Bubble Testing under Deterministic Trends*

Xiaohu Wang
Chinese University of Hong Kong

Jun Yu
Singapore Management University

September 9, 2019

Abstract

This paper develops the asymptotic theory of the ordinary least squares estimator of the autoregressive (AR) coefficient in various AR models, when data is generated from trend-stationary models in different forms. It is shown that, depending on how the autoregression is specified, the commonly used right-tailed unit root tests may tend to reject the null hypothesis of unit root in favor of the explosive alternative. A new procedure to implement the right-tailed unit root tests is proposed. It is shown that when the data generating process is trend-stationary, the test statistics based on the proposed procedure cannot find evidence of explosiveness. Whereas, when the data generating process is mildly explosive, the unit root tests find evidence of explosiveness. Hence, the proposed procedure enables robust bubble testing under deterministic trends. Empirical implementation of the proposed procedure using data from the stock and the real estate markets in the US reveals some interesting findings. While our proposed procedure flags the same number of bubbles episodes in the stock data as the method developed in Phillips, Shi and Yu (2015a, PSY), the estimated termination dates by the proposed procedure match better with the data. For real estate data, all negative bubble episodes flagged by PSY are no longer regarded as bubbles by the proposed procedure.

JEL classification: C12, C22, G01
Keywords: Autoregressive regressions, right-tailed unit root test, explosive and mildly explosive processes, deterministic trends, coefficient-based statistic, t-statistic.

1 Introduction

Other than its catastrophic effect on global financial markets of many kinds, the recent global financial crisis also turned the field of economics, including economic theory of financial markets, econometric theory of financial time series, and economic policies, on its head. Acknowledging an important empirical observation that often a financial bubble

*Xiaohu Wang, Department of Economics, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong. Email: xiaohu.wang@cuhk.edu.hk. Jun Yu, School of Economics and Lee Kong Chian School of Business, Singapore Management University, 90 Stamford Road, Singapore 178903. Email: yujun@smu.edu.sg. Wang acknowledges the support from the Hong Kong Research Grants Council General Research Fund under No. 14503718.

precedes immediately a financial crisis, a great deal of efforts that have been made in econometrics is to test for the presence of bubbles in financial time series and to estimate the origination date and the termination date of each bubble in real time. Widely used methods which have achieved great success in detecting and dating periodically collapsing bubbles are the recursive right-tailed unit root testing procedures proposed in Phillips, Wu and Yu (2011), Phillips and Yu (2011), and Phillips, Shi and Yu (2015a, 2016b, PSY hereafter). Harvey, et al. (2016), and Harvey, Leybourne and Zu (2017a, 2017b) extended the procedures to deal with the case with heteroscedasticity. A recent survey and comparisons of alternative methods in bubble testing and dating can be found in Homm and Breitung (2012).

In an environment of time-invariant discount rate, the standard no-arbitrage condition implies that

$$p_t = f_t + b_t, \tag{1}$$

where p_t is an asset price at time t, $f_t := \sum_{i=1}^{\infty}(1+r)^{-i}E_t(d_{t+i})$ is a "fundamental" component with d_{t+i} representing the payment received over the period from $t+i-1$ to $t+i$ due to the ownership of the asset, r is the discount rate ($r > 0$), and b_t is a bubble component which satisfies

$$E_t(b_{t+1}) = (1+r)b_t. \tag{2}$$

Since $\beta := 1+r > 1$, when $b_t > 0$, $E_t(b_{t+1}) = \beta b_t > b_t$, indicating the presence of an upward explosive behavior in b_{t+1}, i.e. a positive bubble. When f_{t+1} does not involve any explosive behavior, the upward explosive behavior in b_{t+1} suggests that p_{t+1} manifests in an explosive positive bubble behavior. When $b_t < 0$, $E_t(b_{t+1}) = \beta b_t < b_t$, indicating the presence of a downward explosive behavior in b_{t+1}, i.e. a negative bubble. When f_{t+1} does not involve any explosive behavior, the downward explosive behavior in b_{t+1} leads to an explosive negative bubble behavior in p_{t+1}. In this case p_{t+1} could still be positive as long as $f_{t+1} > -b_{t+1}$.

The testing and dating method for bubbles relies on the technique of fitting to time series data (i.e. prices adjusted by fundamental values) the following autoregressive (AR) model where the AR coefficient (β) may take a different value in a bubble (or even a crisis) period from that in a normal period:

$$y_t = \alpha + \beta y_{t-1} + \varepsilon_t. \tag{3}$$

To test $H_0 : \beta = 1$ against $H_1 : \beta > 1$, both the t-test and the coefficient-based test have been used. Both tests are based on the ordinary least squares (LS) estimate of β.[1] When the tests are implemented recursively, they can detect when the time series switches from a unit root model to an explosive model, and vice versa.

