \text{(A.1)}$$

LEMMA **A.3**. *Assume (A1)-(A4). Then*

$$\sup_{1 \leq t_1 \leq \ldots \leq t_4 \leq n} |\mathrm{cum}(u_{t_1}, u_{t_2}, u_{t_3}, u_{t_4})| \leq C\chi^{(t_4 - t_1)/3},$$

with χ as in (A3).

Proof of Lemma A.3. Let $\mathcal{F}'_t = \mathcal{F}^*_{t,0}$ if $t > 0$, and $\mathcal{F}'_t = \mathcal{F}_t$ if $t \leq 0$. Define $\mathcal{F}'_{t,m} = \mathcal{F}_{t,\{k \in \mathbb{Z}: -m \leq k \leq 0\}}$ for $m \geq 0$ and $t > 0$. The argument is similar to the proof of Proposition

2 in Wu and Shao (2004). Let $1 \le t_1 \le \ldots \le t_4 \le n$, and $m_k = t_{k+1} - t_k$ for $k \in \{1, 2, 3\}$. Since for a fixed $s \in [0, 1]$, the process $\{G(s, \mathcal{F}_t)\}_t$ is stationary,

$$
\begin{aligned}
& \mathrm{cum}(u_{t_1}, u_{t_2}, u_{t_3}, u_{t_4}) \\
= \; & \mathrm{cum}\{G(t_1/n, \mathcal{F}_{t_1-t_k}), G(t_2/n, \mathcal{F}_{t_2-t_k}) - G(t_2/n, \mathcal{F}'_{t_2-t_k}), G(t_3/n, \mathcal{F}_{t_3-t_k}), G(t_4/n, \mathcal{F}_{t_4-t_k})\} \\
& + \mathrm{cum}\{G(t_1/n, \mathcal{F}_{t_1-t_k}), G(t_2/n, \mathcal{F}'_{t_2-t_k}), G(t_3/n, \mathcal{F}_{t_3-t_k}) - G(t_3/n, \mathcal{F}'_{t_3-t_k}), G(t_4/n, \mathcal{F}_{t_4-t_k})\} \\
& + \mathrm{cum}\{G(t_1/n, \mathcal{F}_{t_1-t_k}), G(t_2/n, \mathcal{F}'_{t_2-t_k}), G(t_3/n, \mathcal{F}'_{t_3-t_k}), G(t_4/n, \mathcal{F}_{t_4-t_k}) - G(t_4/n, \mathcal{F}'_{t_4-t_k})\} \\
& + \mathrm{cum}\{G(t_1/n, \mathcal{F}_{t_1-t_k}), G(t_2/n, \mathcal{F}'_{t_2-t_k}), G(t_3/n, \mathcal{F}'_{t_3-t_k}), G(t_4/n, \mathcal{F}'_{t_4-t_k})\} \\
:= \; & \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4
\end{aligned}
$$

due to the additive property of cumulants [the property (iii) on page 35 in Rosenblatt (1985)]. First we claim that $\mathcal{I}_4 = 0$. If $k = 1$, $\mathcal{F}_{t_1-t_k} = \mathcal{F}_0$ is independent of $\mathcal{F}'_{t_2-t_k}$, $\mathcal{F}'_{t_3-t_k}$, $\mathcal{F}'_{t_4-t_k}$, so $\mathcal{I}_4 = 0$ using the property (ii) on page 35 in Rosenblatt (1985). If $k = 2$, then $\mathcal{F}'_{t_2-t_k} = \mathcal{F}'_0 = \mathcal{F}_0$ by definition, and $\mathcal{F}_{t_1-t_k}$ and $\mathcal{F}_0$ are independent of $\mathcal{F}'_{t_3-t_k}$, $\mathcal{F}'_{t_4-t_k}$, which leads to $\mathcal{I}_4 = 0$. Similarly, if $k = 3$, $\mathcal{F}_{t_1-t_k}$, $\mathcal{F}'_{t_2-t_k} = \mathcal{F}_{t_2-t_k}$, and $\mathcal{F}_0$ are independent of $\mathcal{F}'_{t_4-t_k}$. Thus $\mathcal{I}_4 = 0$ for all $k = 1, 2, 3$. Also, notice that since $\mathcal{F}'_t = \mathcal{F}_t$ if $t \le 0$, it can be shown that $\mathcal{I}_1 = 0$ if $k = 2$ and $\mathcal{I}_1 = \mathcal{I}_2 = 0$ if $k = 3$. Thus the proof is done if the following statement is proved for each $k = 1, 2, 3$:

$$\max_{k \le i \le 3} |\mathcal{I}_i| \le C\chi^{m_k}. \tag{A.2}$$

Once (A.2) is shown, it follows that for each $k = 1, 2, 3$, $|\mathrm{cum}(u_{t_1}, u_{t_2}, u_{t_3}, u_{t_4})| \le C\chi^{m_k}$. Taking the minimum over k for both sides yields $|\mathrm{cum}(u_{t_1}, u_{t_2}, u_{t_3}, u_{t_4})| \le C \min_{k=1,2,3} \chi^{m_k} = C\chi^{\max_{k=1,2,3} m_k} \le C\chi^{(t_4-t_1)/3}$, since $t_4 - t_1 = \sum_{j=2}^4 (t_j - t_{j-1}) \le 3 \max_{j=2,3,4}(t_j - t_{j-1}) = 3 \max_{k=1,2,3} m_k$.

The subsequent arguments prove (A.2). For each $k = 1, 2, 3$, fix any $j = k+1, \ldots, 4$. Let $Y_0 = G(t_j/n, \mathcal{F}_{t_j-t_k}) - G(t_j/n, \mathcal{F}'_{t_j-t_k})$ and Y_1, Y_2, and Y_3 be the other variables in $\mathcal{I}_{j-1}$ so that we can write $\mathcal{I}_{j-1} = \mathrm{cum}(Y_0, Y_1, Y_2, Y_3)$. Since $||Y_0||_4 \le ||\{G(t_j/n, \mathcal{F}_{t_j-t_k}) - G(t_j/n, \mathcal{F}'_{t_j-t_k,0})\}||_4 + \sum_{m=0}^\infty ||\{G(t_j/n, \mathcal{F}'_{t_j-t_k,m}) - G(t_j/n, \mathcal{F}'_{t_j-t_k,m+1})\}||_4 \le C\{\chi^{t_j-t_k} + \sum_{m=0}^\infty \chi^{t_j-t_k+m+1}\} \le C\chi^{t_j-t_k}$ holds by the triangular inequality and (A3), it follows that

$$||Y_0||_4 \le C\chi^{t_j-t_k}, \tag{A.3}$$

where C is a constant that does not depend on t_j, j, or n. Observe that due to (A.1), $\mathcal{I}_{j-1} = E(Y_0Y_1Y_2Y_3) - E(Y_0Y_1)E(Y_2Y_3) - E(Y_0Y_2)E(Y_1Y_3) - E(Y_0Y_3)E(Y_1Y_2)$. By Hölder's inequality, (A.3), and (A2), it follows that $|E(Y_0Y_1Y_2Y_3)| \le ||Y_0||_4 ||Y_1Y_2Y_3||_{4/3} \le C\chi^{t_j-t_k}$ and $|E(Y_0Y_i)| \le ||Y_0||_2 ||Y_i||_2 \le C\chi^{t_j-t_k}$. Thus $|\mathcal{I}_{j-1}| \le C\chi^{t_j-t_k} \le C\chi^{m_k}$, and (A.2) is proved. $\diamondsuit$

Lemma **A.4**. *Assume (A1)-(A4). Under the local alternatives $\rho = 1 + c/n$, $c \le 0$,*

$$\sup_{1\le t\le n} \{E(X_t^2)/t\} \le C \quad \text{and} \quad \sup_{1\le t\le n} \{E(X_t^4)/t^2\} \le C,$$

where C is a positive constant that does not depend on n.

Proof of Lemma A.4. In this proof, all C's indicate a constant that do not depend on t or n. Suppose $1 \le i_1 \le i_2 \le i_3 \le i_4 \le t$ for some $t = 1, \dots, n$. By (A.1) and Lemmas A.2 and A.3,

$$\begin{aligned} E(u_{i_1}u_{i_2}u_{i_3}u_{i_4}) &= E(u_{i_1}u_{i_2})E(u_{i_3}u_{i_4}) + E(u_{i_1}u_{i_3})E(u_{i_2}u_{i_4}) + E(u_{i_1}u_{i_4})E(u_{i_2}u_{i_3}) \\ &\quad + \text{cum}(u_{i_1}, u_{i_2}, u_{i_3}, u_{i_4}) \\ &\le C(\chi^{i_2-i_1}\chi^{i_4-i_3} + \chi^{i_3-i_1}\chi^{i_4-i_2} + \chi^{i_4-i_1}\chi^{i_3-i_2} + \chi^{(i_4-i_1)/3}). \end{aligned}$$

It follows that $E(X_t^4) = E(\sum_{i=1}^t \rho^{t-i}u_i)^4 = 24\sum_{1\le i_1\le i_2\le i_3\le i_4\le t} \rho^{4t-i_1-i_2-i_3-i_4}E(u_{i_1}u_{i_2}u_{i_3}u_{i_4}) \le Ct^2$, where the last inequality holds by observing the following four simple facts:

i. $\sum_{1\le i_1\le i_2\le i_3\le i_4\le t} \chi^{(i_4-i_1)/3} = \sum_{h=0}^{t-1}(t-h)(h+1)^2\chi^{h/3} \le Ct$.

ii. $\{\sum_{i_1,i_2} \chi^{|i_1-i_2|}\}\{\sum_{i_3,i_4} \chi^{|i_3-i_4|}\} \le (Ct)^2$.

iii. χ is strictly positive.

iv. $\rho^{t-i} = (1 + c/n)^{t-i} \le 1$ for any $i \le t$.

Similarly, $E(X_t^2) = E(\sum_{i=1}^t \rho^{t-i}u_i)^2 = 2\sum_{1\le i_1\le i_2\le t} \rho^{2t-i_1-i_2}E(u_{i_1}u_{i_2}) \le C\sum_{h=0}^{t-1}(t-h)\chi^h \le Ct$. ♢

The following lemmas contain key results needed in the proof of Theorem 2.1 and they may be of independent interest.

Lemma **A.5**. *Assume (A1)-(A4).*

(i) $n^{-1/2}S_{\lfloor nr\rfloor} = n^{-1/2}\sum_{i=1}^{\lfloor nr\rfloor} u_i \Rightarrow B_\sigma(r) = \int_0^r \sigma(s)dB(s)$.

(ii) For a fixed $r \in (0, 1]$ and a fixed integer $h \ge 0$, $|n^{-1}\sum_{i=1}^{\lfloor nr\rfloor\wedge(n-h)} u_iu_{i+h} - \gamma_h(r)| = o_p(1)$. Recall that $\gamma_h(r) = \int_0^r c_{\zeta_s}(s; h)ds$, where $\zeta_s = j$ such that $s \in [b_j, b_{j+1})$.

Proof of Lemma A.5. (i) Define a step function $\sigma_n(s) = \sigma(t/n)$ for $s \in [t/n, (t+1)/n)$ and $t = 0, 1, \ldots, n$, with $\sigma_n(1) = \sigma(1)$. Let $\check{B}_{n,\sigma}(r) = \int_0^{\lfloor nr \rfloor/n} \sigma_n(s)dB(s)$ and $\tilde{B}_{n,\sigma}(r) = \int_0^r \sigma_n(s)dB(s)$. Recall that $B_\sigma(r) = \int_0^r \sigma(s)dB(s)$.

