Online Supplement to "Bubble Testing under Polynomial Trends" by Wang and Yu

Proof of Theorem 3.1.1 when $m > 2$

For the limit of the space, the main paper proves the results in Theorem 3.1.1 for the case where $m = 2$. We now complete the proof for the case where $m > 2$.

In any subsample period from $n_1 = \lfloor nr_1 \rfloor \in T$ to $n_2 = \lfloor nr_2 \rfloor \in T$, $p_t = p_{n_c} + \delta (t - n_c)^m + \varepsilon_t$ are generated by a polynomial trend model. Without loss of generality, we give the proof for the case where $n_1 = n_c = \lfloor nr_c \rfloor$. The same approach can be applied to the case where $n_1 > n_c$.

Let $n_s = n_2 - n_1$ and $\sum = \sum_{n_1+3}^{n_2}$. Define $g(t) = \delta (t - n_c)^m$ with $m > 2$. For any $t \geq n_1 + 3$, it has

$$\Delta g(t) = \delta m (t - n_c - 1)^{m-1} + \delta \frac{m(m-1)}{2} (t - n_c - 1)^{m-2} + O\left((t - n_c)^{m-3}\right),$$

and

$$\Delta^2 g(t) = \Delta g(t) - \Delta g(t-1) = \delta m (m-1)(t - n_c - 2)^{m-2} + O\left((t - n_c)^{m-3}\right).$$

We can further get

$$\Delta^2 p_t = \Delta^2 g(t) + \Delta^2 \varepsilon_t,$$

which leads to

$$p_t = p_{t-1} + \Delta p_{t-1} + \omega_t \quad \text{with} \quad \omega_t = \Delta^2 g(t) + \Delta^2 \varepsilon_t.$$

Hence, the centered LS estimators of the parameters in the AR(2) regression as the Model (1.2) in the main paper take the form of

$$\begin{bmatrix} \check{\alpha}_{r_1}^{r_2} \\ \check{\beta}_{r_1}^{r_2} - 1 \\ \check{\psi}_{r_1}^{r_2} - 1 \end{bmatrix} = \left(\sum \begin{bmatrix} 1 & p_{t-1} & \Delta p_{t-1} \\ p_{t-1} & p_{t-1}^2 & p_{t-1}\Delta p_{t-1} \\ \Delta p_{t-1} & p_{t-1}\Delta p_{t-1} & (\Delta p_{t-1})^2 \end{bmatrix} \right)^{-1} \begin{bmatrix} \sum \omega_t \\ \sum p_{t-1}\omega_t \\ \sum \Delta p_{t-1}\omega_t \end{bmatrix}.$$

Since $m > 2$, as $n \to \infty$, it has

$$n_s^{-(m-1)} \sum \omega_t = n_s^{-(m-1)} \sum \Delta^2 g(t) + o_p(1) \xrightarrow{p} \delta m,$$

$$\begin{aligned} & n_s^{-(2m-1)} \sum p_{t-1}\omega_t \\ = \; & n_s^{-(2m-1)} \delta^2 m (m-1) \sum (t - n_c - 1)^m (t - n_c - 2)^{m-2} + o_p(1) \\ \xrightarrow{p} \; & \frac{m(m-1)}{2m-1} \delta^2, \end{aligned}$$

$$\begin{aligned} n_s^{-2(m-1)} \sum \Delta p_{t-1}\omega_t &= n_s^{-2(m-1)} \delta^2 m^2 (m-1) \sum (t - n_c - 2)^{2m-3} + o_p(1) \\ &\xrightarrow{p} \frac{m^2}{2}\delta^2. \end{aligned}$$

Therefore, we get

$$\begin{pmatrix} n_s^{-(m-1)} & 0 & 0 \\ 0 & n_s^{-(2m-1)} & 0 \\ 0 & 0 & n_s^{-2(m-1)} \end{pmatrix} \begin{bmatrix} \sum \omega_t \\ \sum p_{t-1}\omega_t \\ \sum \Delta p_{t-1}\omega_t \end{bmatrix} \xrightarrow{p} \begin{bmatrix} \delta m \\ \frac{m(m-1)}{2m-1}\delta^2 \\ \frac{m^2}{2}\delta^2 \end{bmatrix}.$$

From $p_t = p_{n_c} + \delta\left(t - n_c\right)^m + \varepsilon_t$, it can be easily proved that, as $n \to \infty$,

$$n_s^{-(m+1)} \sum p_{t-1} \xrightarrow{p} \delta/\left(m+1\right), \quad n_s^{-(2m+1)} \sum p_{t-1}^2 \xrightarrow{p} \delta^2/\left(2m+1\right),$$

$$n_s^{-m} \sum \Delta p_{t-1} \xrightarrow{p} \delta, \quad n_s^{-(2m-1)} \sum \left(\Delta p_{t-1}\right)^2 \xrightarrow{p} \delta^2 m^2/\left(2m-1\right),$$

$$\begin{aligned} & n_s^{-2m} \sum p_{t-1}\Delta p_{t-1} \\ = \; & n_s^{-2m}\delta^2 m \sum \left(t - n_c - 1\right)^m \left(t - n_c - 2\right)^{m-1} + o_p\left(1\right) \xrightarrow{p} \delta^2/2. \end{aligned}$$

With the normalization matrices

$$C = \begin{pmatrix} n_s^{-m+2} & 0 & 0 \\ 0 & n_s^2 & 0 \\ 0 & 0 & n_s \end{pmatrix} \quad \text{and} \quad G = \begin{pmatrix} n_s^{(m-1)} & 0 & 0 \\ 0 & n_s^{(2m-1)} & 0 \\ 0 & 0 & n_s^{2(m-1)} \end{pmatrix},$$

the following limiting results can be proved:

$$\begin{aligned} & C \begin{bmatrix} \sum 1 & \sum p_{t-1} & \sum \Delta p_{t-1} \\ \sum p_{t-1} & \sum p_{t-1}^2 & \sum p_{t-1}\Delta p_{t-1} \\ \sum \Delta p_{t-1} & \sum_{t=3}^n p_{t-1}\Delta p_{t-1} & \sum \left(\Delta p_{t-1}\right)^2 \end{bmatrix}^{-1} G \\ = \; & \begin{bmatrix} n_s^{-1}\sum 1 & n_s^{-m-1}\sum p_{t-1} & n_s^{-m}\sum \Delta p_{t-1} \\ n_s^{-m-1}\sum p_{t-1} & n_s^{-2m-1}\sum p_{t-1}^2 & n_s^{-2m}\sum p_{t-1}\Delta p_{t-1} \\ n_s^{-m}\sum \Delta p_{t-1} & n_s^{-2m}\sum p_{t-1}\Delta p_{t-1} & n_s^{-2m+1}\sum \left(\Delta p_{t-1}\right)^2 \end{bmatrix}^{-1} \\ \xrightarrow{p} \; & \begin{bmatrix} 1 & \delta/\left(m+1\right) & \delta \\ \delta/\left(m+1\right) & \delta^2/\left(2m+1\right) & \delta^2/2 \\ \delta & \delta^2/2 & \delta^2 m^2/\left(2m-1\right) \end{bmatrix}^{-1}. \end{aligned}$$