PSY (2015a) applied the proposed recursive method to a long time series data on the monthly S&P 500 stock price index-dividend ratio over the period from January 1871

[1]More recent studies have proposed methods based on the weighted LS estimate to deal with heteroscedasticity; see, for example, Harvey et al. (2017a).

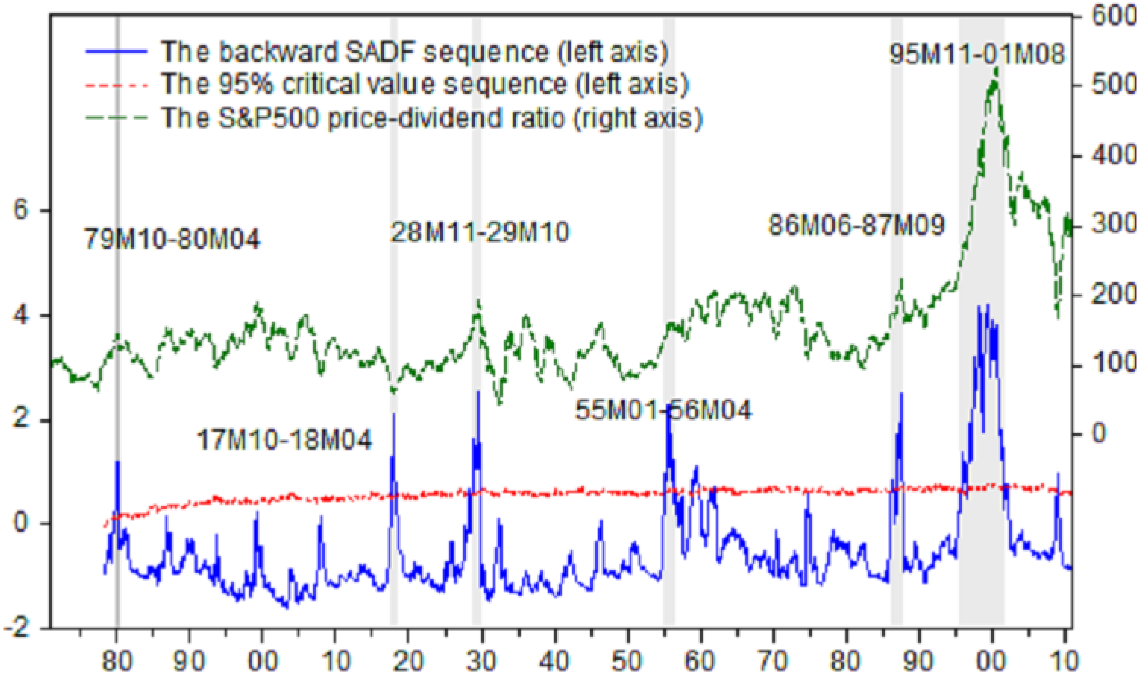

Figure 1: The price-dividend ratio of S&P500, the backward SADF of PSY, and the 95% critical value between January 1871 to December 2010.

to December 2010. Figure 1 reproduces PSY's Figure 7 where the original data, the sequence of the sup t-statistics, and the sequence of the 95% finite sample critical values are plotted. According to PSY, whenever the test statistic exceeds the critical value, it indicates a bubble episode. As a consequence, several well-known bubble episodes in this series have been identified, including the dot-com bubble (1995M11–2001M08).

Interestingly, the PSY strategy also identifies two periods of market downturns as bubble episodes, namely the 1917 stock market crash (1917M08-1918M04) and the subprime mortgage crisis (2009M02-M04). PSY conjectured that "the identification of crashes as bubbles may be caused by very rapid changes in the data". Some recent studies, such as Phillips and Shi (2017) and Harvey, et al. (2016, 2017b), modelled downturns using a stationary AR process with β in Model (3) smaller than 1. However, a stationary AR model is at odd with the testing result found in PSY which suggests $\beta > 1$. Instead of using a stationary AR process to describe downturns, one may argue that the downturns in y_t are generated by a negative bubble process (2), where $\beta > 1$ and $b_t < 0$.