By the triangle inequality, $\sup_{r\in[0,1]} |\check{B}_{n,\sigma}(r) - B_\sigma(r)| \le \sup_{r\in[0,1]} |\check{B}_{n,\sigma}(r) - \tilde{B}_{n,\sigma}(r)| + \sup_{r\in[0,1]} |\tilde{B}_{n,\sigma}(r) - B_\sigma(r)| =: \mathcal{I}_1 + \mathcal{I}_2$. It follows that $\mathcal{I}_1 = o_p(1)$ because $\sup_{r\in[0,1]} |r - \lfloor nr \rfloor/n| \le 1/n$ and $\sup_{r\in[0,1]} |\int_{\lfloor nr \rfloor/n}^r \sigma_n(s)dB(s)| \le C \sup_{t=1,\ldots,n} |B(t/n) - B((t-1)/n)| = o_p(1)$. Notice that by Lemma A.1 (iv), $\sup_{r\in[0,1]} |\sigma_n(r) - \sigma(r)| = \sup_{0\le j\le\tau} \sup_{b_j\le s<b_{j+1}} |\sigma_n(s) - \sigma(s)| = \sup_{0\le j\le\tau} \sup_{b_j\le s<b_{j+1}} |\sigma(\lfloor ns \rfloor/n) - \sigma(s)| \le (\tau+1)C|\lfloor ns \rfloor/n - s|(-\log|\lfloor ns \rfloor/n - s| + 1) = O(n^{-1}\log n) = o(1)$. Thus $\mathcal{I}_2 = o_p(1)$ holds by Kurtz (2001, Proposition 5.19). It follows that

$$\sup_{r\in[0,1]} |\check{B}_{n,\sigma}(r) - B_\sigma(r)| = o_p(1). \tag{A.4}$$

From Proposition 5 in Zhou (2013), on a richer probability space, there exist i.i.d. standard normal random variables $V_1, \ldots, V_n$ such that

$$\sup_{r\in[0,1]} |n^{-1/2} S_{\lfloor nr \rfloor} - \hat{B}_{n,\sigma}(r)| = o_p(1), \tag{A.5}$$

where $\hat{B}_{n,\sigma}(r) = n^{-1/2} \sum_{i=1}^{\lfloor nr \rfloor} \sigma(i/n) V_i$. Since $\{\hat{B}_{n,\sigma}(r)\}_{r\in[0,1]} \overset{\mathcal{D}}{=} \{\sum_{t=1}^{\lfloor nr \rfloor} \sigma(t/n)[B(t/n) - B\{(t-1)/n\}]\}_{r\in[0,1]} \overset{\mathcal{D}}{=} \{\check{B}_{n,\sigma}(r)\}_{r\in[0,1]}$,

$$\hat{B}_{n,\sigma}(r) \Rightarrow B_\sigma(r) \tag{A.6}$$

by (A.4). Then (i) follows from (A.5) and (A.6).

(ii) Define $Y_i = Y_{i,n} = u_i u_{i+h} - E(u_i u_{i+h})$. We claim that $|n^{-1} \sum_{i=1}^{\lfloor nr \rfloor \wedge (n-h)} Y_{i,n}| = o_p(1)$. Observe that by (A.1), for $i \ge i'$, $E(Y_i Y_{i'}) = \mathrm{cov}(u_i u_{i+h}, u_{i'} u_{i'+h}) = E(u_i u_{i'})E(u_{i+h} u_{i'+h}) + E(u_i u_{i'+h})E(u_{i+h} u_{i'}) + \mathrm{cum}(u_i, u_{i+h}, u_{i'}, u_{i'+h}) \le C\chi^{2|i-i'|} + C\chi^{|i-i'-h|+|i+h-i'|} + C\chi^{|i+h-i'|/3} \le C\chi^{|i-i'|/3}$, where the first inequality is due to Lemmas A.2 and A.3. Then, by Chebyshev's inequality, for any $\delta > 0$, $P(|\sum_{i=1}^{\lfloor nr \rfloor \wedge (n-h)} Y_{i,n}| > n\delta) \le (n\delta)^{-2} E(\sum_{i=1}^{\lfloor nr \rfloor \wedge (n-h)} Y_{i,n})^2 \le (n\delta)^{-2} \sum_{i,i'=1}^{\lfloor nr \rfloor \wedge (n-h)} E(Y_{i,n} Y_{i',n}) \le C(n\delta)^{-2} \sum_{i,i'=1}^{\lfloor nr \rfloor \wedge (n-h)} \chi^{|i-i'|/3} \le (n\delta)^{-2} Cn = o(1)$. Therefore, $|n^{-1} \sum_{i=1}^{n-h} \{u_i u_{i+h} - E(u_i u_{i+h})\}| = o_p(1)$.

Now it remains to show that $|n^{-1} \sum_{i=1}^{\lfloor nr \rfloor \wedge (n-h)} E(u_i u_{i+h}) - \gamma_h(r)| = o(1)$. For $r \in (0,1]$, let $\mathcal{B}_r = \{i : i/n < b_j < (i+h)/n$ for some b_j and $1 \le i \le \lfloor nr \rfloor \wedge (n-h)\}$ and τ_r be the number of break points in $(0, r)$. Since $n^{-1} \sum_{i=1}^{\lfloor nr \rfloor \wedge (n-h)} E(u_i u_{i+h}) = n^{-1} \sum_{i \notin \mathcal{B}_r} E(u_i u_{i+h}) + n^{-1} \sum_{i \in \mathcal{B}_r} E(u_i u_{i+h}) = \mathcal{I}_{1,r} + \mathcal{I}_{2,r}$, it suffices to show that

$$\sup_{r\in(0,1]} |\mathcal{I}_{1,r} - \gamma_h(r)| = o(1) \quad \text{and} \quad \sup_{r\in(0,1]} |\mathcal{I}_{2,r}| = o(1). \tag{A.7}$$

For $\mathcal{I}_{1,r}$, it follows from Lemma A.1 (i) that $|\mathcal{I}_{1,r} - \gamma_h(r)| \leq n^{-1}\sum_{i\notin\mathcal{B}_r}|E(u_i u_{i+h}) - c_{\zeta_{i/n}}(i/n;h)| + n^{-1}\sum_{i\in\mathcal{B}_r}|c_{\zeta_{i/n}}(i/n;h)| \leq Ch/n$ holds for a constant C that does not depend on r. For $\mathcal{I}_{2,r}$, $\sup_{r\in[0,1)}|\mathcal{I}_{2,r}| \leq \sup_{r\in[0,1)} C\tau_r h/n \leq C\tau h/n = O(h/n)$. Thus (A.7) holds, and the proof is complete. $\diamondsuit$

Lemma **A.6**. *Assume (A1)-(A4). Let $S_t = \sum_{i=1}^t u_{i,n}$. The following statements hold jointly.*

(i) For any $r \in (0,1]$, $n^{-2}\sum_{t=1}^{\lfloor nr\rfloor} S_{t-1}^2 \xrightarrow{\mathcal{D}} \int_0^r B_\sigma^2(s)ds$.

(ii) For any $r \in (0,1]$, $n^{-1}\sum_{t=1}^{\lfloor nr\rfloor} S_{t-1}u_t \xrightarrow{\mathcal{D}} 2^{-1}\{B_\sigma^2(r) - \gamma_0(r)\}$.

(iii) For any $r \in (0,1]$, $n^{-1}\sum_{t=1}^{\lfloor nr\rfloor\wedge(n-h)} S_{t-1}u_{t+h} \xrightarrow{\mathcal{D}} 2^{-1}\{B_\sigma^2(r) - \gamma_0(r)\} - \sum_{k=1}^h \gamma_k(r)$ for any fixed integer $h \geq 1$.

(iv) For any $r \in (0,1]$, $n^{-3/2}\sum_{t=1}^{\lfloor nr\rfloor} S_{t-1} \xrightarrow{\mathcal{D}} \int_0^r B_\sigma(s)ds$.

Proof of Lemma A.6. The proof can be done by standard arguments using the identity $2S_{t-1}u_t = S_t^2 - S_{t-1}^2 - u_t^2$, the continuous mapping theorem, and Lemma A.5. $\diamondsuit$

Lemma **A.7**. *Assume (A1)-(A4). Under the local alternatives $\rho = 1 + c/n$, $c \leq 0$,*

$$n^{-2}\sum_{t=1}^{\lfloor nr\rfloor} X_{t-1}^2 \Rightarrow \int_0^r J_{c,\sigma}^2(s)ds, \tag{A.8}$$

$$n^{-1}\sum_{t=1}^{n} X_{t-1}u_t \xrightarrow{\mathcal{D}} \int_0^1 J_{c,\sigma}(r)\sigma(r)dB(r) + 2^{-1}\left\{\int_0^1 \sigma^2(r)dr - \sigma_u^2\right\}, \tag{A.9}$$

and

$$s_n^2 = (n-2)^{-1}\sum_{t=1}^{n}(X_t - \widehat{\rho}_n X_{t-1})^2 \xrightarrow{\mathcal{P}} \gamma_0(1) = \sigma_u^2 \qquad \text{for } c = 0. \tag{A.10}$$

Proof of Lemma A.7. First observe that $e^{c/n} = 1 + c/n + O(n^{-2})$ so that $\rho_n = e^{c/n} + O(n^{-2})$. Then X_t is asymptotically equivalent to $\sum_{j=1}^t e^{(t-j)c/n}u_j$, i,e, $X_t = \sum_{j=1}^t \rho_n^{t-j}u_j = \sum_{j=1}^t e^{(t-j)c/n}u_j + O_p(n^{-3/2})$. Following the argument in Phillips (1987), page 539, and using Lemma A.5 (i), it can be shown that $n^{-1/2}\sum_{j=1}^{\lfloor nr\rfloor} e^{(t-j)c/n}u_j \Rightarrow J_{c,\sigma}(r)$, which implies

$$n^{-1/2}X_{\lfloor nr\rfloor} \Rightarrow J_{c,\sigma}(r). \tag{A.11}$$

Then (A.8) follows from the continuous mapping theorem. For (A.9), squaring both sides of (3) yields $X_t^2 = (1+cn^{-1})^2 X_{t-1}^2 + u_t^2 + 2(1+cn^{-1})X_{t-1}u_t$ so that $\sum_{t=1}^n X_t^2 = (1 + 2cn^{-1})\sum_{t=1}^n X_{t-1}^2 + \sum_{t=1}^n u_t^2 + 2\sum_{t=1}^n X_{t-1}u_t + O_p(1)$. Thus

$$\begin{aligned} 2n^{-1}\textstyle\sum_{t=1}^n X_{t-1}u_t &= n^{-1}X_n^2 - 2cn^{-2}\textstyle\sum_{t=1}^n X_{t-1}^2 - n^{-1}\sum_{t=1}^n u_t^2 + O_p(n^{-1}) \\ &\xrightarrow{\mathcal{D}} J_{c,\sigma}^2(1) - 2c\textstyle\int_0^1 J_{c,\sigma}^2(r)dr - \sigma_u^2 \\ &= 2\textstyle\int_0^1 J_{c,\sigma}(r)\sigma(r)dB(r) + \{\int_0^1 \sigma^2(r)dr - \sigma_u^2\}, \end{aligned}$$

which implies (A.9). Here, the last equality is due to

$$J_{c,\sigma}^2(1) = \int_0^1 \sigma^2(r)dr + 2c\int_0^1 J_{c,\sigma}^2(r)dr + 2\int_0^1 J_{c,\sigma}(r)\sigma(r)dB(r),$$

which follows from Itô's formula.[1]

For (A.10), notice that $s_n^2 = (n-2)^{-1}\sum_{t=1}^n (X_t - \widehat{\rho}_n X_{t-1})^2 = (n-2)^{-1}\sum_{t=1}^n u_t^2 + (n-2)^{-1}(\widehat{\rho}_n - \rho)^2\sum_{t=1}^n X_{t-1}^2 + 2(n-2)^{-1}(\rho - \widehat{\rho}_n)\sum_{t=1}^n X_{t-1}u_t := \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3$. Here $\mathcal{I}_1 \xrightarrow{\mathcal{P}} \gamma_0(1)$ by Lemma A.5 (ii). For $\mathcal{I}_2$, under the null hypothesis $\rho = 1$, $S_t = X_t = \sum_{i=1}^t u_i$. By Lemma A.6 (i) and (ii), $\widehat{\rho}_n - \rho = O_p(n^{-1})$ and $\sum_{t=1}^n X_{t-1}^2 = O_p(n^2)$ so that $\mathcal{I}_2 = O_p(n^{-1})$. Under the null, $\sum_{t=1}^n X_{t-1}u_t = O_p(n^{3/2})$ by the Cauchy-Schwartz inequality, which leads to $\mathcal{I}_3 = O_p(n^{-1/2})$. Thus the proof is complete.