Consequently, it is obtained that

$$\begin{pmatrix} n_s^{-m+2} & 0 & 0 \\ 0 & n_s^2 & 0 \\ 0 & 0 & n_s \end{pmatrix} \begin{bmatrix} \breve{\alpha}_{r_1}^{r_2} \\ \breve{\beta}_{r_1}^{r_2} - 1 \\ \breve{\psi}_{r_1}^{r_2} - 1 \end{bmatrix} \xrightarrow{p} \begin{bmatrix} 1 & \delta/\left(m+1\right) & \delta \\ \delta/\left(m+1\right) & \delta^2/\left(2m+1\right) & \delta^2/2 \\ \delta & \delta^2/2 & \delta^2 m^2/\left(2m-1\right) \end{bmatrix}^{-1} \begin{bmatrix} \delta m \\ \delta^2 \frac{m(m-1)}{2m-1} \\ \delta^2 m^2/2 \end{bmatrix},$$

which leads to

$$n_s^2\left(\breve{\beta}_{r_1}^{r_2} - 1\right) \xrightarrow{p} -\frac{m\left(2m+1\right)\left(m+1\right)\left(m-2\right)}{\left(2m-1\right)\left(m-1\right)} \quad \text{as} \quad n \to \infty.$$

Hence, $n_s\left(\breve{\beta}_{r_1}^{r_2} - 1\right) = o_p\left(1\right) \xrightarrow{p} 0$, as $n \to \infty$.

Next, we develop the limit of the DF t statistic $t\left(\check{\beta}_{r_1}^{r_2}\right)$. Note that $\check{e}_t = p_t - \check{\alpha}_{r_1}^{r_2} - \check{\beta}_{r_1}^{r_2} p_{t-1} - \check{\psi}_{r_1}^{r_2} \Delta p_{t-1}$ and $p_t = p_{t-1} + \Delta p_{t-1} + \omega_t$ with $\omega_t = \Delta^2 g\left(t\right) + \Delta^2 \varepsilon_t$. Thus, the first-order condition of the LS regression leads to

$$\sum \check{e}_t^2 = \sum \check{e}_t p_t = \sum \check{e}_t \omega_t.$$

Applying the fact of $\omega_t = \delta m\left(m-1\right)\left(t-n_c-2\right)^{m-2} + O\left(\left(t-n_c\right)^{m-3}\right) + \Delta^2 \varepsilon_t$, as $n \to \infty$, it has

$$\begin{aligned}
& n_s^{-(2m-3)} \sum \check{e}_t^2 \\
= \; & n_s^{-(2m-3)} \sum \left(p_t - \check{\alpha}_{r_1}^{r_2} - \check{\beta}_{r_1}^{r_2} p_{t-1} - \check{\psi}_{r_1}^{r_2} \Delta p_{t-1}\right) \omega_t \\
= \; & n_s^{-(2m-3)} \sum \left(\omega_t - \check{\alpha}_{r_1}^{r_2} - \left(\check{\beta}_{r_1}^{r_2} - 1\right) p_{t-1} - \left(\check{\psi}_{r_1}^{r_2} - 1\right) \Delta p_{t-1}\right) \omega_t \\
= \; & n_s^{-(2m-3)} \sum \omega_t^2 - \begin{pmatrix} n_s^{-m+2} \check{\alpha}_{r_1}^{r_2} & n_s^2 \left(\check{\beta}_{r_1}^{r_2} - 1\right) & n_s \left(\check{\psi}_{r_1}^{r_2} - 1\right) \end{pmatrix} \begin{bmatrix} n_s^{-(m-1)} \sum \omega_t \\ n_s^{-(2m-1)} \sum p_{t-1} \omega_t \\ n_s^{-(2m-2)} \sum \Delta p_{t-1} \omega_t \end{bmatrix} \\
= \; & O_p\left(1\right),
\end{aligned}$$

where the last equation comes from the fact that

$$\begin{aligned}
n_s^{-(2m-3)} \sum \omega_t^2 &= n_s^{-(2m-3)} \sum \delta^2 m^2 \left(m-1\right)^2 \left(t-n_c-2\right)^{2m-4} + o_p\left(1\right) \\
&= \delta^2 \frac{m^2 \left(m-1\right)^2}{2m-3} + o_p\left(1\right),
\end{aligned}$$

and the limiting results for $\begin{pmatrix} \check{\alpha}_{r_1}^{r_2} & \left(\check{\beta}_{r_1}^{r_2} - 1\right) & \left(\check{\psi}_{r_1}^{r_2} - 1\right) \end{pmatrix}$ developed earlier. Consequently, it is obtained that

$$\begin{aligned}
& n_s^5 \left[se\left(\check{\beta}_{r_1}^{r_2}\right)\right]^2 \\
= \; & \begin{pmatrix} 0 & n_s^2 & 0 \end{pmatrix} \left(\sum \begin{bmatrix} 1 & p_{t-1} & \Delta p_{t-1} \\ p_{t-1} & p_{t-1}^2 & p_{t-1} \Delta p_{t-1} \\ \Delta p_{t-1} & p_{t-1} \Delta p_{t-1} & \left(\Delta p_{t-1}\right)^2 \end{bmatrix}\right)^{-1} \begin{pmatrix} 0 \\ n_s^{(2m-1)} \\ 0 \end{pmatrix} \frac{\sum \check{e}_t^2}{n_s^{(2m-4)} \left(n_s - 5\right)} \\
= \; & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} C \left(\sum \begin{bmatrix} 1 & p_{t-1} & \Delta p_{t-1} \\ p_{t-1} & p_{t-1}^2 & p_{t-1} \Delta p_{t-1} \\ \Delta p_{t-1} & p_{t-1} \Delta p_{t-1} & \left(\Delta p_{t-1}\right)^2 \end{bmatrix}\right)^{-1} G \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \frac{\sum \check{e}_t^2}{n_s^{(2m-4)} \left(n_s - 5\right)} \\
= \; & O_p\left(1\right) \quad \text{as } n \to \infty.
\end{aligned}$$

Together with the fact of $n_s^2 \left(\check{\beta}_{r_1}^{r_2} - 1\right) \xrightarrow{p} -\frac{m(2m+1)(m+1)(m-2)}{(2m-1)(m-1)} < 0$, as $n \to \infty$, we have

$$t\left(\check{\beta}_{r_1}^{r_2}\right) = \frac{n_s^2 \left(\check{\beta}_{r_1}^{r_2} - 1\right)}{n_s^{5/2} se\left(\check{\beta}_{r_1}^{r_2}\right)} \sqrt{n_s} \xrightarrow{p} -\infty.$$

Proof of Theorem 3.1.3

To prove the results in this part of the theorem, we only need to show that, when $n_1 = n_c$ and $n_2 = n$, it has $(n_2 - n_1)\left(\check{\beta}_{r_1}^{r_2} - 1\right) \to \infty$ and $t\left(\check{\beta}_{r_1}^{r_2}\right) = \left(\check{\beta}_{r_1}^{r_2} - 1\right)/se\left(\check{\beta}_{r_1}^{r_2}\right) \to \infty$ as $n \to \infty$. From n_c to n, $\{p_t\}$ are generated by a mildly explosive process as $p_t = \rho_{n_e} p_{t-1} + \varepsilon_t$ with $\rho_{n_e} = 1 + \gamma / n_e^{\theta}$ and $n_e = n - n_c$. For any $t > n_c + 2$, it has

$$\Delta p_{t-1} = p_{t-1} - \rho_{n_e}^{-1}\left(p_{t-1} - \varepsilon_{t-1}\right) = \frac{\gamma}{n_e^{\theta}\rho_{n_e}} p_{t-1} + \rho_{n_e}^{-1}\varepsilon_{t-1}.$$

Phillips and Magdalinos (2007b) showed that $p_n = O_p\left(\rho_{n_e}^{n_e} n_e^{\theta/2}\right)$ as $n_e \to \infty$. Thus, the first term in the above equation dominates the second term when t is large. As a result, Model (1.2) encounters the problem of perfect multi-collinearity asymptotically.