To see the time series property in the downturns more clearly, we plot the monthly data between October 2006 and March 2009 in Figure 2. The plot seems to suggest an alternative model of a quadratic trend to describe the downturns, rather than a stationary AR model. A similar observation can be made for the data in the 1917 stock market crash period (1917M08-1918M04). These observations naturally raise two questions. First, if the time series data is actually generated by a downward quadratic trend model, is it possible for right-tailed unit root tests to reject the null hypothesis of unit root in favor

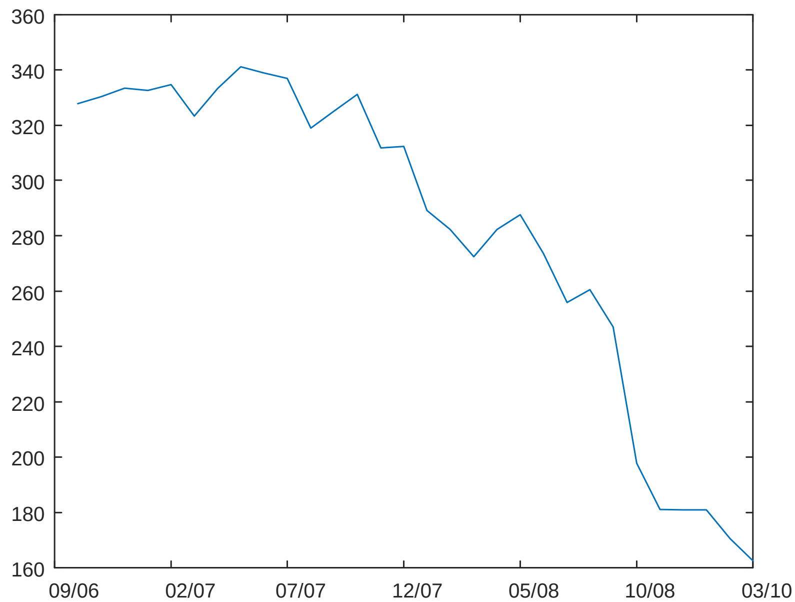

Figure 2: Monthly price-dividend ratio of S&P500 between October 2006 to March 2009.

of the explosive alternative? This question extends to the testing results of the explosive positive bubble. That is, when an explosive positive bubble is detected by a right-tailed unit root test, is it possible that the data is actually generated from a model with an upward polynomial trend? Second, if the answer to the first question is yes, how to modify the right-tailed unit root tests so that they are robust under deterministic trends. These two questions motivate us to study the ability of two right-tailed unit root tests based on AR regressions to distinguish the explosive process from the trend-stationary process, and to introduce robust unit root testing procedures.

In the paper, we first show that if the data is generated from a stationary process with a quadratic trend or a higher order trend (upward or downward), the LS estimate of β from regression (3) tends to be larger than 1 so that the right-tailed unit root tests (both the t-test and the coefficient-based test) often reject the unit root null hypothesis in favor of the explosive alternative. We then propose to construct the right-tailed unit root tests based on the following regression with an explicit requirement that $k \geq 1$:

$$y_t = \check{\alpha} + \check{\beta} y_{t-1} + \sum\nolimits_{i=1}^{k} \check{\psi}_i \Delta y_{t-i} + \check{e}_t. \tag{4}$$

It is proved that, when the data generating process (DGP) is trend-stationary, the estimate $\check{\beta}$ in regression (4) converges to 1 at a rate faster than $O_p(n)$. As a result, both the t-test and the coefficient-based test based on $\check{\beta}$ tend not to reject the null hypothesis of unit root. Whereas, when the data is generated by a mildly explosive process, the test statistics based on $\check{\beta}$ diverge to positive infinity, and hence, the right-tailed unit root tests reject the unit root null hypothesis in favor of the explosive alternative.

To see the difference in empirical implications of the tests based on (3) and on (4) with $k = 1$, we first apply the coefficient-based test recursively to the monthly S&P 500 stock price index-dividend ratio from January 1871 to December 2010. It is found that the test based on (4) with $k = 1$ detects the same bubbles as the test based on (3), including five positive bubble episodes and two negative bubble episodes. Since the right-tailed test based on (4) with $k = 1$ can better distinguish the explosive process from the trend-stationary process, we can conclude that the seven episodes including the two negative bubble episodes are not likely due to polynomial trends.

In the second empirical study, we apply the coefficient-based test recursively to the monthly price-rent ratio of the US housing market (the Case-Shiller US National Home Price Index divided by the rent of primary residence) for the period from January 1981 to June 2017. The test based on (3) detects two positive bubble episodes (1986M05 to 1990M04, and 1998M04 to 2008M01). Moreover, it also flags two collapsing periods as bubble episodes. However, the right-tailed coefficient-based test based on (4) with $k = 1$ detects only the two positive bubble periods, indicating that the data in the two collapsing periods may be better fitted by models with downward polynomial trends. In addition, the two positive bubble episodes are estimated to end much earlier by our proposed method than those based on (3). The estimated termination dates of the two positive bubbles by the proposed method match better with the turning points in the data.