$\diamondsuit$

Proof of Theorem 2.1. The proof is straightforward using the continuous mapping theorem, Lemma A.7, and Slutsky's theorem. $\diamondsuit$

We now prove bootstrap consistency. The proof can be done using the large-block small-block argument as presented in the proof of Theorem 3.1 in Shao (2010). Let $L_n = \lfloor (n/l_n)^{1/2} \rfloor$ be the length of a large-block and l_n be that of a small-block. Note that $L_n \to \infty$ and $l_n = o(L_n)$. Our goal is to assign points $t \in \{1, 2, \ldots, \lfloor nr \rfloor\}$ to alternating large and small blocks. Let $K_n = K_{n,r} = \lfloor \lfloor nr \rfloor (L_n + l_n)^{-1} \rfloor$ be the number of the large (small) blocks. Define the kth large-block $\mathcal{L}_k = \{j \in \mathbb{N} : (k-1)(L_n + l_n) + 1 \le j \le k(l_n + L_n) - l_n\}$ for $1 \le k \le K_n$, and the kth small-block $\mathcal{S}_k = \{j \in \mathbb{N} : k(L_n + l_n) - l_n + 1 \le j \le k(l_n + L_n)\}$ for $1 \le k \le K_n - 1$ and $\mathcal{S}_{K_n} = \{j \in \mathbb{N} : K_n(L_n + l_n) - l_n + 1 \le j \le \lfloor nr \rfloor\}$.

Let $U_k = \sum_{j \in \mathcal{L}_k} W_j u_j$ and $V_k = \sum_{j \in \mathcal{S}_k} W_j u_j$, $k = 1, ..., K_n$. Define $\mathcal{B}_L = \{k : \mathcal{L}_k$ contains a break point b_j for some $j = 0, \ldots, \tau\}$ and $\mathcal{B}_S = \{k : \mathcal{S}_k$ contains a break point b_j for some $j = 0, \ldots, \tau\}$. Notice that there are only finitely many (less than τ) elements in $\mathcal{B}_L$ and $\mathcal{B}_S$.

[1]Recall that $J_{c,\sigma}(r)$ is defined as $dJ_{c,\sigma}(r) = cJ_{c,\sigma}(r)dr + \sigma(r)dB(r)$. Using Itô's formula, we can derive $J_{c,\sigma}^2(r) = J_{c,\sigma}^2(0) + \int_0^r 2cJ_{c,\sigma}^2(s)ds + \int_0^r 2\sigma(s)J_{c,\sigma}(s)dB(s) + \int_0^r \sigma^2(s)ds$, which leads to the desired result.

Lemma **A.8**. *Assume (A1)-(A4) and (B1)-(B2). Then*

$$\sup_{r\in[0,1]}\left|n^{-1}\sum_{k=1}^{K_n}\sum_{j,j'\in\mathcal{L}_k} cov(u_j,u_{j'})a\{(j-j')/l_n\}-\int_0^r\sigma^2(s)ds\right|=o(1). \tag{A.12}$$

Proof of Lemma A.8. Suppose $k\notin\mathcal{B}_L$. We shall first show that

$$\sup_{r\in[0,1]}\left|n^{-1}\sum_{k\notin\mathcal{B}_L}\sum_{j,j'\in\mathcal{L}_k}\text{cov}(u_j,u_{j'})a\{(j-j')/l_n\}-\int_0^r\sigma^2(s)ds\right|=o(1). \tag{A.13}$$

Recall that $\zeta_s=j$ such that $s\in[b_j,b_{j+1})$ and $\zeta_1=\tau$, and $c(s;h)=c_{\zeta_s}(s;h)$. Since $a(\cdot)=0$ outside of its support [-1,1], by Lemma A.1 (i) and (ii), it follows that $L_n^{-1}\sum_{j,j'\in\mathcal{L}_k}\text{cov}(u_j,u_{j'})a\{(j-j')/l_n\}=c(k/K_n;0)+O(L_n/n)+2\sum_{h=1}^{l_n}(1-h/L_n)a(h/l_n)\{c(k/K_n;h)+O(L_n/n)\}=\sigma^2(k/K_n)-2\sum_{h=1}^{\infty}d_hc(k/K_n;h)+O(l_nL_n/n)$, where $d_h=1-(1-h/L_n)a(h/l_n)$ if $0\le h\le l_n$ and 1 if $h>l_n$. By (B2) and Lemma A.1 (iii), $\sum_{h=1}^{\infty}d_hc(k/K_n;h)\le Cl_n^{-q}\{k_q+o(1)\}\sum_{h=1}^{\infty}h^qc(k/K_n;h)+C\overline{a}L_n^{-1}\sum_{h=1}^{\infty}hc(k/K_n;h)\le C(l_n^{-q}+L_n^{-1})=o(1)$, where $\overline{a}=\sup_{s\in[-1,1]}a(s)$ and C is a constant that does not depend on k or r.

Therefore,

$$\sup_{k\notin\mathcal{B}_L}\left|L_n^{-1}\sum_{j,j'\in\mathcal{L}_k}\text{cov}(u_j,u_{j'})a\{(j-j')/l_n\}-\sigma^2(k/K_n)\right|\le C\{l_n^{-q}+L_n^{-1}\}=o(1) \tag{A.14}$$

so that $\sup_{r\in[0,1]}|n^{-1}\sum_{k\notin\mathcal{B}_L}\sum_{j,j'\in\mathcal{L}_k}\text{cov}(u_j,u_{j'})a\{(j-j')/l_n\}-n^{-1}\sum_{k\notin\mathcal{B}_L}\sigma^2(k/K_n)L_n|=o(1)$. Since $\sup_{r\in[0,1]}|n^{-1}\sum_{k\notin\mathcal{B}_L}\sigma^2(k/K_n)L_n-\sum_{k\notin\mathcal{B}_L}\int_{(k-1)/K_n}^{k/K_n}\sigma^2(s)ds|=o(1)$ by Lemma A.1 (iv) and $\sup_{r\in[0,1]}|\sum_{k\notin\mathcal{B}_L}\int_{(k-1)/K_n}^{k/K_n}\sigma^2(s)ds-\int_0^r\sigma^2(s)ds|=o(1)$, (A.13) is proved. If $k\in\mathcal{B}_L$, (A2) implies that

$$\left|n^{-1}\sum_{k\in\mathcal{B}_L}\sum_{j,j'\in\mathcal{L}_k}\text{cov}(u_j,u_{j'})a\{(j-j')/l_n\}\right|=O(n^{-1}\tau L_n^2)=o(1). \tag{A.15}$$

Thus (A.12) follows from (A.13) and (A.15). ◊

Lemma **A.9**. *Assume (A1)-(A4) and (B1)-(B2). For a fixed constant* $r\in(0,1]$,

$$n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor}W_tu_t\xrightarrow{\mathcal{D}}N\left(0,\int_0^r\sigma^2(s)ds\right)\qquad \textit{in probability.} \tag{A.16}$$

Proof of Lemma A.9. The left-hand side of (A.16) can be decomposed into large- and small-block parts as $n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor} W_t u_t = n^{-1/2}\sum_{k=1}^{K_n} U_k + n^{-1/2}\sum_{k=1}^{K_n} V_k$. Note that $E^*(U_k) = 0$ for all $k = 1, ..., K_n$ and since W_t's are l_n-dependent, $U_1, ..., U_{K_n}$ are independent random variables conditional on $\mathcal{X}_n$. The same property holds for $V_1, ..., V_{K_n}$.

First it will be shown that the large-block part converges to the limit in (A.16), i.e.,

$$n^{-1/2}\sum_{k=1}^{K_n} U_k \xrightarrow{\mathcal{D}} N\left(0, \int_0^r \sigma^2(s)ds\right) \quad \text{in probability.} \tag{A.17}$$

Using the same argument as in the equation (A.3) in Shao (2010) and Hölder's inequality, it follows that

$$\sum_{k=1}^{K_n} E^*|U_k|^4 \leq C l_n^2 L_n \sum_{k=1}^{K_n}\sum_{j\in\mathcal{L}_k} |u_j|^4. \tag{A.18}$$

The argument in Shao (2010) applies here because everything is conditional on $\mathcal{X}_n$, and the property of W_t remains the same. From (A2), $E|u_j|^4 \leq C$ for $j = 1, \ldots, n$, so that $\sum_{k=1}^{K_n}\sum_{j\in\mathcal{L}_k} |u_j|^4 \leq \sum_{j=1}^{n} |u_j|^4 = O_p(n)$. It follows that $\sum_{k=1}^{K_n} E^*|U_k|^4 = O_p(l_n^2 L_n n) = O_p\{(nl_n)^{3/2}\}$. Since for any $\epsilon > 0$, $E^*\{U_k^2 \mathbf{1}(|U_k| > n^{1/2}\epsilon)\} \leq (n^{1/2}\epsilon)^{-2} E^*\{|U_k|^4 \mathbf{1}(|U_k| > n^{1/2}\epsilon)\} \leq n^{-1}\epsilon^{-2}E^*|U_k|^4$ holds for all k, it follows that $n^{-1}\sum_{k=1}^{K_n} E^*\{U_k^2\mathbf{1}(|U_k| > n^{1/2}\epsilon)\} = O_p\{(l_n^3/n)^{1/2}\} = o_p(1)$. Then (A.17) follows from Lemma A.8.

Next it will be shown that the contribution from small-blocks $n^{-1/2}\sum_{k=1}^{K_n} V_k$ is negligible, i.e.,

$$n^{-1/2}\sum_{k=1}^{K_n} V_k = o_p^*(1). \tag{A.19}$$

For $k \notin \mathcal{B}_S$, by Lemma A.1 (i) and (iii), $E\{E^*(V_k^2)\} = E\left[\sum_{j,j'\in\mathcal{S}_k} u_j u_{j'} a\{(j-j')/l_n\}\right] = \sum_{j,j'\in\mathcal{S}_k} \mathrm{cov}(u_j, u_{j'}) a\{(j-j')/l_n\} \leq l_n \sum_{h=0}^{l_n-1}\{c(k/K_n; h) + C(l_n/n)\}a(h/l_n) \leq Cl_n$. For $k = K_n$, using a similar argument, $E\{E^*(V_{K_n}^2)\} \leq CL_n$. For $k \in \mathcal{B}_S$ and $k \neq K_n$, $E\{E^*(V_{K_n}^2)\} \leq C\tau l_n^2$. Since $\tau < \infty$, it follows that $\sum_{k=1}^{K_n} E\{E^*(V_k^2)\} \leq C(K_n l_n + l_n^2 + L_n) = o(n)$. Then (A.19) follows from the Markov inequality, independence of V_k's, and linearity of expectation. The proof is completed in view of (A.17) and (A.19). ◊

The following two lemmas are used in the proof of Theorem 3.1,

Lemma **A.10**. *Assume (A1)-(A4) and (B1)-(B2). Then for* $0 < r_1 < r_2 \leq 1$ *and* $n \geq n_0$ *for some positive integer* n_0*, conditional on the data* $\mathcal{X}_n$*,*

$$E^*\left|n^{-1/2}\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} W_t u_t\right|^4 \leq \overline{C}(\mathcal{X}_n)\left\{(r_2-r_1)^2 + n^{-p_1}(r_2-r_1)\right\}, \tag{A.20}$$

for some $p_1 > 0$, $\overline{C}(\mathcal{X}_n)$ that does not depend on r_1 or r_2, and $\overline{C}(\mathcal{X}_n) = O_p(1)$. Furthermore,

$$n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} W_t u_t \Rightarrow B_\sigma(r) \qquad \textit{in probability.} \tag{A.21}$$

Proof of Lemma A.10. First (A.20) will be proved using the large-block small-block argument. Recall that $U_k = \sum_{j \in \mathcal{L}_k} W_j u_j$ and $V_k = \sum_{j \in \mathcal{S}_k} W_j u_j$ for $k = 1, ..., K_n$, $L_n = \lfloor (n/l_n)^{1/2} \rfloor$, and $K_{n,r} = O(\lfloor \lfloor nr \rfloor (L_n + l_n)^{-1} \rfloor)$. Let $K_1 = K_{n,r_1}$ and $K_2 = K_{n,r_2}$ for convenience. Define $p_2 = (1 - 3\kappa)/2 > 0$ and $p_3 = \kappa q$, where κ and q are from (B1) and (B2), respectively. Define $p_1 = \min(p_2, p_3)$. By the Cr-inequality,

$$E^* \Bigg| \sum_{t=\lfloor nr_1 \rfloor + 1}^{\lfloor nr_2 \rfloor} W_t u_t \Bigg|^4 = E^* \Bigg| \sum_{k=K_1+1}^{K_2} U_k + \sum_{k=K_1+1}^{K_2} V_k \Bigg|^4 \leq 2^3 \left(E^* \Bigg| \sum_{k=K_1+1}^{K_2} U_k \Bigg|^4 + E^* \Bigg| \sum_{k=K_1+1}^{K_2} V_k \Bigg|^4 \right).$$

Since U_k and V_k are independent conditional on the data and have mean 0,

$$E^* \Bigg| \sum_{k=K_1+1}^{K_2} U_k \Bigg|^4 = \sum_{k=K_1+1}^{K_2} E^*(U_k^4) + \sum_{k \neq k'} E^*(U_k^2 U_{k'}^2) \leq \sum_{k=K_1+1}^{K_2} E^*(U_k^4) + \left\{ \sum_{k=K_1+1}^{K_2} E^*(U_k^2) \right\}^2,$$

and similarly for V_k.