To address the problem of asymptotic perfect multi-collinearity, we consider the transformed regression of

$$p_t = \check{\alpha}_{r_1}^{*,r_2} + \check{\beta}_{r_1}^{*,r_2} p_{t-1} + \check{\psi}_{r_1}^{*,r_2}\left(\rho_{n_e}^{-1} u_{t-1}\right) + \check{e}_t^{*}, \tag{A.1}$$

whose centered LS estimators have the following relationship with the centered LS estimators of the proposed regression model (1.2):

$$\begin{pmatrix} \check{\alpha}_{r_1}^{r_2} - 0 \\ \check{\beta}_{r_1}^{r_2} - \rho_{n_e} \\ \check{\psi}_{r_1}^{r_2} - 0 \end{pmatrix} = D' \begin{pmatrix} \check{\alpha}_{r_1}^{*,r_2} - 0 \\ \check{\beta}_{r_1}^{*,r_2} - \rho_{n_e} \\ \check{\psi}_{r_1}^{*,r_2} - 0 \end{pmatrix} \quad \text{with} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{\gamma}{n_e^{\theta}\rho_{n_e}} & 1 \end{pmatrix},$$

which leads to

$$\check{\beta}_{r_1}^{r_2} - \rho_{n_e} = \left(\check{\beta}_{r_1}^{*,r_2} - \rho_{n_e}\right) - \frac{\gamma}{n_e^{\theta}\rho_{n_e}}\left(\check{\psi}_{r_1}^{*,r_2} - 0\right).$$

Note that

$$\begin{pmatrix} \check{\alpha}_{r_1}^{*,r_2} - 0 \\ \check{\beta}_{r_1}^{*,r_2} - \rho_{n_e} \\ \check{\psi}_{r_1}^{*,r_2} - 0 \end{pmatrix} = \left[\sum \begin{pmatrix} 1 & p_{t-1} & \rho_{n_e}^{-1} u_{t-1} \\ p_{t-1} & p_{t-1}^2 & p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) \\ \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) & p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) & \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)^2 \end{pmatrix}\right]^{-1} \begin{pmatrix} \sum \varepsilon_t \\ \sum p_{t-1}\varepsilon_t \\ \sum \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)\varepsilon_t \end{pmatrix}$$

where $\sum = \sum_{t=n_c+2}^{n}$. Phillips and Magdalinos (2007b) proved that

$$n_e^{-2\theta}\rho_{n_e}^{-2n_e}\sum p_{t-1}^2 \Rightarrow \frac{\eta^2}{4\gamma^2} \quad \text{and} \quad n_e^{-\theta}\rho_{n_e}^{-n_e}\sum p_{t-1}\varepsilon_t \Rightarrow \frac{\eta\xi}{2\gamma},$$

where η and ξ are two independent standard normal variates. We then have

$$\begin{aligned} n_e^{-\theta}\rho_{n_e}^{-n_e}\sum p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) &= n_e^{-\theta}\rho_{n_e}^{-n_e}\left(\sum p_{t-2}\varepsilon_{t-1} + \rho_{n_e}^{-1}\sum \varepsilon_{t-2}\varepsilon_{t-1}\right) \\ &= n_e^{-\theta}\rho_{n_e}^{-n_e}\left(\sum p_{t-1}\varepsilon_t + p_{n_c}\varepsilon_{n_c+1} - p_{n-1}\varepsilon_n + \rho_{n_e}^{-1}\sum \varepsilon_{t-2}\varepsilon_{t-1}\right) \\ &= n_e^{-\theta}\rho_{n_e}^{-n_e}\sum p_{t-1}\varepsilon_t + o_p(1) \Rightarrow \frac{\eta\xi}{2\gamma}. \end{aligned}$$

Together with the result in Wang and Yu (2016) that $n_e^{-3\theta/2}\rho_{n_e}^{-n_e}\sum p_{t-1} \Rightarrow \frac{\eta}{\gamma\sqrt{2\gamma}}$, we can have

$$\Psi^{-1}\begin{pmatrix}\sum \varepsilon_t \\ \sum p_{t-1}\varepsilon_t \\ \sum\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)\varepsilon_t\end{pmatrix} \Rightarrow \begin{pmatrix}\sigma W(1) \\ \eta\xi/(2\gamma) \\ 0\end{pmatrix} \quad \text{with} \quad \Psi = \begin{pmatrix} n_e^{1/2} & 0 & 0 \\ 0 & n_e^{\theta}\rho_{n_e}^{n_e} & 0 \\ 0 & 0 & n_e\end{pmatrix},$$

$$\begin{pmatrix} n_e^{1/2} & 0 & 0 \\ 0 & n_e^{\theta}\rho_{n_e}^{n_e} & 0 \\ 0 & 0 & 1\end{pmatrix}\left[\sum\begin{pmatrix} 1 & p_{t-1} & \rho_{n_e}^{-1}\varepsilon_{t-1} \\ p_{t-1} & p_{t-1}^2 & p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) \\ \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) & p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) & \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)^2\end{pmatrix}\right]^{-1}\Psi$$
$$\Rightarrow \begin{pmatrix} 1 & 0 & \sigma W(1) \\ 0 & \eta^2/\left(4\gamma^2\right) & \eta\xi/(2\gamma) \\ 0 & 0 & \sigma^2\end{pmatrix}^{-1}.$$

Consequently, it has

$$\begin{pmatrix} n_e^{1/2} & 0 & 0 \\ 0 & n_e^{\theta}\rho_{n_e}^{n_e} & 0 \\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\check{\alpha}_{r_1}^{*,r_2} - 0 \\ \check{\beta}_{r_1}^{*,r_2} - \rho_{n_e} \\ \check{\psi}_{r_1}^{*,r_2} - 0\end{pmatrix} \Rightarrow \begin{pmatrix} 1 & 0 & \sigma W(1) \\ 0 & \eta^2/\left(4\gamma^2\right) & \eta\xi/(2\gamma) \\ 0 & 0 & \sigma^2\end{pmatrix}^{-1}\begin{pmatrix}\sigma W(1) \\ \eta\xi/(2\gamma) \\ 0\end{pmatrix},$$

which leads to

$$n_e^{\theta}\rho_{n_e}^{n_e}\left(\check{\beta}_{r_1}^{*,r_2} - \rho_{n_e}\right) \Rightarrow 2\gamma\frac{\xi}{\eta} \quad \text{and} \quad \check{\psi}_{r_1}^{*,r_2} \xrightarrow{p} 0.$$