The rest of the paper is organized as follows. Sections 2-4 introduce the stationary processes with different kinds of trends, including the linear trend, the quadratic trend, and the cubic trend. For each trend-stationary process, the large sample theory of the LS estimator of the AR coefficient and the two unit root test statistics are developed, when (misspecified) AR models are fitted. The large sample results show that the right-tailed unit root tests based on the regression (3) cannot distinguish the explosive process from the trend-stationary process. In Section 5 we show that the right-tailed unit root tests based on the regression (4) with $k \geq 1$ can successfully distinguish the mildly explosive process from the trend stationary process. Section 6 presents simulation evidence to support the asymptotic results. The proposed procedure is used to analyze two real time series in Section 7. Section 8 concludes. A brief summary and the proof of all theoretical results are included in the Appendix.

2 Linear Trend Model and Explosiveness

We first study the asymptotic performance of the right-tailed unit root tests based on the LS regressions of (4) with different values of k including $k = 0$, when the data is actually generated from the linear trend model as

$$y_t = \delta t + u_t, \quad t = 1, 2, ..., n, \tag{5}$$

where δ is a non-zero constant (positive or negative), $u_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$, $c_0 = 1$, $\sum_{j=0}^{\infty} j|c_j| < \infty$, and $\{\varepsilon_t\}$ is a sequence of independent and identically distributed (iid) random variables with $E(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \sigma^2$, and $E\left(\varepsilon_t^4\right) < \infty$. Let $\gamma_0 = Var(u_t)$ and $\gamma_1 = Cov(u_t, u_{t-1})$.

We consider the three LS regression equations:

$$y_t = \widehat{\beta} y_{t-1} + \hat{e}_t, \tag{6}$$

$$y_t = \tilde{\alpha} + \tilde{\beta} y_{t-1} + \tilde{e}_t, \tag{7}$$

$$y_t = \check{\alpha} + \check{\beta} y_{t-1} + \sum\nolimits_{i=1}^{k} \check{\psi}_i \Delta y_{t-i} + \check{e}_t, \tag{8}$$

where $\widehat{\beta}$, $\left(\tilde{\alpha}, \tilde{\beta}\right)$, and $\left(\check{\alpha}, \check{\beta}, \check{\psi}_1, \cdots, \check{\psi}_k\right)$ are the conventional LS estimators. We define the following regression t-statistics which are commonly used to test the unit root hypothesis against the explosive alternative:

$$t_{\widehat{\beta}} = \frac{\widehat{\beta} - 1}{se\left(\widehat{\beta}\right)}, \quad t_{\tilde{\beta}} = \frac{\tilde{\beta} - 1}{se\left(\tilde{\beta}\right)}, \quad \text{and} \quad t_{\check{\beta}} = \frac{\check{\beta} - 1}{se\left(\check{\beta}\right)},$$

where $se\left(\widehat{\beta}\right)$, $se\left(\tilde{\beta}\right)$ and $se\left(\check{\beta}\right)$ are the standard errors of $\widehat{\beta}$, $\tilde{\beta}$ and $\check{\beta}$, respectively. The coefficient-based statistics are defined as $n\left(\widehat{\beta} - 1\right)$, $n\left(\tilde{\beta} - 1\right)$ and $n\left(\check{\beta} - 1\right)$, respectively, which can also be used to test the unit root hypothesis against explosiveness.

Note that the process y_t defined in (5) can be rewritten as

$$y_t = \delta + y_{t-1} + \Delta u_t, \tag{9}$$

where $\Delta u_t = u_t - u_{t-1}$. Hence, the true DGP of y_t is not covered by the regression (6) for any value of $\widehat{\beta}$ due to the presence of an intercept in (9). Theorem 2.1 reports the asymptotic theory of the t-statistic and the coefficient-based statistic when the regression (6) is fitted to y_t.

Theorem 2.1 *When the LS regression (6) is fitted to the time series y_t generated from the linear trend model defined in (5), we have, as $n \to \infty$,*

(a) $n\left(\widehat{\beta} - 1\right) = 3/2 + O_p\left(n^{-1}\right)$;

(b) $t_{\widehat{\beta}}/\sqrt{n} \xrightarrow{p} \sqrt{3}\,|\delta| / \sqrt{\delta^2 + 8(\gamma_0 - \gamma_1)}$.

Theorem 2.1 suggests that, although $\widehat{\beta} \xrightarrow{p} 1$ at the rate of n, the average value of $\widehat{\beta}$ is greater than one. Moreover, $n\left(\widehat{\beta} - 1\right) \approx 1.5$ which is larger than 1.28, the 95% critical value of the Dickey-Fuller (DF) distribution.[2] In addition, the t-statistic $t_{\widehat{\beta}}$ diverges to $+\infty$ with the speed of $\sqrt{n}$. Hence, when the data is generated from a linear trend model, the coefficient-based statistic and, more so, the t-statistic in regression (6) tend to reject the unit root null hypothesis in favor of explosiveness, no matter what the sign of δ **is**.