For the large-block part, from (A.18) and (A2),

$$n^{-2} \sum_{k=K_1+1}^{K_2} E^*(U_k^4) \leq n^{-2} C l_n^2 L_n \sum_{k=K_1+1}^{K_2} \sum_{j \in \mathcal{L}_k} |u_j|^4 \leq C_1(\mathcal{X}_n) n^{-p_2} (r_2 - r_1), \tag{A.22}$$

where $C_1(\mathcal{X}_n) = O_p(1)$. By (A.12), (A.14), and (A.15), for any $0 \leq r_1 < r_2 \leq 1$, $E \big| n^{-1} \sum_{k=K_1+1}^{K_2} E^*(U_k^2) - \int_{r_1}^{r_2} \sigma^2(s) ds \big| \leq C\{l_n^{-q} + L_n^{-1}\} \leq C(n^{-p_3} + n^{-p_2}) \leq C n^{-p_1}$. Note that the constant C does not depend on r_1 or r_2. Therefore,

$$n^{-2} \left\{ \sum_{k=K_1+1}^{K_2} E^*(U_k^2) \right\}^2 \leq C_2(\mathcal{X}_n)(r_2 - r_1)^2 + C_3(\mathcal{X}_n) n^{-p_1} (r_2 - r_1), \tag{A.23}$$

where $c = \{\sup_{s \in [0,1]} \sigma^2(s)\}^2 < \infty$ is a constant and $C_2(\mathcal{X}_n)$ and $C_3(\mathcal{X}_n)$ are both $O_p(1)$.

For the small block part, note that $K_2 - K_1 \leq Cn(r_2 - r_1)/L_n = C(r_2 - r_1)(nl_n)^{1/2}$ by the definition of K_1, K_2, and L_n, and $E^*(V_k^4) = O_p(l_n^4)$ by (A2) and (B1). Therefore,

$$n^{-2} \sum_{k=K_1+1}^{K_2} E^*(V_k^4) = O_p\{n^{-2} l_n^4 (K_2 - K_1)\} = C_4(\mathcal{X}_n)(l_n^3/n) n^{-p_2} (r_2 - r_1), \tag{A.24}$$

where $C_4(\mathcal{X}_n) = O_p(1)$. Also, it has been shown that $n^{-1}\sum_{k=K_1+1}^{K_2} E^*(V_k^2) = O_p\{(K_2 - K_1)l_n/n\} = O_p(1)n^{-p_2}(r_2 - r_1)$, which implies that

$$\left\{n^{-1}\sum_{k=K_1+1}^{K_2} E^*(V_k^2)\right\}^2 = C_5(\mathcal{X}_n)n^{-2p_2}(r_2 - r_1)^2, \qquad \text{(A.25)}$$

where $C_5(\mathcal{X}_n) = O_p(1)$. It is worth noting that $C_j(\mathcal{X}_n)$, $j = 1, \ldots, 5$ in (A.22), (A.23), (A.24), and (A.25), does not depend on r_1 or r_2. Therefore an upper bound for the left-hand side of (A.20) is

$$2^3\left[\left\{C_2(\mathcal{X}_n) + C_5(\mathcal{X}_n)n^{-2p_2}\right\}(r_2 - r_1)^2 + \left\{C_1(\mathcal{X}_n) + C_3(\mathcal{X}_n) + C_4(\mathcal{X}_n)(l_n^3/n)\right\}n^{-p_1}(r_2 - r_1)\right],$$

so that (A.20) holds for large enough n with $\overline{C}(\mathcal{X}_n) = 2^3\max\{C_2(\mathcal{X}_n), C_1(\mathcal{X}_n)+C_3(\mathcal{X}_n)\}+1$.

For (A.21), the finite-dimensional convergence,

$$\left(n^{-1/2}\sum_{t=1}^{\lfloor nr_1\rfloor} W_t u_t, \ldots, n^{-1/2}\sum_{t=1}^{\lfloor nr_k\rfloor} W_t u_t\right) \xrightarrow{\mathcal{D}} \left\{\int_0^{r_1}\sigma(s)dB(s), \ldots, \int_0^{r_k}\sigma(s)dB(s)\right\}$$

in probability for any $k \in \mathbb{N}$ and $r_1, \ldots, r_k$, follows from a similar argument presented in Lemma A.9 and the Cramér-Wold device. The tightness follows from (A.20) and the argument of Theorem 2.1 in Shao and Yu (1996). This completes the proof for (A.21). $\Diamond$

Lemma **A.11**. *Under the conditions (A1)-(A4) and (B1)-(B2),*

$$n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor} X_{t-1}W_t(\rho - \widehat{\rho}_n) \Rightarrow 0 \qquad \textit{in probability}$$

under the local alternatives $\rho = 1 + c/n$, $c \leq 0$.

Proof of Lemma A.11. The proof follows once the following two statements are established:

$$\left|n^{-1/2}(\rho - \widehat{\rho}_n)\sum_{t=1}^{\lfloor nr\rfloor} X_{t-1}W_t\right| = o_p^*(1) \qquad \text{for any } r \in [0,1] \qquad \text{(A.26)}$$

and

$$E^*\left|n^{-1/2}(\rho - \widehat{\rho}_n)\sum_{t=\lfloor nr_1\rfloor+1}^{\lfloor nr_2\rfloor} X_{t-1}W_t\right|^4 \leq \overline{C}(\mathcal{X}_n)\{(r_2 - r_1)^2 + n^{-p_1}(r_2 - r_1)\}, \qquad \text{(A.27)}$$

where $p_1 > 0$, $\overline{C}(\mathcal{X}_n)$ is a constant that does not depend on r_1 or r_2 such that $\overline{C}(\mathcal{X}_n) = O_p(1)$. Note that $n(\widehat{\rho}_n - \rho) = (n^{-1}\sum_{t=1}^n X_{t-1}u_t)/(n^{-2}\sum_{t=1}^n X_{t-1}^2) = O_p(1)$ under the local alternatives by Lemma A.7 and the continuous mapping theorem.

Equation (A.26) holds trivially if $r = 0$. For any fixed $r \in (0, 1]$, by Chebyshev's inequality, $P^*(|\sum_{t=1}^{\lfloor nr \rfloor} X_{t-1}W_t| > \lambda) \leq \lambda^{-2}E^*|\sum_{t=1}^{\lfloor nr \rfloor} X_{t-1}W_t|^2 = C\lambda^{-2}\sum_{t=1}^{\lfloor nr \rfloor}\sum_{h=0}^{l_n} X_{t-1}X_{t+h-1}a(h/l_n)$ for any $\lambda > 0$. Observe that $E|X_{t-1}X_{t+h-1}| \leq ||X_{t-1}||_2||X_{t+h-1}||_2 \leq C(t+h)$ by the Cauchy-Schwarz inequality and Lemma A.4. For any $\delta > 0$, by letting $\lambda = n^{3/2}\delta$, it follows that $E\{P^*(|n^{-3/2}\sum_{t=1}^{\lfloor nr \rfloor} X_{t-1}W_t| > \delta)\} \leq Cn^{-3}\delta^{-2}\sum_{t=1}^n\sum_{h=0}^{l_n}(t+h) \leq Cn^{-3}\delta^{-2}(n^2 l_n) = O(n^{-1}l_n) = o(1)$. Thus (A.26) is established.

Equation (A.27) can be shown using the large- and small- block argument. Define indices for large and small blocks $\mathcal{S}_k$ and $\mathcal{L}_k$ as before. Decompose $\sum_{t=1}^{\lfloor nr \rfloor} X_{t-1}W_t = \sum_{k=1}^{K_{n,r}} \mathbf{U}_k + \sum_{k=1}^{K_{n,r}} \mathbf{V}_k$ into large and small blocks. Recall that $K_{n,r} = \lfloor \lfloor nr \rfloor (L_n + l_n)^{-1} \rfloor$ is the number of large and small blocks, $L_n = \lfloor (n/l)^{1/2} \rfloor$ is the length of the large block, and $l_n \asymp Cn^\kappa$ with $\kappa \in (0, 1/3)$. Let $K_1 = K_{n,r_1}$ and $K_2 = K_{n,r_2}$.

Following the same argument used in the proof of (A.20), the upper bounds of $\sum_{k=K_1+1}^{K_2} E^*(\mathbf{U}_k^4)$, $\sum_{k=K_1+1}^{K_2} E^*(\mathbf{U}_k^2)$, $\sum_{k=K_1+1}^{K_2} E^*(\mathbf{V}_k^4)$, and $\sum_{k=K_1+1}^{K_2} E^*(\mathbf{V}_k^2)$ shall be examined. In the subsequent argument, $C(\mathcal{X}_n)$, $C_1(\mathcal{X}_n)$, $C_2(\mathcal{X}_n)$, $C_3(\mathcal{X}_n)$, and $C_4(\mathcal{X}_n)$ are all $O_p(1)$ and do not depend on r_2 or r_1. In particular, $C(\mathcal{X}_n)$ may have different values in different places.

Following the same argument as in (22) or (A.3) in Shao (2010), $\sum_{k=K_1+1}^{K_2} E^*(\mathbf{U}_k^4) \leq Cl_n^2 L_n \sum_{k=K_1+1}^{K_2}\sum_{j\in\mathcal{L}_k}|X_{j-1}|^4 \leq C(\mathcal{X}_n)l_n^2 L_n \sum_{j=\lfloor nr_1 \rfloor+1}^{\lfloor nr_2 \rfloor} j^2 \leq C(\mathcal{X}_n)l_n^2 L_n(\lfloor nr_2 \rfloor^3 - \lfloor nr_1 \rfloor^3) \leq C(\mathcal{X}_n)l_n^2 L_n n^3(r_2 - r_1)$, where the second inequality is due to Lemma A.4. Since $l_n^2 L_n n^{-3} = l^{3/2}n^{-5/2} \asymp Cn^{-(3\kappa+5)/2}$, letting $p_1 = (3\kappa+5)/2$, it follows that

$$n^{-6}\sum_{k=K_1+1}^{K_2} E^*(\mathbf{U}_k^4) \leq C_1(\mathcal{X}_n)n^{-p_1}(r_2 - r_1). \tag{A.28}$$

By Lemma A.4, $E\{E^*(\mathbf{U}_k^2)\} = E\{E^*(\sum_{t\in\mathcal{L}_k} X_{t-1}W_t)^2\} \leq \sum_{t\in\mathcal{L}_k}\sum_{h=-l}^{l}|E(X_{t-1}X_{t-1+h})|a(h/l) \leq C\sum_{t\in\mathcal{L}_k}\sum_{h=-l}^{l} t$ so that $\sum_{k=K_1+1}^{K_2} E\{E^*(\mathbf{U}_k^2)\} \leq Cl_n(\lfloor nr_2 \rfloor^2 - \lfloor nr_1 \rfloor^2) \leq Cl_n n^2(r_2 - r_1)$ and

$$n^{-6}\left\{\sum_{k=K_1+1}^{K_2} E^*(\mathbf{U}_k^2)\right\}^2 \leq l_n^2 n^{-2}C_2(\mathcal{X}_n)(r_2 - r_1)^2. \tag{A.29}$$

The same arguments work for small blocks, replacing $\mathbf{U}_k$ in (A.28) and (A.29) with $\mathbf{V}_k$, which complete the proof of (A.27). ♢

We are now ready to prove Theorems 3.1 and 3.2.