We thus have

$$\begin{aligned} n_e\left(\check{\beta}_{r_1}^{r_2} - 1\right) &= n_e\left(\check{\beta}_{r_1}^{r_2} - \rho_{n_e}\right) + n_e\left(\rho_{n_e} - 1\right) \\ &= n_e\left(\check{\beta}_{r_1}^{*,r_2} - \rho_{n_e}\right) - \frac{n_e\gamma}{n_e^{\theta}\rho_{n_e}}\left(\check{\psi}_{r_1}^{*,r_2} - 0\right) + n_e\left(\rho_{n_e} - 1\right) \\ &= n_e\left(\check{\beta}_{r_1}^{*,r_2} - \rho_{n_e}\right) - \frac{n_e\gamma}{n_e^{\theta}\rho_{n_e}}\left(\check{\psi}_{r_1}^{*,r_2} - 0\right) + n_e\frac{\gamma}{n_e^{\theta}} \\ &= n_e\left(\check{\beta}_{r_1}^{*,r_2} - \rho_{n_e}\right) + \frac{n_e\gamma}{n_e^{\theta}}\left(1 - \rho_{n_e}^{-1}\check{\psi}_{r_1}^{*,r_2}\right) \xrightarrow{p} +\infty, \end{aligned}$$

where the last limit comes from the fact that

$$n_e\left(\check{\beta}_{r_1}^{*,r_2} - \rho_{n_e}\right) = O_p\left(\frac{n_e}{n_e^{\theta}\rho_{n_e}}\right) = o_p(1) \quad \text{and } 1 - \rho_{n_e}^{-1}\check{\psi}_{r_1}^{*,r_2} \xrightarrow{p} 1.$$

To prove the limit of $t\left(\check{\beta}_{r_1}^{r_2}\right) = \left(\check{\beta}_{r_1}^{r_2} - 1\right)/se\left(\check{\beta}_{r_1}^{r_2}\right)$, we first study the limit of $se\left(\check{\beta}_{r_1}^{r_2}\right)$. It is easy to prove that

$$\begin{aligned} \frac{1}{n_e}\sum \check{e}_t^2 &= \frac{1}{n_e}\sum\left[p_t - \check{\alpha}_{r_1}^{r_2} - \check{\beta}_{r_1}^{r_2}p_{t-1} - \check{\psi}_{r_1}^{r_2}\Delta p_{t-1}\right]^2 \\ &= \frac{1}{n_e}\sum\left[p_t - \check{\alpha}_{r_1}^{*,r_2} - \check{\beta}_{r_1}^{*,r_2}p_{t-1} - \check{\psi}_{r_1}^{*,r_2}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)\right]^2 = \frac{1}{n_e}\sum\left(\check{e}_t^*\right)^2. \end{aligned}$$

Given that the true DGP of p_t is covered by the transformed regression model (6.11) and that the LS estimates of the parameters are consistent, it can be shown that

$$\frac{1}{n_e}\sum\left(\check{e}_t^*\right)^2 = \frac{1}{n_e}\sum \varepsilon_t^2 + o_p(1) \xrightarrow{p} \sigma^2.$$

Note that

$$
\begin{aligned}
& n_e^{2\theta}\begin{bmatrix}0 & 1 & 0\end{bmatrix}\left[\sum\begin{pmatrix}1 & p_{t-1} & \Delta p_{t-1}\\ p_{t-1} & p_{t-1}^2 & p_{t-1}\Delta p_{t-1}\\ \Delta p_{t-1} & p_{t-1}\Delta p_{t-1} & (\Delta p_{t-1})^2\end{pmatrix}\right]^{-1}\begin{bmatrix}0\\1\\0\end{bmatrix}\\
= \;& n_e^{2\theta}\begin{bmatrix}0 & 1 & 0\end{bmatrix}D'\left[\sum D\begin{pmatrix}1 & p_{t-1} & \Delta p_{t-1}\\ p_{t-1} & p_{t-1}^2 & p_{t-1}\Delta p_{t-1}\\ \Delta p_{t-1} & p_{t-1}\Delta p_{t-1} & (\Delta p_{t-1})^2\end{pmatrix}D'\right]^{-1}D\begin{bmatrix}0\\1\\0\end{bmatrix}\\
= \;& n_e^{2\theta}\begin{bmatrix}0 & 1 & -\frac{\gamma}{n_e^{\theta}\rho_{n_e}}\end{bmatrix}\left[\sum\begin{pmatrix}1 & p_{t-1} & \rho_{n_e}^{-1}\varepsilon_{t-1}\\ p_{t-1} & p_{t-1}^2 & p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)\\ \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) & p_{t-1}\left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right) & \left(\rho_{n_e}^{-1}\varepsilon_{t-1}\right)^2\end{pmatrix}\right]^{-1}\begin{bmatrix}0\\1\\-\frac{\gamma}{n_e^{\theta}\rho_{n_e}}\end{bmatrix}\\
= \;& n_e^{2\theta}\begin{bmatrix}0 & 1 & -\frac{\gamma}{n_e^{\theta}\rho_{n_e}}\end{bmatrix}\begin{pmatrix}n_e^{-1/2} & 0 & 0\\ 0 & n_e^{-\theta}\rho_{n_e}^{-n_e} & 0\\ 0 & 0 & 1\end{pmatrix}\left\{\begin{pmatrix}1 & 0 & \sigma W(1)\\ 0 & \eta^2/\left(4\gamma^2\right) & \eta\xi/(2\gamma)\\ 0 & 0 & \sigma^2\end{pmatrix}^{-1} + o_p(1)\right\}\\
& \times\begin{pmatrix}n_e^{-1/2} & 0 & 0\\ 0 & n_e^{-\theta}\rho_{n_e}^{-n_e} & 0\\ 0 & 0 & n_e^{-1}\end{pmatrix}\begin{bmatrix}0\\1\\-\frac{\gamma}{n_e^{\theta}\rho_{n_e}}\end{bmatrix}\\
= \;& n_e^{2\theta}\begin{bmatrix}0 & n_e^{-\theta}\rho_{n_e}^{-n_e} & -\frac{c}{n_e^{\theta}\rho_{n_e}}\end{bmatrix}\left\{\begin{pmatrix}1 & 0 & \sigma W(1)\\ 0 & \eta^2/\left(4\gamma^2\right) & \eta\xi/(2\gamma)\\ 0 & 0 & \sigma^2\end{pmatrix}^{-1} + o_p(1)\right\}\begin{bmatrix}0\\ n_e^{-\theta}\rho_{n_e}^{-n_e}\\ -\frac{\gamma}{n_e^{1+\theta}\rho_{n_e}}\end{bmatrix}\\
= \;& \begin{bmatrix}0 & \rho_{n_e}^{-n_e} & -\frac{\gamma}{\rho_{n_e}}\end{bmatrix}\left\{\begin{pmatrix}1 & 0 & \sigma W(1)\\ 0 & \eta^2/\left(4\gamma^2\right) & \eta\xi/(2\gamma)\\ 0 & 0 & \sigma^2\end{pmatrix}^{-1} + o_p(1)\right\}\begin{bmatrix}0\\ \rho_{n_e}^{-n_e}\\ -\frac{\gamma}{n_e\rho_{n_e}}\end{bmatrix}\\
\xrightarrow{p} \;& \begin{bmatrix}0 & 0 & -\gamma\end{bmatrix}\begin{pmatrix}1 & 0 & \sigma W(1)\\ 0 & \eta^2/\left(4\gamma^2\right) & \eta\xi/(2\gamma)\\ 0 & 0 & \sigma^2\end{pmatrix}^{-1}\begin{bmatrix}0\\0\\0\end{bmatrix} = 0.
\end{aligned}
$$