[2] The DF distribution is given by $\int_0^1 W(r)\, dW(r) \,/\, \int_0^1 [W(r)]^2\, dr$ where $W(r)$ is a standard Brownian motion; see Hamilton (1994) for the textbook treatment and Table B.5 Case 1 in Hamilton (1994) for finite sample critical values and asymptotic critical values.

The regression equations (7) and (8), although misspecified, include the true DGP of y_t when particular values of the parameters are chosen: $\left(\tilde{\alpha}, \tilde{\beta}\right) = (\delta, 1)$ for regression (7), and $\left(\check{\alpha}, \check{\beta}, \check{\psi}_1, \cdots, \check{\psi}_k\right) = (\delta, 1, 0, \cdots, 0)$ for regression (8). Theorem 2.2 reports the asymptotic theory for the t-statistic and the coefficient-based statistic when the regression (7) is fitted to the data y_t.

Theorem 2.2 *When the LS regression (7) is fitted to the time series y_t generated from the linear trend model defined in (5), we have, as $n \to \infty$,*

(a) $n^2\left(\tilde{\beta} - 1\right) = -12\left(\gamma_0 - \gamma_1\right)/\delta^2 + 6\left(u_n + u_1\right)/\delta + o_p(1)$;

(b) $\sqrt{n}t_{\tilde{\beta}} = \frac{-6(\gamma_0-\gamma_1)+3\delta(u_n+u_1)}{|\delta|\sqrt{6(\gamma_0-\gamma_1)}} + o_p(1)$.

Theorem 2.2 suggests that $n\left(\tilde{\beta} - 1\right) \xrightarrow{p} 0$, indicating the null of unit root cannot be rejected if a positive asymptotic critical value is used.[3] Interestingly, it can be seen that $\tilde{\beta}$ has a downward bias as

$$E\left(\tilde{\beta}\right) = 1 - 12\left(\gamma_0 - \gamma_1\right)/(n^2\delta^2) + o\left(n^{-2}\right).$$

Hence, the average value of $t_{\tilde{\beta}}$ is expected to be negative. This is confirmed as

$$E\left(t_{\tilde{\beta}}\right) = \frac{-6\left(\gamma_0 - \gamma_1\right)}{|\delta|\sqrt{6n\left(\gamma_0 - \gamma_1\right)}} + o\left(n^{-1/2}\right).$$

Note that $t_{\tilde{\beta}}$ converges to zero, indicating the null of unit root cannot be rejected if a positive critical value is used.[4] When the critical values for both tests are chosen to be positive constants or to diverge to positive infinity at the same speed, since $n\left(\check{\beta} - 1\right)$ converges in probability to zero faster than $t_{\check{\beta}}$, we expect that the coefficient-based test has better asymptotic power than the t-test to distinguish the linear trend process from the explosive process.

Based on the large sample results reported in Theorem 2.3 below, similar conclusions can be made when the regression (8) with $k \geq 1$ is used to fit the time series y_t generated from the linear trend model (5). That is, both the coefficient-based test and the t-test tend not to reject the null hypothesis of unit root and the coefficient-based test has better asymptotic power than the t-test to distinguish the linear trend process from the explosive process.

Theorem 2.3 *When the LS regression (8) with $k \geq 1$ is fitted to the time series y_t generated from the linear trend model defined in (5), we have, as $n \to \infty$,*

(a) $n^2\left(\check{\beta} - 1\right) = O_p(1)$;

(b) $\sqrt{n}t_{\check{\beta}} = O_p(1)$,

where the form of the limiting distribution depends on the value of k in the regression.

[3] From Table B.5 in Hamilton (1994), the 97.5% and the 99% critical value are 0.41 and 1.04.

[4] From Table B.6 in Hamilton (1994), the 97.5% and the 99% critical value are 0.23 and 0.60.