Proof of Theorem 3.1. Observe that $n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor} u_t^* = n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor}\widehat{u}_t W_t = n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor}(X_t - \widehat{\rho}_n X_{t-1})W_t = n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor}(\rho X_{t-1} + u_t - \widehat{\rho}_n X_{t-1})W_t = \left\{n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor} X_{t-1}W_t\right\}(\rho - \widehat{\rho}_n) + n^{-1/2}\sum_{t=1}^{\lfloor nr\rfloor} W_t u_t =: \mathcal{I}_{1,r} + \mathcal{I}_{2,r}$. Noting that $\mathcal{I}_{1,r} \Rightarrow 0$ in probability by Lemma A.11 and $\mathcal{I}_{2,r} \Rightarrow B_\sigma(r)$ in probability by Lemma A.10, the proof is complete. ♢

Proof of Theorem 3.2. We claim that under the local alternatives,

$$n^{-1}\sum_{t=1}^{n}\{(u_t^*)^2 - E^*(u_t^*)^2\} = o_p^*(1) \quad \text{and} \tag{A.30}$$

$$n^{-1}\sum_{t=1}^{n}\{E^*(u_t^*)^2 - u_t^2\} = o_p(1). \tag{A.31}$$

Once (A.30) and (A.31) are established, it follows that $n^{-1}\sum_{t=1}^{n}\{(u_t^*)^2 - u_t^2\} = o_p^*(1)$. Then using a similar argument as in the proof of Lemma A.6 (i) and (ii), Theorem 3.2 follows from an application of the continuous mapping theorem, Theorem 3.1, and the fact that $n^{-1}\sum_{t=1}^{n} u_t^2 \xrightarrow{\mathcal{P}} \sigma_u^2$, which is due to Lemma A.5 (ii) and the argument in the proof of Theorem 5.1 in Paparoditis and Politis (2003).

To prove (A.31), write $n^{-1}\sum_{t=1}^{n}\{E^*(u_t^*)^2 - u_t^2\} = n^{-1}\sum_{t=1}^{n}(\widehat{u}_t^2 - u_t^2) = n^{-1}\sum_{t=1}^{n}[\{u_t + (\rho - \widehat{\rho}_n)X_{t-1}\}^2 - u_t^2] = (\rho - \widehat{\rho}_n)^2 n^{-1}\sum X_{t-1}^2 + 2(\rho - \widehat{\rho}_n)n^{-1}\sum X_{t-1}u_t =: \mathcal{I}_1 + \mathcal{I}_2$. Lemma A.7 implies that $\mathcal{I}_k = O_p(n^{-1})$ for all $k = 1, 2$ under the local alternatives.

Now we shall prove (A.30). Observe that $\sum_{t=1}^{n}\{(u_t^*)^2 - E^*(u_t^*)^2\} = \sum_{t=1}^{n}\widehat{u}_t^2(W_t^2 - 1)$. For any $\delta > 0$, $P^*\{|\sum_{t=1}^{n}\widehat{u}_t^2(W_t^2 - 1)| > n\delta\} \le (n\delta)^{-2}E^*\{\sum_{t=1}^{n}\widehat{u}_t^2(W_t^2 - 1)\}^2 \le (n\delta)^{-2}C\{\sum_{t=1}^{n}\sum_{h=0}^{l_n}\widehat{u}_t^2\widehat{u}_{t+h}^2\}$, and it remains to show $\sum_{t=1}^{n}\sum_{h=0}^{l_n}\widehat{u}_t^2\widehat{u}_{t+h}^2 = o_p(n^2)$. Since $\widehat{u}_t = u_t + (\rho - \widehat{\rho}_n)X_{t-1}$,

$$\begin{aligned}
\sum_{t=1}^{n}\sum_{h=0}^{l_n}\widehat{u}_t^2\widehat{u}_{t+h}^2 \;=\; & \sum_{t=1}^{n}\sum_{h=0}^{l_n} u_t^2 u_{t+h}^2 \\
& + 2(\rho - \widehat{\rho}_n)\sum_{t=1}^{n}\sum_{h=0}^{l_n}\{u_t^2 u_{t+h}X_{t+h-1} + u_{t+h}^2 u_t X_{t-1}\} \\
& + (\rho - \widehat{\rho}_n)^2\sum_{t=1}^{n}\sum_{h=0}^{l_n}\{u_t^2 X_{t+h-1}^2 + u_{t+h}^2 X_{t-1}^2 + 4u_t u_{t+h}X_{t-1}X_{t+h-1}\} \\
& + 2(\rho - \widehat{\rho}_n)^3\sum_{t=1}^{n}\sum_{h=0}^{l_n}\{u_{t+h}X_{t-1}^2 X_{t+h-1} + u_t X_{t+h-1}^2 X_{t-1}\} \\
& + (\rho - \widehat{\rho}_n)^4\sum_{t=1}^{n}\sum_{h=0}^{l_n} X_{t-1}^2 X_{t+h-1}^2 \\
=:\; & \mathcal{I}_1 + \mathcal{I}_2 + \mathcal{I}_3 + \mathcal{I}_4 + \mathcal{I}_5.
\end{aligned}$$

We claim that $\mathcal{I}_j = o_p(n^2)$ for all $j = 1, \ldots, 5$.

For $\mathcal{I}_1$, since $\sup_{t_1,t_2} E|u_{t_1}^2 u_{t_2}^2| \le C$, $\mathcal{I}_1 = O_p(nl_n) = o_p(n^2)$.

For $\mathcal{I}_2$, write $\mathcal{I}_2 =: \mathcal{I}_{2,1} + \mathcal{I}_{2,2}$. Observe that

$$\begin{aligned}\mathcal{I}_{2,1} &= \sum_{t=1}^n u_t^2 \left(\sum_{h=0}^{l_n} X_{t+h-1}u_{t+h}\right) \le \{\sum_{t=1}^n u_t^4\}^{1/2} \left\{\sum_{t=1}^n \left(\sum_{h=0}^{l_n} X_{t+h-1}u_{t+h}\right)^2\right\}^{1/2} \\ &= \{O_p(n)\}^{1/2} \left\{\sum_{t=1}^n \sum_{h=0}^{l_n} \sum_{h'=0}^{l_n} X_{t+h-1}u_{t+h}X_{t+h'-1}u_{t+h'}\right\}^{1/2}\end{aligned}$$

by Hölder's inequality. Since $E|X_{t+h-1}u_{t+h}X_{t+h'-1}u_{t+h'}| \le ||X_{t+h-1}u_{t+h}||_2||X_{t+h'-1}u_{t+h'}||_2 \le (||X_{t+h-1}^2||_2||u_{t+h}^2||_2||X_{t+h'-1}^2||_2||u_{t+h'}^2||_2)^{1/2} \le C\{(t+h-1)(t+h'-1)\}^{1/2}$ by Hölder's inequality, (A2), and Lemma A.4, it follows that $E(|\sum_{t=1}^n \sum_{h=0}^{l_n} \sum_{h'=0}^{l_n} X_{t+h-1}u_{t+h}X_{t+h'-1}u_{t+h'}|) \le C\sum_{t=1}^n \{(t+l_n)^{3/2} - t^{3/2}\}^2 = O(l_n^2 n^2)$. Thus $\mathcal{I}_{2,1} = O_p\{n^{1/2}(l_n^2 n^2)^{1/2}\} = O_p(n^{3/2}l_n) = o_p(n^2)$. Similarly, it can be shown that $\mathcal{I}_{2,2} = o_p(n^2)$, which leads to $\mathcal{I}_2 = o_p(n^2)$. The proof for $\mathcal{I}_3$, $\mathcal{I}_4$, and $\mathcal{I}_5$ can be done using Hölder's inequality, (A2), and Lemma A.4 for all summands. Specifically, for $\mathcal{I}_3 =: \mathcal{I}_{3,1} + \mathcal{I}_{3,2} + \mathcal{I}_{3,3}$, observe that for any $t_1, t_2, t_3, t_4 \in \{1, \ldots, n\}$,

$$E|u_{t_1}u_{t_2}X_{t_3}X_{t_4}| \le ||u_{t_1}u_{t_2}||_2||X_{t_3}X_{t_4}||_2 \le \{E(u_{t_1}^4)E(u_{t_2}^4)E(X_{t_3}^4)E(X_{t_4}^4)\}^{1/4} \le Cn.$$

Thus $\mathcal{I}_3 = O_p(n^{-2}nnl_n) = o_p(n)$. For $\mathcal{I}_4 =: \mathcal{I}_{4,1} + \mathcal{I}_{4,2}$, observe that for any $t_1, t_2, t_3 \in \{1, \ldots, n\}$,

$$\begin{aligned}E|u_{t_1}X_{t_2}^2X_{t_3}| &\le ||u_{t_1}||_4||X_{t_2}^2X_{t_3}||_{4/3} \le C\{E(X_{t_2}^{8/3}X_{t_3}^{4/3})\}^{3/4} \le C(||X_{t_2}^{8/3}||_{3/2}||X_{t_3}^{4/3}||_3)^{3/4} \\ &\le C\{(EX_{t_2}^4)^{2/3}(EX_{t_3}^4)^{1/3}\}^{3/4} = C(EX_{t_2}^4)^{1/2}(EX_{t_3}^4)^{1/4} \le Ct_2t_3 \le Cn^2.\end{aligned}$$

Thus $\mathcal{I}_4 = O_p(n^{-3}n^2nl_n) = o_p(n)$. For $\mathcal{I}_5$, notice that $E(X_{t-1}^2X_{t+h-1}^2) \le ||X_{t-1}^2||_2||X_{t+h-1}^2||_2 = \{E(X_{t-1}^4)E(X_{t+h-1}^4)\}^{1/2} \le C(t-1)(t+h-1) \le Cn^2$. Thus $\mathcal{I}_5 = O_p(n^{-4}n^2nl_n) = o_p(n^2)$, which completes the proof. $\diamondsuit$

B The Choice of l and the Minimum Volatility Method

In this section, we shall investigate the effect of the choice of l on the finite sample behavior of the DWB and RDWB methods. A data-driven approach, the minimum volatility (MV) method, is first proposed. The deterministic choice of l, as suggested in Section 4 of the paper, is compared to the MV method.

The idea behind the MV method is similar in spirit to that in Politis et al. (1999). The rationale behind the MV method is that the approximation of the limiting distribution should be stable if the bandwidth parameter l is in an appropriate range. We shall propose the following MV algorithm in the context of finding the optimal bandwidth parameter for the DWB method.

Algorithm **B.1**. [The Minimum Volatility (MV) Method]

1. Choose some candidates $l_1, \ldots, l_k$.
2. For each l_i $(i = 1, \ldots, k)$, generate the bootstrap sample $y_{t,n}^{*(i)}$ $(t = 1, \ldots, n)$ and calculate $\mathbf{T}_n^{(1,i)}$
3. Repeat B times so that we have $(\mathbf{T}_n^{*(1,i)}, \ldots, \mathbf{T}_n^{*(B,i)})$ for each l_i.
4. Let D_i be the empirical distribution function of $(\mathbf{T}_n^{*(1,i)}, \ldots, \mathbf{T}_n^{*(B,i)})$, i.e., $D_i(x) = B^{-1} \sum_{b=1}^{B} \mathbf{1}(\mathbf{T}_n^{*(b,i)} \leq x)$. For $i = 1, \ldots, k-1$, calculate the Kolmogorov-Smirnov distance between D_i and D_{i+1}, $H_i = \sup_{x \in \mathbb{R}} |D_i(x) - D_{i+1}(x)|$.
5. The optimal l is $l_{\widehat{i}}$, where $\widehat{i} = \text{argmin}_{i=1,\ldots,k-1} H_i$.

The MV procedure above is described for the $\mathbf{T}_n$ statistics and DWB for simplicity. The same method can be applied to $\mathbf{t}_n$ and RDWB as well. Note that the MV choice of l depends on the data $\{X_{t,n}\}$. Tables B.1 and B.2 present the details of how the choice of l affects the empirical size, along with the average of the chosen l for selected DGPs. Here, the candidates are $l = 1, \ldots, \lfloor 12(n/100)^{1/4} \rfloor$. Thus, the maximum value of l that is considered equals 13 if $n = 100$, and 17 if $n = 400$. Although the MV method may not necessarily choose a theoretically optimal l, it seems to provide a reasonable practical guidance as long as the range of the candidates for l is appropriate.