We then have

$$n_e^{2\theta}\left[se\left(\check{\beta}_{r_1}^{r_2}\right)\right]^2 = n_e^{2\theta}\begin{bmatrix}0 & 1 & 0\end{bmatrix}\left[\sum\begin{pmatrix}1 & p_{t-1} & \Delta p_{t-1}\\ p_{t-1} & p_{t-1}^2 & p_{t-1}\Delta p_{t-1}\\ \Delta p_{t-1} & p_{t-1}\Delta p_{t-1} & (\Delta p_{t-1})^2\end{pmatrix}\right]^{-1}\begin{bmatrix}0\\1\\0\end{bmatrix}\frac{\sum \check{e}_t^2}{n_e-4} \xrightarrow{p} 0.$$

Therefore, as $n \to \infty$,

$$t\left(\check{\beta}_{r_1}^{r_2}\right) = \frac{\check{\beta}_{r_1}^{r_2} - 1}{se\left(\check{\beta}_{r_1}^{r_2}\right)} = \frac{n_e^{\theta}\left(\check{\beta}_{r_1}^{r_2} - 1\right)}{n_e^{\theta} se\left(\check{\beta}_{r_1}^{r_2}\right)} = \frac{n_e^{\theta}\left(\check{\beta}_{r_1}^{r_2} - \rho_{n_e}\right) + n_e^{\theta}\left(\rho_{n_e} - 1\right)}{n_e^{\theta} se\left(\check{\beta}_{r_1}^{r_2}\right)} \xrightarrow{p} +\infty.$$

The proof is completed.

Tests based on AR(k) regression

In this subsection we extend our results for the AR(2) regression reported in the main paper to the AR(k) case with $k > 2$:

$$p_t = \alpha + \beta p_{t-1} + \sum_{i=1}^{k-1} \psi_i \Delta p_{t-i} + \varepsilon_t, t = 1, ..., n. \tag{A.2}$$

Let $\breve{\beta}$ denote the LS estimate of β. We will show that both the DF coefficient-based test and the DF t-test can distinguish successfully between the polynomial trend model and the explosive AR model. In other words, they are robust to polynomial trend. For simplicity, we only develop the asymptotics of the test statistics in full sample analysis. The results for subsample analysis in the context of models with a structural break can be obtained similarly.

It is well-known in the literature that, if the true DGP of p_t is an AR model with the lag length smaller or equal to k and the errors $\{\varepsilon_t\}$ being an iid sequence, then under the null hypothesis of unit root, as $n \to \infty$,

$$n\left(\breve{\beta} - 1\right) \Rightarrow \frac{\int \widetilde{W}(s)\, dW(s)}{\int \left[\widetilde{W}(s)\right]^2 ds} \quad \text{and} \quad t\left(\breve{\beta}\right) \Rightarrow \frac{\int \widetilde{W}(s)\, dW(s)}{\left\{\int \left[\widetilde{W}(s)\right]^2 ds\right\}^{1/2}} \tag{A.3}$$

where $\int = \int_0^1$ and

$$\widetilde{W}(s) := W(s) - \int W(t)\, dt.$$

Theorem 1 below develops the large sample theory of $n\left(\breve{\beta} - 1\right)$ and $t\left(\breve{\beta}\right)$, when $\{p_t\}$ is generated entirely from a mildly explosive AR(1) model or from a polynomial trend model.

Theorem 1 *Assume the AR(k) model (A.2) with $k > 2$ is estimated by LS.*
(a) If the true DGP of p_t is a mildly explosive process, that is, $p_t = \rho_n p_{t-1} + \varepsilon_t$ with $\rho_n = 1 + \gamma/n^\theta$, $\gamma > 0$, and $\theta \in (0, 1)$, then, as $n \to \infty$,

$$n\left(\breve{\beta} - 1\right) \xrightarrow{p} +\infty \quad \textit{and} \quad t\left(\breve{\beta}\right) \xrightarrow{p} +\infty.$$

(b) If the true DGP is $p_t = \alpha + \delta t^m + \varepsilon_t$ with $m \geq 2$. When $k \geq m$, as $n \to \infty$, it has,

$$n\left(\breve{\beta} - 1\right) \xrightarrow{p} 0 \quad \textit{and} \quad t\left(\breve{\beta}\right) = O_p\left(n^{-1/2}\right) \xrightarrow{p} 0.$$

(c) If the true DGP is $p_t = \theta + \delta t^m + u_t$ with $m \geq 2$. When $2 < k < m$, as $n \to \infty$, it has,

$$n\left(\breve{\beta} - 1\right) \xrightarrow{p} 0$$

and

$$t\left(\breve{\beta}\right) = O_p\left(\sqrt{n}\right) \to -\infty \textit{ as long as } \lim_{n\to\infty} n^k\left(\breve{\beta} - 1\right) < 0.$$

According to Part (a) of Theorem 1, when the true DGP of p_t is a mildly explosive process, the right-tailed unit root tests, obtained based on the AR(k) regression in (A.2) with $k > 2$, can find evidence of explosiveness. According to Part (b)-(c) of Theorem 1, when the true DGP of p_t has a polynomial trend with $m \geq 2$, the right-tailed unit root tests, obtained based on the AR(k) regression in (A.2) with $k > 2$, do not find evidence of explosiveness. This is because $n\left(\check{\beta}-1\right) \xrightarrow{p} 0$ and $t\left(\check{\beta}\right)$ converges to zero or diverges to minus infinity, both less than the CV of the corresponding null asymptotic distribution. Consequently, we can claim that both the coefficient-based test and the t test can distinguish between the polynomial trend model and the explosive AR model.

For Part (c) of Theorem 1, in general, the value of $\lim_{n\to\infty} n^k\left(\check{\beta}-1\right)$ varies with k and m. We conjecture that the sign of $\lim_{n\to\infty} n^k\left(\check{\beta}-1\right)$ is always negative. For the special case where $k = 2$ and $m > k$, the proof of Theorem 3.1.1 given above has proved that $\lim_{n\to\infty} n^k\left(\check{\beta}-1\right) < 0$, and hence, $t\left(\check{\beta}\right) \to -\infty$.

We now prove Theorem **1.** Part (a) is easier to prove. Note that in the proof of Theorem 3.1.3 in the main paper, the results of $n\left(\check{\beta}-1\right) \xrightarrow{p} +\infty$ and $t(\check{\beta}) \to +\infty$ have been proved for the case of $k = 2$. By taking the same procedure, the results for $k > 2$ can be proved directly.