3 Quadratic Trend Model and Explosiveness

In this section, we study the asymptotic performance of the right-tailed unit root tests based on the LS regressions of (6), (7) and (8), respectively, when the data $\{y_t\}$ is generated from a quadratic trend model as

$$y_t = \delta t^2 + u_t, \ t = 1, 2, ..., n, \tag{10}$$

where δ is a positive or negative constant, $\{u_t\}$ is a weakly stationary process as defined in (5). An equivalent representation of y_t is

$$y_t = -\delta + 2\delta t + y_{t-1} + \Delta u_t. \tag{11}$$

Due to the presence of both the intercept and the linear trend in (11), when the regression (6) is considered, no value of $\widehat{\beta}$ can cover the true DGP of y_t; when the regression (7) is considered, no value of $\left(\tilde{\alpha}, \tilde{\beta}\right)$ can cover the true DGP of y_t. However, since $\Delta y_t - \Delta y_{t-1} = \Delta^2 y_t = 2\delta + \Delta^2 u_t$, the DGP of y_t can be rewritten as

$$y_t = 2\delta + y_{t-1} + \Delta y_{t-1} + \Delta^2 u_t. \tag{12}$$

Therefore, the DGP of y_t is covered by the regression (8) with parameters $\left(\breve{\alpha}, \breve{\beta}, \breve{\psi}_1, \cdots, \breve{\psi}_k\right) = (2\delta, 1, 1, 0, \cdots, 0)$.

Theorems 3.1-3.3 report the asymptotic theory of the t-statistic and the coefficient-based statistic when regressions (6)-(8) are fitted to y_t, respectively.

Theorem 3.1 *When the LS regression (6) is fitted to the data $\{y_t\}$ generated from the quadratic trend model defined in (10), we have, as $n \to \infty$,*

(a) $n\left(\widehat{\beta} - 1\right) = 5/2 + O_p\left(n^{-1}\right)$;

(b) $t_{\widehat{\beta}}/\sqrt{n} \xrightarrow{p} \sqrt{15}$.

Theorem 3.1 shows that, on average $n\left(\widehat{\beta} - 1\right) \approx 5/2 > 2.03$, the 99% critical value of the DF distribution. In addition, the t-statistic diverges to $+\infty$ at the speed of $\sqrt{n}$. Note that the leading terms both in $\widehat{\beta} - 1$ and $t_{\widehat{\beta}}$ are independent of δ. Therefore, no matter what the sign of δ has, the coefficient-based test, and more so, the t-test in regression (6) tend to reject the unit root null hypothesis in favor of explosiveness.

Theorem 3.2 *When the LS regression (7) is fitted to the data $\{y_t\}$ generated from the quadratic trend model defined in (10), we have, as $n \to \infty$,*

(a) $n\left(\tilde{\beta} - 1\right) = 15/8 + O_p\left(n^{-1}\right)$;

(b) $t_{\tilde{\beta}}/\sqrt{n} \xrightarrow{p} \sqrt{15}$.

When the regression (7) is considered, the 99% asymptotic critical values of the DF coefficient-based test and the DF t-test are 1.04 and 0.6, respectively. However, the large sample results in Theorem 3.2 show that $n\left(\tilde{\beta}-1\right) \approx 1.875 > 1.04$, and $t_{\tilde{\beta}} \approx \sqrt{15n} \to +\infty$. As a result, for any value of δ (> 0 or < 0), both the coefficient-based test and the t-test in the regression (7) tend to reject the unit root null hypothesis in favor of explosiveness. It means that these two tests, especially the t-test, have no power to distinguish the quadratic trend process from the explosive process under the LS regression (7). Since the regression (7) is the one that is commonly used in the literature on testing and dating bubbles, this finding has important empirical implications.

Theorem 3.3 *When the LS regression (8) with $k \geq 1$ is fitted to the data $\{y_t\}$ generated from the quadratic trend model defined in (10), we have, as $n \to \infty$,*

(a) $n^3\left(\check{\beta}-1\right) = O_p(1)$;

(b) $\sqrt{n}t_{\check{\beta}} = O_p(1)$,

where the form of the limiting distribution depends on the value of k in the regression (8).

The large sample results reported in Theorem 3.3 show that $n\left(\check{\beta}-1\right) \xrightarrow{p} 0$ and $t_{\check{\beta}} \xrightarrow{p} 0$, which suggest that both the coefficient-based test and the t-test in regression (8) with $k \geq 1$ tend not to reject the unit root null hypothesis if the critical values for both tests are chosen to be positive constants or to diverge to positive infinity. Since $n\left(\check{\beta}-1\right)$ converges in probability to zero faster than $t_{\check{\beta}}$, we expect that the coefficient-based test has better asymptotic power than the t-test to distinguish the quadratic trend process from the explosive process under the LS regression (8) with $k \geq 1$.

4 Cubic Trend Model and Explosiveness

In this section, we assume the data $\{y_t\}$ is generated from a cubic trend model as

$$y_t = \delta t^3 + u_t, \ t = 1, 2, ..., n, \tag{13}$$

where δ is a non-zero constant and $\{u_t\}$ is a weakly stationary process defined as in (6).