On the other hand, the MV method is computationally costly, with the computational time proportional to the number of candidate bandwidths we include and the number of bootstrap replications. Tables B.1 and B.2 indicate that the empirical rejection rates for DWB and RDWB are not too sensitive to the choice of l, as long as l is not too small. We propose to use the middle value, $l = \lfloor 6(n/100)^{1/4} \rfloor$, as a computationally efficient practical alternative. Table B.3 further compares this deterministic choice with the MV method for RDWB, which is recommended in the paper for its finite sample performance. It seems that the two choices of l are comparable in almost all DGPs for RDWB. This behavior is observed not just for the size but also for the power in our unreported simulations. Therefore, we shall recommend RDWB with the aforementioned deterministic choice of l.

C Power Curves for All DGPs

Figures C.1-C.4 present all power curves for the DWB, RWB, and RDWB methods and for $\mathbf{T}_n$ and $\mathbf{t}_n$ statistics.

References

Chang, Y. and J. Y. Park (2003). A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis 24*(4), 379–400.

Kurtz, T. G. (2001). *Lectures on Stochastic Analysis*. University of Wisconsin-Madison: available at http://www.math.wisc.edu/ kurtz/735/main735.pdf.

Paparoditis, E. and D. N. Politis (2003). Residual-based block bootstrap for unit root testing. *Econometrica 71*(3), 813–855.

Phillips, P. C. B. (1987). Towards a unified asymptotic theory for autoregression. *Biometrika 74*(3), 535–547.

Politis, D. N., J. P. Romano, and M. Wolf (1999). *Subsampling*. New York: Springer.

Rosenblatt, M. (1985). *Stationary Squences and Random Fields*. Birkhũser.

Shao, Q.-M. and H. Yu (1996). Weak convergence for weighted empirical processes of dependent sequences. *The Annals of Probability 24*(4), 2098–2127.

Shao, X. (2010). The dependent wild bootstrap. *Journal of the American Statistical Association 105*(489), 218–235.

Shao, X. and W. B. Wu (2007). Asymptotic spectral theory for nonlinear time series. *The Annals of Statistics 35*(4), 1773–1801.

Wu, W. B. (2005). Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America 102*(40), 14150–14154.

Wu, W. B. and X. Shao (2004). Limit theorems for iterated random functions. *Journal of Applied Probability 41*(2), 425–436.

Zhou, Z. (2013). Heteroscedasticity and autocorrelation robust structural change detection. *Journal of the American Statistical Association 108*(502), 726–740.

Table B.1: Empirical sizes for DWB and RDWB with various choices of l for selected DGPs, matching with the four panels in Figures 1 and 2. The table represents MA models with 2000 Monte-Carlo replications, 1000 Bootstrap replications, and $\rho = 1$. The l_{MV} rows indicate the average of optimal l chosen by the MV method. The MV rows indicates the empirical sizes with DWB and RDWB using the MB method. The nominal level is 5%.

		$(MA_{4,1})$				$(MA_{2,1})$				$(MA_{1,3})$				$(MA_{6,3})$			
		DWB		RDWB		DWB		RDWB		DWB		RDWB		DWB		RDWB	
n	l	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$
	1	1.5	1.5	4.4	4.3	93.5	93.2	20.5	20.4	0.9	0.9	4.2	4.9	43.9	40.5	12.8	12.6
	2	3.3	3.2	5.1	4.7	84.9	84.7	20.6	20.9	2.9	3.2	4.8	4.9	34.0	32.5	12.3	12.0
	3	3.5	3.6	4.3	4.4	81.2	81.0	19.9	20.2	4.0	4.2	4.4	5.0	33.6	31.6	12.2	11.8
	4	4.0	3.7	4.9	4.5	81.0	80.8	20.0	20.0	4.5	4.7	4.6	5.2	33.9	32.0	11.9	11.8
	5	3.9	3.9	4.7	4.5	81.0	81.2	20.0	20.1	4.4	4.5	4.8	5.3	34.5	33.4	12.2	12.0
	6	4.0	4.0	5.0	4.9	82.2	81.9	19.8	20.0	4.3	4.8	4.9	5.3	35.4	33.7	12.1	12.2
	7	3.8	4.0	4.9	5.0	82.7	82.9	19.1	19.5	4.6	4.8	5.1	5.4	36.6	34.8	12.4	12.5
100	8	3.8	4.0	4.9	4.8	83.9	83.8	19.1	19.4	4.5	4.5	5.3	5.5	37.6	35.6	12.5	12.4
	9	3.9	4.0	5.0	5.1	84.9	85.0	19.2	19.4	4.0	4.5	5.5	5.8	38.7	36.3	12.8	13.0
	10	3.9	4.2	5.1	5.0	85.2	85.0	19.2	19.2	4.2	4.4	5.5	5.8	39.8	38.0	13.0	13.2
	11	3.8	4.0	5.1	5.5	86.1	86.1	19.5	19.8	4.2	4.2	5.6	5.8	40.9	38.5	13.1	13.2
	12	3.8	3.8	5.5	5.2	86.9	87.0	19.1	19.2	4.3	4.3	5.6	5.9	41.8	39.6	13.2	13.3
	13	3.6	4.2	5.5	5.2	87.3	87.4	20.0	19.9	4.0	4.2	5.8	5.7	42.6	40.8	13.4	13.2
	MV	4.0	4.2	5.0	4.8	83.5	83.7	19.8	19.9	4.4	4.6	5.1	5.6	37.6	36.0	12.3	12.2
	l_{MV}	7.3	7.3	6.9	6.5	7.5	7.4	6.4	6.3	8.0	7.8	7.2	7.0	8.1	7.9	7.3	7.2
	1	0.8	1.0	4.9	4.8	97.2	97.0	12.8	12.8	0.5	0.6	5.3	5.5	44.9	41.9	16.2	15.3
	2	2.4	2.6	4.7	4.9	89.0	88.7	13.1	13.1	2.5	2.7	5.2	5.4	31.2	29.7	14.4	13.5
	3	3.1	3.5	4.5	4.6	82.8	82.2	12.7	12.7	3.2	3.4	5.3	5.3	26.6	25.8	13.2	12.4
	4	3.8	3.8	4.5	4.8	79.7	79.6	12.8	12.8	3.6	3.7	5.5	5.2	25.4	24.6	12.1	11.8
	5	3.8	3.9	4.7	4.7	78.0	78.2	12.1	12.2	3.8	4.2	5.4	5.7	25.0	24.2	11.9	11.5
	6	3.8	4.0	4.2	4.4	77.8	77.9	12.3	12.4	4.0	4.0	5.5	5.9	25.2	24.8	11.5	10.9
	7	3.7	4.0	4.6	4.7	78.0	77.9	12.4	12.2	4.0	4.5	5.8	5.8	25.7	24.7	11.6	11.2
	8	3.6	4.0	4.6	4.7	78.5	78.4	12.2	12.0	4.5	4.3	5.9	5.9	25.9	25.5	11.6	11.1
	9	3.8	4.1	4.6	4.7	78.7	78.6	11.8	11.8	4.5	4.5	5.9	5.9	26.8	26.2	11.3	11.0
400	10	3.8	4.0	4.7	4.8	78.8	79.0	11.7	11.7	4.3	4.7	6.3	6.2	27.0	26.5	11.9	11.5
	11	4.0	4.0	4.7	4.6	79.6	79.5	11.6	11.6	4.3	4.9	6.2	6.3	27.8	27.3	12.3	11.7
	12	3.8	3.9	4.6	4.7	80.5	80.2	11.5	11.5	4.5	4.6	6.3	6.4	28.2	27.8	12.0	11.3
	13	3.9	4.3	4.7	4.8	81.0	80.9	11.3	11.5	4.8	4.7	6.4	6.3	28.8	28.2	12.2	11.7
	14	3.7	4.0	4.4	4.8	81.3	81.3	11.2	11.3	4.6	5.0	6.7	6.3	30.0	29.3	12.2	11.8
	15	3.8	4.1	4.8	4.9	82.1	82.2	11.3	11.3	4.6	4.9	6.7	6.6	30.3	29.1	12.6	11.9
	16	3.8	4.3	4.8	4.8	82.7	82.7	11.0	11.0	4.5	4.8	6.7	6.8	30.7	30.2	12.8	12.2
	17	3.6	4.0	4.8	4.5	83.2	83.0	11.0	10.8	4.4	4.8	6.8	6.9	31.8	30.6	12.5	12.2
	MV	3.5	3.7	4.5	4.6	79.5	79.2	11.8	11.8	4.2	4.3	6.2	5.9	27.1	26.8	12.0	11.8
	l_{MV}	9.1	9.1	8.5	8.2	9.7	9.6	8.4	8.2	9.6	9.6	8.8	8.6	10.0	9.8	9.2	9.0

Table B.2: Empirical sizes for DWB and RDWB with various choices of l for selected DGPs, matching with the four panels in Figures 1 and 2. The table represents AR models with 2000 Monte-Carlo replications, 1000 Bootstrap replications, and $\rho = 1$. The l_{MV} rows indicate the average of optimal l chosen by the MV method. The MV rows indicates the empirical sizes with DWB and RDWB using the MB method. The nominal level is 5%.

		$(AR_{4,1})$				$(AR_{2,1})$				$(AR_{1,3})$				$(AR_{6,3})$			
		DWB		RDWB		DWB		RDWB		DWB		RDWB		DWB		RDWB	
n	l	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$	$\mathbf{T}_n$	$\mathbf{t}_n$
	1	0.2	0.2	3.5	3.5	63.6	62.4	7.8	7.6	0.0	0.0	5.5	6.9	31.6	29.7	9.4	9.3
	2	0.6	0.5	3.5	3.5	34.9	35.0	7.7	7.7	0.0	0.2	5.7	7.0	20.7	19.4	8.7	8.6
	3	1.1	1.1	3.6	3.5	43.8	43.9	7.6	7.6	0.2	0.3	5.5	7.0	23.1	22.1	9.0	8.9
	4	1.5	1.5	3.5	3.4	40.2	40.1	7.3	7.4	0.2	0.4	5.6	7.1	23.4	22.4	9.2	9.0
	5	1.5	1.5	3.6	3.5	44.6	44.5	7.6	7.7	0.5	0.6	5.8	7.0	24.8	24.1	9.2	9.3
	6	1.7	1.7	3.6	3.7	45.1	45.0	7.5	7.7	0.5	0.6	5.9	7.2	26.1	25.4	9.3	9.3
	7	1.7	1.8	3.5	3.8	47.0	47.2	7.6	7.9	0.5	0.7	5.9	7.4	27.5	26.0	9.9	9.6
100	8	1.8	1.9	3.8	4.0	48.1	48.3	8.0	8.1	0.6	0.6	6.0	7.3	27.7	26.8	10.0	9.8
	9	1.8	2.0	4.2	4.1	49.4	49.5	8.2	8.2	0.5	0.6	5.8	7.0	29.3	27.9	10.0	9.8
	10	1.7	1.9	3.8	4.2	50.8	50.6	7.9	8.1	0.8	0.7	5.7	7.3	30.1	28.8	10.3	10.0
	11	1.7	1.7	4.0	4.0	52.0	51.7	8.2	8.3	0.8	0.6	5.9	7.3	30.6	29.0	10.2	10.0
	12	1.8	1.6	4.0	4.0	52.9	52.8	8.2	8.3	0.8	0.8	6.0	7.4	31.3	29.6	10.4	10.1
	13	1.8	1.5	3.9	4.1	53.4	53.5	8.6	8.6	0.8	0.7	6.2	7.6	31.8	30.8	10.8	10.4
MV		1.7	1.8	3.8	3.6	48.9	49.1	7.9	8.0	0.8	0.6	5.9	7.3	27.6	27.0	9.8	9.4
l_{MV}		8.0	8.0	7.3	6.7	8.7	8.5	6.9	6.7	8.6	8.4	7.0	6.9	8.3	8.0	7.2	7.2
	1	0.0	0.0	6.5	5.7	65.9	65.2	5.3	5.3	0.0	0.0	3.8	4.6	31.1	29.2	10.9	10.1
	2	0.5	0.4	5.8	5.3	24.9	24.9	5.5	5.7	0.0	0.0	3.8	4.7	14.1	13.8	8.9	8.3
	3	0.7	0.8	5.8	5.3	37.8	37.5	5.4	5.3	0.0	0.0	4.0	4.8	16.2	15.7	9.0	8.4
	4	0.8	0.9	5.3	5.0	28.3	28.3	5.5	5.5	0.0	0.2	4.4	4.9	14.8	14.8	8.6	8.2
	5	1.5	1.4	5.2	4.9	33.5	33.2	5.5	5.5	0.2	0.3	4.2	4.9	16.1	15.8	9.0	8.8
	6	1.5	1.6	5.0	5.1	30.9	31.0	5.5	5.5	0.3	0.6	4.3	5.0	16.3	16.2	8.8	8.4
	7	1.8	1.8	4.8	4.8	33.4	33.1	5.6	5.5	0.4	0.7	4.2	4.9	16.5	16.4	9.1	8.5
	8	2.1	2.0	5.1	4.8	32.9	33.0	5.7	5.5	0.7	1.1	4.3	5.1	17.3	16.8	8.8	8.6
	9	2.1	2.2	4.9	4.8	34.4	34.3	5.5	5.5	1.0	1.2	4.3	4.9	17.9	17.5	9.3	8.8
400	10	2.4	2.3	5.1	4.9	34.8	34.8	5.3	5.4	1.1	1.4	4.3	5.2	18.4	18.1	9.3	8.9
	11	2.3	2.4	5.1	5.1	35.9	35.9	5.8	5.6	1.2	1.6	4.5	5.4	18.9	18.6	9.3	8.9
	12	2.5	2.5	5.1	4.9	36.7	36.6	5.9	5.9	1.5	1.7	4.8	5.1	19.5	19.1	9.9	9.3
	13	2.6	2.4	5.1	5.0	37.1	37.3	5.8	5.8	1.8	1.8	4.7	5.4	20.4	19.9	9.8	9.2
	14	2.5	2.5	5.0	4.9	38.0	37.9	5.5	5.6	1.7	2.1	4.6	5.3	20.9	20.1	10.1	9.3
	15	2.5	2.5	5.1	4.9	38.6	38.6	5.7	5.7	1.8	2.0	4.8	5.5	21.3	20.7	10.0	9.2
	16	2.5	2.4	5.3	4.9	39.1	39.5	5.9	5.9	1.9	2.1	4.8	5.6	21.8	21.4	10.4	9.8
	17	2.5	2.6	5.0	5.2	40.3	40.0	5.8	5.8	2.1	2.1	4.8	5.7	22.4	21.8	10.4	9.8
MV		2.1	2.1	5.2	4.8	36.0	35.4	5.5	5.5	1.2	1.4	4.4	4.9	18.6	18.3	9.8	8.9
l_{MV}		10.2	10.3	9.6	8.8	11.5	11.5	8.4	8.3	10.9	10.9	8.6	8.6	10.1	9.8	9.2	9.0