We then prove Part (b). For the case where $m \leq k$, the process $p_t = \theta + \delta t^m + u_t$ can be represented as an AR(k) regression given in (A.2) with proper choice of the AR coefficients and the setup of the regression errors. For example, if $m = 2$, p_t can be rewritten as

$$p_t = 2\delta + p_{t-1} + \Delta p_{t-1} + \Delta^2 u_t,$$

which is covered by the AR(k) regression given in (A.2) with $k \geq 2$. According to Theorem 3.1.1 in the main paper, when $k = 2$, we have $t\left(\check{\beta}\right) = O_p\left(n^{-1/2}\right)$ as $n \to \infty$. By taking the same procedure, the result of $t\left(\check{\beta}\right) = O_p\left(n^{-1/2}\right)$ can be proved for general values of m and k with $m = k$.

We now consider the case where $m < k$. In the following, we take $m = 2$ and $k = 3$ as an example to carry out the proof. The result for other cases of $m < k$ can be obtained by taking the same approach.

When $k = 3$, Model (A.2) becomes

$$p_t = \alpha + \beta p_{t-1} + \psi_1 \Delta p_{t-1} + \psi_2 \Delta p_{t-2} + \epsilon_t.$$

Hence, the above regression with $\begin{pmatrix}\alpha & \beta & \psi_1 & \psi_2\end{pmatrix} = \begin{pmatrix}2\delta & 1 & 1 & 0\end{pmatrix}$ and $\epsilon_t = \Delta^2 \varepsilon_t$ gives the true DGP of p_t. However, the regression confronts with the problem of asymptotic multicollinearity as

$$\Delta p_{t-1} = 2\delta t - 3\delta + \Delta\varepsilon_{t-1} = O\left(t\right) \quad \text{and } \Delta p_{t-2} = 2\delta t - 5\delta + \Delta\varepsilon_{t-2} = O\left(t\right).$$

To get the asymptotics of the LS estimators $\begin{pmatrix}\check{\alpha} & \check{\beta} & \check{\psi}_1 & \check{\psi}_2\end{pmatrix}$, we consider an alternative regression:

$$p_t = \check{\alpha}^* + \check{\beta}^* p_{t-1} + \check{\psi}_1^* \Delta p_{t-1} + \check{\psi}_2^* \Delta^2 \varepsilon_{t-1} + \check{e}_t^*. \tag{A.4}$$

Using the fact of $\Delta p_{t-1} - \Delta p_{t-2} = 2\delta + \Delta^2 \varepsilon_{t-1}$, it is easily obtained that

$$\begin{pmatrix} \breve{\alpha} - 2\delta \\ \breve{\beta} - 1 \\ \breve{\psi}_1 - 1 \\ \breve{\psi}_2 - 0 \end{pmatrix} = M \begin{pmatrix} \breve{\alpha}^* - 2\delta \\ \breve{\beta}^* - 1 \\ \breve{\psi}_1^* - 1 \\ \breve{\psi}_2^* - 0 \end{pmatrix} \quad \text{with} \quad M := \begin{pmatrix} 1 & 0 & 0 & -2\delta \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

By using the large sample properties of the quadratic trend model as obtained in the proof of Theorem 3.1.1 in the main paper, it is easy to get that

$$\begin{pmatrix} n & 0 & 0 & 0 \\ 0 & n^{m+1} & 0 & 0 \\ 0 & 0 & n^m & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \breve{\alpha}^* - 2\delta \\ \breve{\beta}^* - 1 \\ \breve{\psi}_1^* - 1 \\ \breve{\psi}_2^* - 0 \end{pmatrix} = O_p(1),$$

which leads to the result of

$$\breve{\beta} - 1 = \breve{\beta}^* - 1 = O_p\left(n^{-m-1}\right) \quad \text{as} \quad n \to \infty.$$

Next, we consider the limit of the sum of squared regression errors of (A.2). As $n \to \infty$:

$$\begin{aligned} n^{-1} \sum_{t=1}^{n} (\breve{e}_t)^2 &= n^{-1} \sum_{t=1}^{n} \left(p_t - \breve{\alpha} - \breve{\beta} p_{t-1} - \breve{\psi}_1 \Delta p_{t-1} - \breve{\psi}_2 \Delta p_{t-2}\right)^2 \\ &= n^{-1} \sum_{t=1}^{n} \left(p_t - \breve{\alpha}^* - \breve{\beta}^* p_{t-1} - \breve{\psi}_1^* \Delta p_{t-1} - \breve{\psi}_2^* \Delta^2 \varepsilon_{t-1}\right)^2 \\ &= O_p(1). \end{aligned}$$

Moreover,

$$\begin{aligned} &\left[se\left(\breve{\beta}\right)\right]^2 \\ = \; &\begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \left\{ \sum_{t=1}^{n} \begin{pmatrix} 1 \\ p_{t-1} \\ \Delta p_{t-1} \\ \Delta p_{t-2} \end{pmatrix} \begin{pmatrix} 1 & p_{t-1} & \Delta p_{t-1} & \Delta p_{t-2} \end{pmatrix} \right\}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \frac{\sum_{t=1}^{n} (\breve{e}_t)^2}{n} \\ = \; &\begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} M \left\{ \sum_{t=1}^{n} \begin{pmatrix} 1 \\ p_{t-1} \\ \Delta p_{t-1} \\ \Delta^2 \varepsilon_{t-1} \end{pmatrix} \begin{pmatrix} 1 & p_{t-1} & \Delta p_{t-1} & \Delta^2 \varepsilon_{t-1} \end{pmatrix} \right\}^{-1} M' \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \frac{\sum_{t=1}^{n} (\breve{e}_t)^2}{n} \\ = \; &O_p\left(n^{-(2m+1)}\right). \end{aligned}$$

Finally, we get

$$t\left(\breve{\beta}\right) = \frac{\breve{\beta} - 1}{se\left(\breve{\beta}\right)} = n^{-1/2} \frac{n^{m+1}\left(\breve{\beta} - 1\right)}{n^{m+1/2} se\left(\breve{\beta}\right)} = O_p\left(n^{-1/2}\right).$$

We now prove Part (c). For the case where $m > k$, the process $p_t = \theta + \delta t^m + u_t$ cannot be represented as an $AR(k)$ regression with stationary errors, no matter how the AR coefficients are chosen. Theorem 3.1.1 in the main paper has given the asymptotic theory of the test statistics when $m > k$ and $k = 2$. In the following, we use the case of $k = 3$ as an example to show that the results can be extended to the case where k take any values less than m.