An equivalent representation of the DGP of y_t is

$$y_t = 3\delta t^2 - 3\delta t + \delta + y_{t-1} + \Delta u_t. \tag{14}$$

Due to the presence of the intercept, the linear trend, and the quadratic trend in (14), the regression (6) does not cover the true DGP of y_t for any value of $\widehat{\beta}$, and the regression (7) does not cover the true DGP of y_t for any value of $\left(\tilde{\alpha}, \tilde{\beta}\right)$.

Note that $\Delta^2 y_t = 6\delta(t-1) + \Delta^2 u_t$, which leads to

$$y_t = -6\delta + y_{t-1} + \Delta y_{t-1} + 6\delta t + \Delta^2 u_t.$$

Together with the result of $\Delta^2 y_{t-1} = 6\delta(t-2) + \Delta^2 u_{t-1}$, an alternative representation of the DGP of y_t can be obtained as

$$y_t = 6\delta + y_{t-1} + 2\Delta y_{t-1} - \Delta y_{t-2} + \Delta^3 u_t. \tag{15}$$

Hence, the true DGP is covered by the regression (8) with $k \geq 2$ and $\left(\check{\alpha}, \check{\beta}, \check{\psi}_1, \check{\psi}_2, \cdots, \check{\psi}_k\right) = (6\delta, 1, 2, -1, 0, \cdots, 0)$. Since $\Delta y_{t-1} = 3\delta t^2 + O_p(t)$ and $\Delta y_{t-2} = 3\delta t^2 + O_p(t)$, the regression (8) with $k \geq 2$ faces with the problem of asymptotic perfect collinearity. However, as we prove in the Appendix, this problem plays no effect on the large sample property of $\check{\beta}$.

Theorems 4.1-4.3 report the asymptotic theory of the t-statistic and the coefficient-based statistic when regressions of (6)-(8) are fitted, respectively, to the data $\{y_t\}$ generated from the cubic trend model (13).

Theorem 4.1 *When the LS regression (6) is fitted to the data $\{y_t\}$ generated from the cubic trend model (13), we have, as $n \to \infty$,*

(a) $n\left(\hat{\beta} - 1\right) = 7/2 + O_p\left(n^{-1}\right)$;

(b) $t_{\hat{\beta}}/\sqrt{n} \xrightarrow{p} \sqrt{35}$.

Theorem 4.1 shows that, $n\left(\widehat{\beta} - 1\right) \approx 7/2 > 2.03$, the 99% critical value of the DF distribution. Moreover, the t-statistic $t_{\widehat{\beta}}$ diverges towards $+\infty$ at the speed of $\sqrt{n}$. Therefore, no matter what sign δ has, the coefficient-based test and, more so, the t-test in the regression (6) tend to reject the unit root null hypothesis in favor of explosiveness.

Theorem 4.2 *When the LS regression (7) is fitted to the data $\{y_t\}$ generated from the cubic trend model (13), we have, as $n \to \infty$,*

(a) $n\left(\tilde{\beta} - 1\right) = 28/9 + O_p\left(n^{-1}\right)$;

(b) $t_{\tilde{\beta}}/\sqrt{n} \xrightarrow{p} \sqrt{35}$.

From Theorem 4.2, it can be seen that $t_{\tilde{\beta}} \approx \sqrt{35n} \to +\infty$ as $n \to \infty$, and $n\left(\tilde{\beta} - 1\right) \approx 28/9 > 1.04$, the 99% critical value of the DF distribution. Thus, for any non-zero value of δ (> 0 or < 0), both the coefficient-based test, and more so, the t-test in regression (7) tend to reject the unit root null hypothesis in favor of explosiveness.

Theorem 4.3 *When the LS regression (8) is fitted to the data $\{y_t\}$ generated from the cubic trend model (13), we have, as $n \to \infty$,*

(a) when $k = 1$,

$$n^2\left(\check{\beta} - 1\right) = -8.4 + O_p\left(n^{-1}\right), \quad \textit{and} \quad t_{\check{\beta}}/\sqrt{n} = -\sqrt{21}/2 + O_p\left(n^{-1}\right);$$

(b) when $k > 1$,

$$n^4\left(\check{\beta} - 1\right) = O_p(1), \quad \textit{and} \quad \sqrt{n}t_{\check{\beta}} = O_p(1).$$

where the limits of $\check{\beta}$ and $t_{\check{\beta}}$ in (b) depends on the value of k in the regression (8).

From (15), it is known that the true DGP of y_t is covered by the regression (8) with $k>1$. Part (b) of Theorem 4.3 shows that $n\left(\breve{\beta}-1\right) \xrightarrow{p} 0$ and $t_{\breve{\beta}} \xrightarrow{p} 0$. Hence, both the coefficient-based test and the t-test from the regression (8) with $k>1$ tend not to reject the unit root null hypothesis if the critical values for both tests are chosen to be positive constants or to diverge to positive infinity. Since $n\left(\breve{\beta}-1\right)$ converges in probability to zero much faster than $t_{\breve{\beta}}$, we expect that the coefficient-based test has better asymptotic power than the t-test to distinguish the cubic trend process from the explosive process.