Table B.3: Empirical Sizes for RDWB with l chosen by the MV method and the deterministic choice (DC) $l = \lfloor 6(n/100)^{1/4} \rfloor$, based on 2000 Monte-Carlo replications and 1000 Bootstrap replications under $\rho = 1$ for all ($\mathrm{MA}_{i,j}$) and ($\mathrm{AR}_{i,j}$) models. The nominal level is 5%.

		MA models								AR models							
		$n = 100$				$n = 400$				$n = 100$				$n = 400$			
		$\mathbf{T}_n$		$\mathbf{t}_n$		$\mathbf{T}_n$		$\mathbf{t}_n$		$\mathbf{T}_n$		$\mathbf{t}_n$		$\mathbf{T}_n$		$\mathbf{t}_n$	
i	j	DC	MV	DC	MV	DC	MV	DC	MV	DC	MV	DC	MV	DC	MV	DC	MV
	1	4.7	4.9	4.7	4.7	4.5	5.1	4.4	4.9	4.0	3.6	4.2	3.8	3.7	3.6	3.9	3.6
	2	5.0	4.9	4.9	5.0	4.5	4.5	4.5	4.4	4.6	4.6	4.7	5.0	4.0	3.8	4.1	4.0
1	3	4.9	5.1	5.3	5.6	5.9	6.2	5.9	5.9	5.9	5.9	7.2	7.3	4.3	4.4	5.1	4.9
	4	5.3	5.2	5.3	5.4	5.9	6.7	6.0	6.2	2.8	2.9	2.9	3.2	5.6	5.9	5.6	5.8
	5	4.1	4.4	4.0	4.2	4.9	4.7	5.0	5.0	3.5	3.7	3.8	4.0	4.1	4.2	4.0	4.0
	1	19.8	19.8	20.0	19.9	12.2	11.8	12.0	11.8	7.5	7.9	7.7	8.0	5.7	5.5	5.5	5.5
	2	22.3	22.3	22.5	22.4	10.6	10.2	10.7	10.2	7.3	7.8	7.2	8.0	5.5	5.3	5.5	5.3
2	3	21.4	22.0	20.6	21.3	12.4	13.0	12.4	12.8	11.0	11.6	10.6	11.2	9.8	9.7	9.8	9.4
	4	18.4	18.9	18.3	18.8	11.1	11.2	11.2	11.3	6.6	7.6	6.7	7.4	6.8	7.2	6.8	7.2
	5	23.4	23.4	23.5	23.6	12.8	13.3	12.8	13.4	9.0	9.2	9.0	9.2	6.0	6.2	6.0	6.4
	1	5.0	5.2	5.0	5.1	4.7	4.5	4.5	4.5	4.0	4.3	4.3	4.8	4.8	4.9	4.9	5.1
	2	5.2	5.5	5.0	5.2	4.8	5.0	5.1	4.9	4.1	4.0	4.1	4.6	3.5	3.6	3.6	3.6
3	3	6.1	6.2	6.3	6.6	7.4	7.4	7.1	7.6	6.8	6.9	8.9	8.9	4.9	5.0	5.7	5.7
	4	5.5	5.9	5.8	6.6	6.2	6.5	6.3	6.5	3.3	3.8	3.5	3.6	5.5	5.7	5.3	5.5
	5	6.0	5.7	5.8	5.9	5.2	5.9	5.2	5.2	3.5	3.4	4.0	3.9	4.9	4.8	4.7	4.6
	1	5.0	5.0	4.9	4.8	4.6	4.5	4.7	4.6	3.6	3.8	3.7	3.6	5.1	5.2	4.8	4.8
	2	5.2	5.2	5.0	5.2	4.0	4.2	4.2	4.2	3.9	4.0	3.9	4.2	5.4	5.1	5.2	5.6
4	3	5.3	5.6	6.6	6.5	7.7	8.4	7.5	7.4	6.0	6.2	11.3	11.7	7.4	7.5	8.6	8.8
	4	4.8	4.5	4.9	4.8	6.2	6.2	6.2	6.2	3.6	3.6	3.6	3.5	5.4	5.8	5.1	5.2
	5	4.7	4.9	4.2	4.3	4.9	4.7	4.7	4.8	3.6	4.0	3.8	4.0	5.4	5.8	5.1	5.5
	1	6.2	6.2	6.4	6.6	5.7	5.7	5.2	5.8	5.2	5.5	5.3	5.5	5.5	5.5	5.0	5.5
	2	5.1	4.9	5.2	5.0	6.2	6.5	6.5	6.7	3.9	3.8	3.7	3.8	6.0	6.5	6.3	6.7
5	3	7.3	7.4	9.2	9.2	10.1	10.3	10.0	10.3	6.9	7.0	14.1	13.9	9.2	9.6	10.5	10.9
	4	5.5	5.4	5.5	5.3	5.7	6.0	5.9	6.0	4.5	4.9	4.2	4.7	5.8	6.0	5.5	5.5
	5	5.3	5.9	5.5	5.5	5.1	5.1	5.1	5.3	4.2	4.5	4.3	4.5	6.2	6.4	6.4	6.6
	1	10.0	10.4	9.5	9.8	8.0	8.5	8.1	8.2	6.3	6.2	6.2	6.3	5.5	5.8	5.5	5.7
	2	9.8	9.6	9.9	9.5	6.8	7.2	7.0	7.0	6.5	6.6	6.5	6.5	4.8	4.6	4.8	4.8
6	3	12.1	12.3	12.2	12.2	11.6	12.0	11.1	11.8	9.3	9.8	9.3	9.4	8.8	9.8	8.6	8.9
	4	8.6	8.4	8.6	8.9	6.7	6.6	6.7	6.6	6.4	6.8	6.4	6.7	5.9	5.8	5.8	5.9
	5	9.8	10.0	9.7	10.1	6.7	6.9	6.7	7.0	6.2	6.3	6.5	6.3	5.1	5.0	5.0	5.0

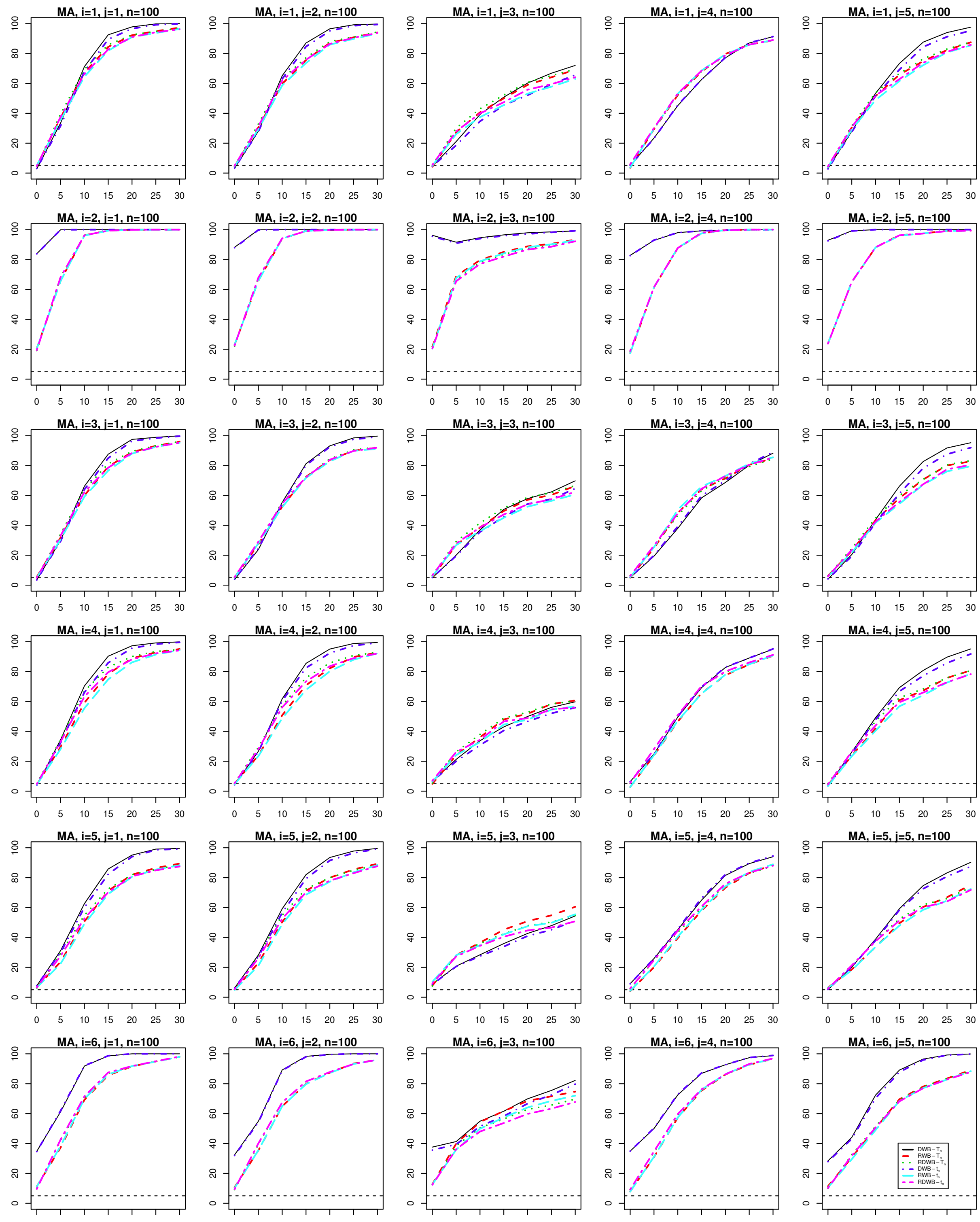

Figure C.1: Rejection frequencies (%) versus $-c$, where $\rho = 1 + c/n$ for DWB, RWB, and RDWB unit root tests in MA models. The sample size is $n = 100$ and the nominal level is 5%.