When $k = 3$, the $AR(k)$ regression of (A.2) has the representation as

$$\begin{aligned} p_t &= \alpha + \beta p_{t-1} + \psi_1 \Delta p_{t-1} + \psi_2 \Delta p_{t-2} + \varepsilon_t \\ &= \alpha + \beta p_{t-1} + (\psi_1 + \psi_2) \Delta p_{t-1} - \psi_2 \Delta^2 p_{t-1} + \varepsilon_t \\ &= \alpha^* + \beta^* y_{t-1} + \psi_1^* \Delta p_{t-1} + \psi_2^* \Delta^2 p_{t-1} + \varepsilon_t. \end{aligned}$$

Whereas, the polynomial trend process has the representation as

$$\begin{aligned} p_t &= \theta + \delta t^m + u_t = \theta + \delta \left[\sum_{j=0}^{m} C_m^j (t-1)^{m-j} \right] + u_t \\ &= p_{t-1} + \delta \left[\sum_{j=1}^{m} C_m^j (t-1)^{m-j} \right] + \Delta u_t \\ &= p_{t-1} + \Delta p_{t-1} + \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=1}^{m-j} C_{m-j}^i (t-2)^{m-j-i} \right] + \Delta^2 u_t \\ &= p_{t-1} + \Delta p_{t-1} + \Delta^2 p_{t-1} + \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=1}^{m-j} C_{m-j}^i \sum_{\tau=1}^{m-j} C_{m-j-i}^{\tau} (t-3)^{m-j-i-\tau} \right] + \Delta^3 u_t \\ &= p_{t-1} + \Delta p_{t-1} + \Delta^2 p_{t-1} + O_p\left(t^{m-3}\right), \end{aligned}$$

where the fourth and fifth equalities come from the following two equations, respectively,

$$\begin{aligned} \Delta p_t &= \delta \left[\sum_{j=1}^{m} C_m^j (t-1)^{m-j} \right] + \Delta u_t = \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=0}^{m-j} C_{m-j}^i (t-2)^{m-j-i} \right] + \Delta u_t \\ &= \Delta p_{t-1} + \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=1}^{m-j} C_{m-j}^i (t-2)^{m-j-i} \right] + \Delta^2 u_t, \end{aligned}$$

and

$$
\begin{aligned}
\Delta^2 p_t &= \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=1}^{m-j} C_{m-j}^i (t-2)^{m-j-i} \right] + \Delta^2 u_t \\
&= \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=1}^{m-j} C_{m-j}^i \sum_{\tau=0}^{m-j} C_{m-j-i}^{\tau} (t-3)^{m-j-i-\tau} \right] + \Delta^2 u_t \\
&= \Delta^2 p_{t-1} + \delta \left[\sum_{j=1}^{m} C_m^j \sum_{i=1}^{m-j} C_{m-j}^i \sum_{\tau=1}^{m-j} C_{m-j-i}^{\tau} (t-3)^{m-j-i-\tau} \right] + \Delta^3 u_t.
\end{aligned}
$$

Therefore, as long as $m > 3$, the process p_t cannot be represented as an AR(3) model with stationary errors. This makes the sum of squared regression errors have the following limit:

$$
n^{-2(m-k)-1} \sum_{t=1}^{n} (\breve{e}_t)^2 = O_p(1).
$$

To study the asymptotic behavior of $\breve{\beta}$ from the AR(3) regression, instead of the original regression

$$
p_t = \breve{\alpha} + \breve{\beta} p_{t-1} + \breve{\psi}_1 \Delta p_{t-1} + \breve{\psi}_2 \Delta p_{t-2} + \breve{e}_t,
$$

we consider the transformed regression

$$
p_t = \breve{\alpha}^* + \breve{\beta}^* p_{t-1} + \breve{\psi}_1^* \Delta p_{t-1} + \breve{\psi}_2^* \Delta^2 p_{t-1} + \breve{e}_t^*,
$$

to avoid the multicollinearity problem. The following relationship is easy to get

$$
\begin{pmatrix} \breve{\alpha} - 0 \\ \breve{\beta} - 1 \\ \breve{\psi}_1 - 1 \\ \breve{\psi}_2 - 0 \end{pmatrix} = D \begin{pmatrix} \breve{\alpha}^* - 0 \\ \breve{\beta}^* - 1 \\ \breve{\psi}_1^* - 1 \\ \breve{\psi}_2^* - 0 \end{pmatrix} \quad \text{with} \quad D := \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.
$$

Given the facts of $y_{t-1} = O(t^m)$, $\Delta y_{t-1} = O\left(t^{m-1}\right)$, $\Delta^2 y_{t-2} = O\left(t^{m-2}\right)$, and $\breve{e}_t^* = O\left(t^{m-3}\right)$, together with the property of $\lim_{n\to\infty} n^{-s+1} \sum_{t=1}^{n} t^s = 1/(s+1)$, it can be proved that, as $n \to \infty$,

$$
\begin{pmatrix} n^{-m+k} & & & \\ & n^k & & \\ & & n^{k-1} & \\ & & & n^{k-2} \end{pmatrix} \begin{pmatrix} \breve{\alpha}^* - 0 \\ \breve{\beta}^* - 1 \\ \breve{\psi}_1^* - 1 \\ \breve{\psi}_2^* - 0 \end{pmatrix} \xrightarrow{p} \text{a vector of constants,}
$$

where the limit depends on the values of m, k, and δ. Consequently, it has

$$
n^k \left(\breve{\beta} - 1\right) = n^k \left(\breve{\beta}^* - 1\right) = \text{a constant} + o_p(1), \text{ as } n \to \infty.
$$

Next, consider the limit of the standard error of $\check{\beta}$, which is

$$
\begin{aligned}
& \left[se\left(\check{\beta}\right)\right]^2 \\
= & \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \left\{ \sum_{t=1}^{n} \begin{pmatrix} 1 \\ p_{t-1} \\ \Delta p_{t-1} \\ \Delta p_{t-2} \end{pmatrix} \begin{pmatrix} 1 & p_{t-1} & \Delta p_{t-1} & \Delta p_{t-2} \end{pmatrix} \right\}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \frac{\sum_{t=1}^{n} \left(\check{e}_t\right)^2}{n} \\
= & \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} D \left\{ \sum_{t=1}^{n} \begin{pmatrix} 1 \\ p_{t-1} \\ \Delta p_{t-1} \\ \Delta^2 u_{t-1} \end{pmatrix} \begin{pmatrix} 1 & p_{t-1} & \Delta p_{t-1} & \Delta^2 u_{t-1} \end{pmatrix} \right\}^{-1} D' \begin{pmatrix} 0 \\ n^{2(m-k)} \\ 0 \\ 0 \end{pmatrix} \frac{\sum_{t=1}^{n} \left(\check{e}_t\right)^2}{n^{2(m-k)} n} \\
= & O_p\left(n^{-(2k+1)}\right).
\end{aligned}
$$

Finally, it is obtained that

$$
t\left(\check{\beta}\right) = \frac{\check{\beta} - 1}{se\left(\check{\beta}\right)} = n^{1/2} \frac{n^k \left(\check{\beta} - 1\right)}{n^{k+1/2} se\left(\check{\beta}\right)} = O_p\left(n^{1/2}\right).
$$

More Simulation Results

Assume the data is generated from the following structural break model:

$$
p_t = \begin{cases} p_{t-1} + \varepsilon_t & \text{if } t \in N \equiv [1, n_c] \\ p_{n_c} + \delta\left(t - n_c\right)^m + \varepsilon_t & \text{if } t \in T \equiv (n_c, n] \end{cases}, \tag{A.5}
$$

where $\delta \neq 0$, $m \geq 1$, $\{\varepsilon_t\}$ an independent and identically distributed (iid) sequence with mean zero and finite variance (denoted σ^2), and $E\left(\varepsilon_t^4\right) < \infty$.

In the main paper, we report simulation results of the four tests (DF coefficient-based test, DF t test, *GSADFc* test, and *GSADFt* test) based on AR(1) and AR(2) regressions when $m = 2$. We now report simulation results when $m = 3$ and $m = 1$.

Table 1 reports the simulation results when the AR(1) model is fitted to the data simulated from Model (A.5) with $m = 3$ (i.e. cubic trend). As in the quadratic trend model, the two full-sample tests and the two subsample tests always lead to spurious explosiveness.

Table 2 reports the simulation results when the AR(2) model is fitted to the data simulated from Model (A.5) with $m = 3$. In sharp contrast with the case when the AR(1) model is fitted, the two full-sample tests based on the AR(2) model never suggests explosiveness. While the *GSADFc* and *GSADFt* tests sometimes still suggest explosiveness, the probability of spurious detection when the AR(2) model is fitted is much smaller than the case when the AR(1) model is fitted. Comparing the finite performance of the *GSADFc* and *GSADFt* tests, the *GSADFt* test appears to perform better.

Table 3 reports the simulation results when the AR(1) model is fitted to the data simulated from Model (A.5) with $m = 1$ (i.e. linear trend). When data

Table 1: Statistical results when the AR(1) model is fitted to the data simulated from Model (A.5) with $m = 3$ and $r_c = 0.5$.

δ	0.5	1	0.5	1
Effective sample size	20	20	50	50
Mean of estimated β	1.1661	1.1661	1.0639	1.0639
Mean of c-stat	3.1550	3.1550	3.1289	3.1289
% rejection of H_0 by c-stat	1.0000	1.0000	1.0000	1.0000
Mean of t-stat	24.6749	24.6933	40.5855	40.5867
% rejection of H_0 by t-stat	1.0000	1.0000	1.0000	1.0000
Effective sample size	40	40	100	100
Mean of $GSADFc$	25.0760	32.1382	43.0405	54.9475
% rejection of H_0 by $GSADFc$	1.0000	1.0000	1.0000	1.0000
Mean of $GSADFt$	32.4304	32.3770	69.9372	69.9607
% rejection of H_0 by $GSADFt$	1.0000	1.0000	1.0000	1.0000

Note: Calculations are based on 1,000 replications when the AR(1) model is fitted to the simulated sample path. For the DF coefficient-based and DF t statistics, the simulated sample is from the polynomial trend model with $n = 20, 50$. For the $GSADFc$ and $GSADFt$ statistics, the simulated sample is from the model that switches from the random walk to the polynomial trend with $n = 40, r_0 = 0.3$ and $n = 100, r_0 = 0.2$.

Table 2: Statistical results when the AR(2) model is fitted to the data simulated from Model (A.5) with $m = 3$ and $r_c = 0.5$.

δ	0.5	1	0.5	1
Effective sample size	20	20	50	50
Mean of estimated β	0.9798	0.9792	0.9966	0.9966
Mean of c-stat	-0.3635	-0.3748	-0.1611	-0.1613
% rejection of H_0 by c-stat	0.0000	0.0000	0.0000	0.0000
Mean of t-stat	-4.3713	-7.1312	-13.5980	-15.5304
% rejection of H_0 by t-stat	0.0000	0.0000	0.0000	0.0000
Effective sample size	40	40	100	100
Mean of $GSADFc$	14.8666	16.4577	16.6165	17.2657
% rejection of H_0 by $GSADFc$	0.6540	0.6560	0.7650	0.7250
Mean of $GSADFt$	2.3301	1.9788	2.5469	2.0982
% rejection of H_0 by $GSADFt$	0.3180	0.2350	0.4410	0.2950

Note: Calculations are based on 1,000 replications when the AR(2) model is fitted to the simulated sample path. For the DF coefficient-based and DF t statistics, the simulated sample is from the polynomial trend model with $n = 20, 50$. For the $GSADFc$ and $GSADFt$ statistics, the simulated sample is from the model that switches from the random walk to the polynomial trend with $n = 40, r_0 = 0.3$ and $n = 100, r_0 = 0.2$.

Table 3: Statistical results when the AR(1) model is fitted to the data simulated from Model (A.5) with $m = 1$ and $r_c = 0.5$.

δ	10	20	10	20
Effective sample size	20	20	50	50
Mean of estimated β	0.9997	0.9999	0.9999	1.0000
Mean of c-stat	-0.0066	-0.0018	-0.0030	-0.0009
% rejection of H_0 by c-stat	0.0000	0.0000	0.0000	0.0000
Mean of t-stat	-0.0588	-0.0311	-0.0413	-0.0239
% rejection of H_0 by t-stat	0.0140	0.0160	0.0030	0.0030
Effective sample size	40	40	100	100
Mean of $GSADFc$	13.1264	19.5242	19.1444	29.2403
% rejection of H_0 by $GSADFc$	1.0000	1.0000	1.0000	1.0000
Mean of $GSADFt$	6.9567	7.1405	11.2694	11.7489
% rejection of H_0 by $GSADFt$	1.0000	1.0000	1.0000	1.0000

Note: Calculations are based on 1,000 replications when the AR(1) model is fitted to the simulated sample path. For the DF coefficient-based and DF t statistics, the simulated sample is from the linear trend model with $n = 20, 50$. For the $GSADFc$ and $GSADFt$ statistics, the simulated sample is from the model that switches from the random walk to the linear trend with $n = 40, r_0 = 0.3$ and $n = 100, r_0 = 0.2$.

come entirely from the linear trend, the DF coefficient-based test and t test never find evidence of explosiveness. However, when data come from a model that switches from unit root to linear trend, both $GSADFc$ and $GSADFt$ always find explosiveness. This finding is consistent with the asymptotic results in Phillips and Shi (2019).

Table 4 reports the simulation results when the AR(2) model is fitted to the data simulated from Model (A.5) with $m = 1$. When data come entirely from the linear trend, the DF coefficient-based test and t test based on AR(2) never find evidence of explosiveness, just like the case when AR(1) is fitted. However, when data come from a model that switches from unit root to linear trend, both $GSADFc$ and $GSADFt$ based on AR(2) often do not find evidence explosiveness. The $GSADFt$ test appears to have better power in finite samples in distinguish between the explosive AR model and the model that switch from unit root to linear trend.

Table 4: Statistical results when the AR(2) model is fitted to the data simulated from Model (A.5) with $m = 1$ and $r_c = 0.5$.

δ	10	20	10	20
Effective sample size	20	20	50	50
Mean of estimated β	0.9998	0.9999	1.0000	1.0000
Mean of c-stat	-0.0037	-0.0010	-0.0008	-0.0001
% rejection of H_0 by c-stat	0.0000	0.0000	0.0000	0.0000
Mean of t-stat	-0.0302	-0.0150	-0.0116	-0.0017
% rejection of H_0 by t-stat	0.0320	0.0340	0.0090	0.0110
Effective sample size	40	40	100	100
Mean of $GSADFc$	9.3359	16.3638	9.3344	13.0955
% rejection of H_0 by $GSADFc$	0.3350	0.4680	0.4350	0.4740
Mean of $GSADFt$	1.7612	1.5537	2.4709	1.6814
% rejection of H_0 by $GSADFt$	0.0780	0.0640	0.2390	0.0710

Note: Calculations are based on 1,000 replications when the AR(2) model is fitted to the simulated sample path. For the DF coefficient-based and DF t statistics, the simulated sample is from the linear trend model with $n = 20, 50$. For the $GSADFc$ and $GSADFt$ statistics, the simulated sample is from the model that switches from the random walk to the linear trend with $n = 40, r_0 = 0.3$ and $n = 100, r_0 = 0.2$.