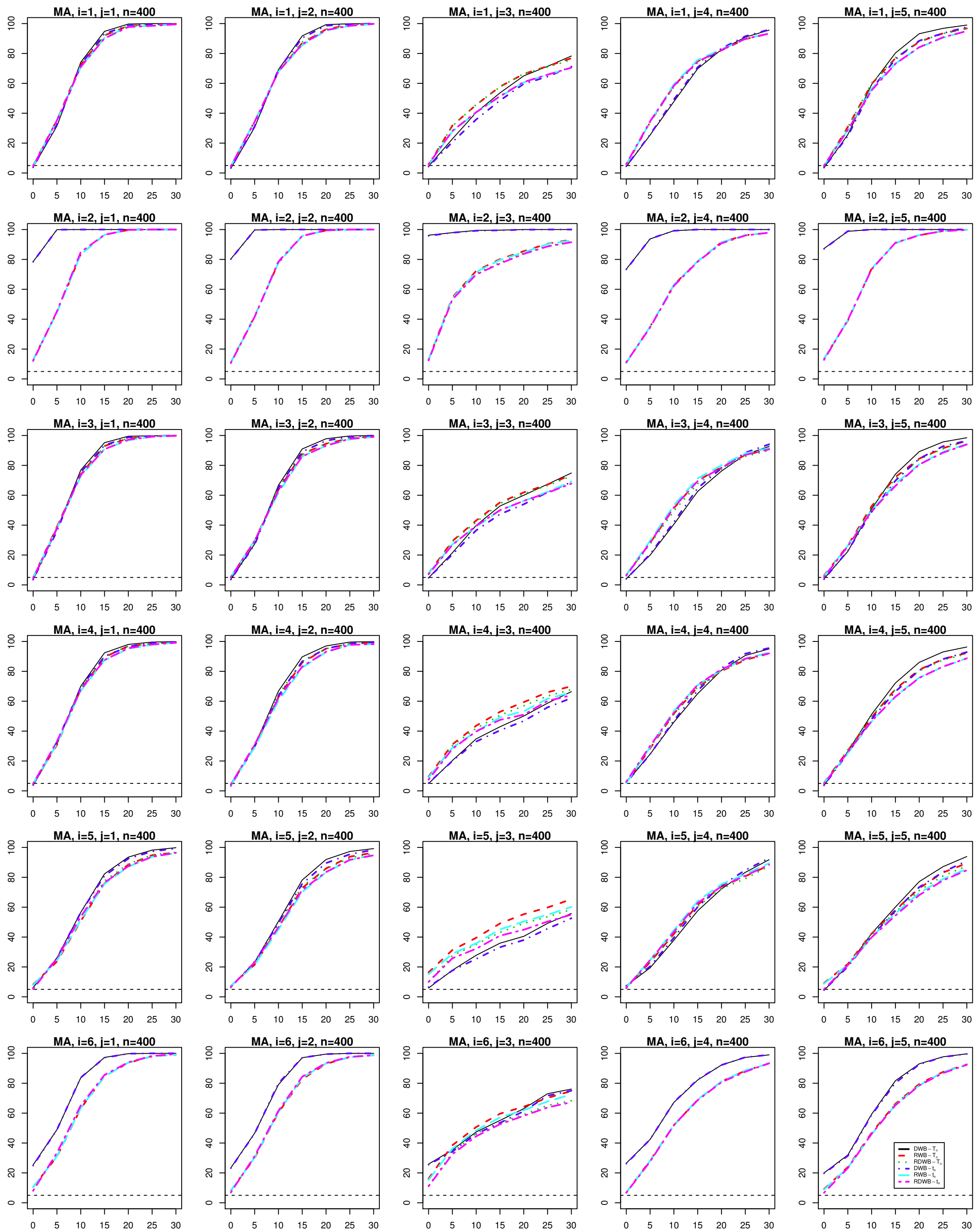

Figure C.2: Rejection frequencies (%) versus $-c$, where $\rho = 1 + c/n$ for DWB, RWB, and RDWB unit root tests in MA models. The sample size is $n = 400$ and the nominal level is 5%.

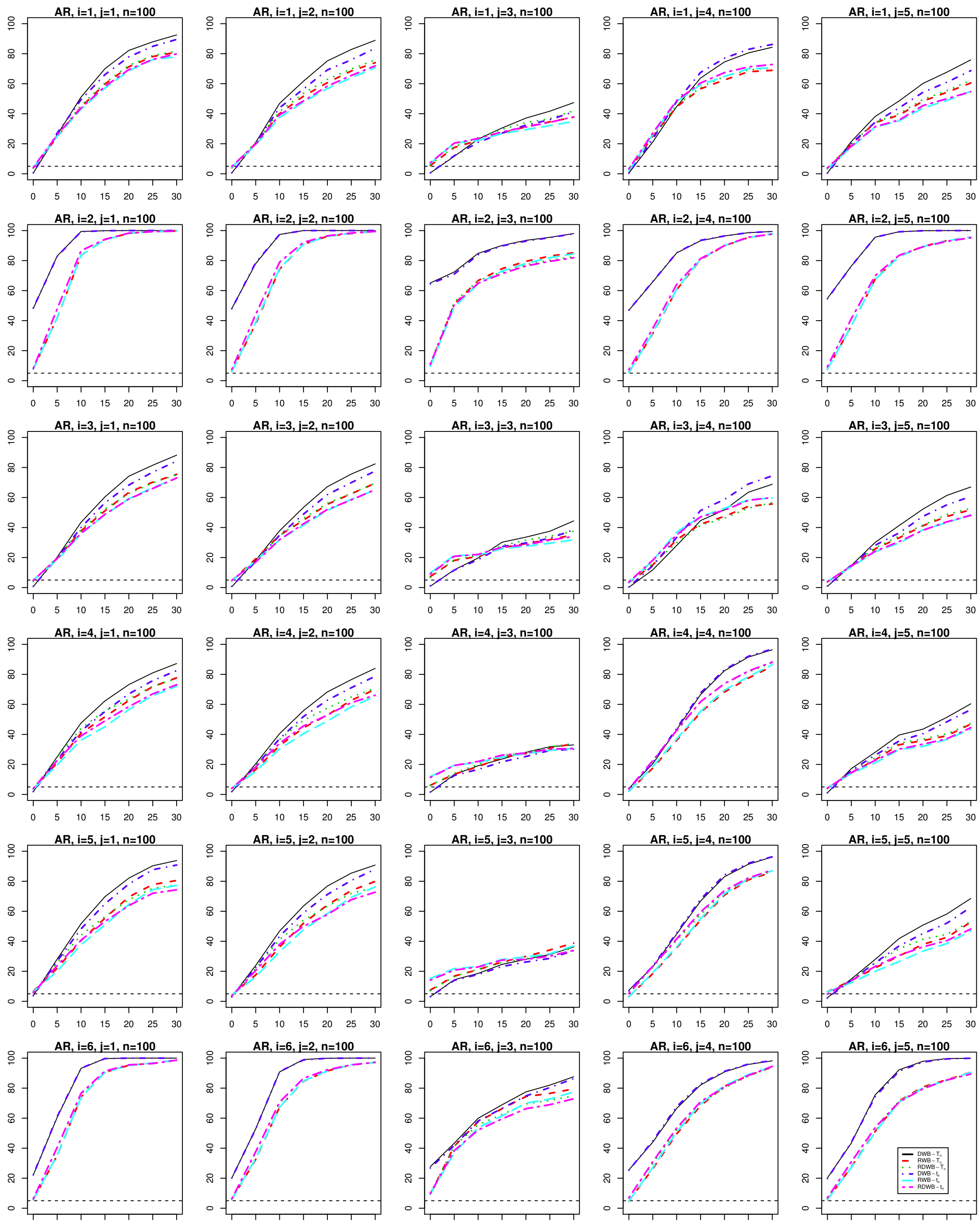

Figure C.3: Rejection frequencies (%) versus $-c$, where $\rho = 1 + c/n$ for DWB, RWB, and RDWB unit root tests in AR models. The sample size is $n = 100$ and the nominal level is 5%.

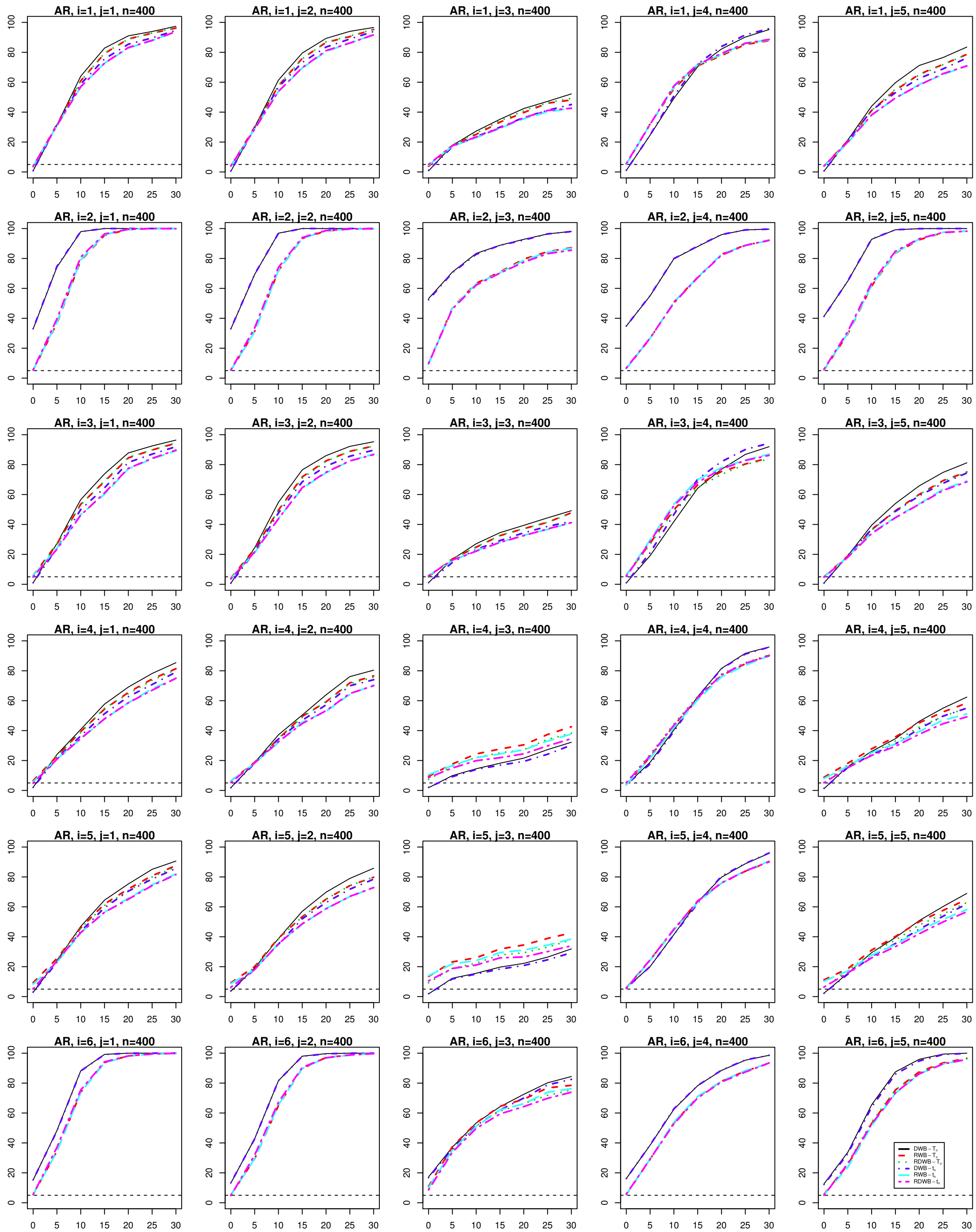

Figure C.4: Rejection frequencies (%) versus $-c$, where $\rho = 1 + c/n$ for DWB, RWB, and RDWB unit root tests in AR models. The sample size is $n = 400$ and the nominal level is 5%.