Also revealed by the representation in (15) is that the regression (8) with $k=1$ does not cover the true DGP of y_t. Part (a) of Theorem 4.3 shows that, although $\breve{\beta}$ converges to 1 less quickly than it does in the regression (8) with $k>1$, it still goes to 1 fast enough to ensure $n\left(\breve{\beta}-1\right) \xrightarrow{p} 0$. Moreover, the leading term of $n\left(\breve{\beta}-1\right)$ is negative. Therefore, the coefficient-based test tends not to reject the unit root null hypothesis if a positive critical value is used. It is also shown that $t_{\breve{\beta}}$ diverges to negative infinity at the speed of $n^{1/2}$, which implies that the t-test tends not to reject the unit root null hypothesis.

It is not difficult to extend the theory to the case where the true DGP has a deterministic trend with an order higher than 3. In particular, if the LS regression (8) with $k=1$ is applied to fit the data, it can still be shown that $n\left(\breve{\beta}-1\right) \xrightarrow{p} 0$ and $t_{\breve{\beta}} \xrightarrow{p} -\infty$. Hence, both the t-test and the coefficient-based test tend not to reject the unit root null hypothesis.

5 A Proposed Regression

As we have proved in earlier sections, if the true DGP is a trend-stationary process, the right-tailed unit root tests based on the regression model (8) with $k \geq 1$ tend not to find the evidence of explosiveness and hence can distinguish a bubble behavior from a trend-stationary behavior. This finding motivates us to propose the following procedure to test the unit root null hypothesis against explosiveness:

$$y_t = \breve{\alpha} + \breve{\beta} y_{t-1} + \sum_{i=1}^{k} \breve{\psi}_i \Delta y_{t-i} + \breve{e}_t \quad \text{with} \quad k \geq 1, \tag{16}$$

where the minimum value that k can take is set to be 1, namely at least one term involving the lagged Δy_t is included in the regression. As shown in the following Theorem 5.1, the right-tailed unit root tests (the coefficient-based test and the t-test) based on the LS regression of (16) are able to distinguish the explosive AR process from the trend-stationary process.

Theorem 5.1 *Let $W(r)$ denote a standard Brownian motion, and $u_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$ be the weakly stationary process as defined in (5) with $\varepsilon_t \overset{iid}{\sim} N(0,\sigma^2)$.*
(a) When the DGP is $y_t = y_{t-1} + \varepsilon_t$, the LS regression (16) with $k \geq 1$ leads to

$$n\left(\breve{\beta}-1\right) \Rightarrow \frac{\int_0^1 W(r)\, dW(r) - W(1)\int_0^1 W(r)\, dr}{\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r)\, dr\right)^2}, \tag{17}$$

and

$$t_{\breve{\beta}} = \frac{\breve{\beta}-1}{se\left(\breve{\beta}\right)} \Rightarrow \frac{\int_0^1 W(r)\, dW(r) - W(1)\int_0^1 W(r)\, dr}{\left\{\int_0^1 [W(r)]^2 dr - \left(\int_0^1 W(r)\, dr\right)^2\right\}^{1/2}}, \tag{18}$$

as $n \to \infty$*;*
(b) When the DGP is an ARIMA$(p,1,0)$ *process with* $p \geq 1$*, the statistics* $n\left(\breve{\beta}-1\right)$ *and* $t_{\breve{\beta}}$ *based on the LS regression (16) with* $k \geq p$ *have the same limiting distributions as those given in (a), respectively;*
(c) When the DGP is $y_t = \delta t^m + u_t$ *with* $m = 1$ *or* 2 *or* 3*,..., the statistics* $n\left(\breve{\beta}-1\right)$ *and* $t_{\breve{\beta}}$ *based the LS regression (16) with* $k \geq 1$ *have the limits as*

$$n\left(\breve{\beta}-1\right) \xrightarrow{p} 0 \quad and \quad t_{\breve{\beta}} \xrightarrow{p} 0 \quad or \ -\infty;$$

(d) When the DGP is mildly explosive, i.e., $y_t = \rho_n y_{t-1} + u_t$ *with* $\rho_n = 1 + c/n^{\theta}$*,* $c > 0$*, and* $\theta \in (0,1)$*, the statistics* $n\left(\breve{\beta}-1\right)$ *and* $t_{\breve{\beta}}$ *based on the LS regression (16) with* $k \geq 1$ *have the limits as*

$$n\left(\breve{\beta}-1\right) \to +\infty \quad and \quad t_{\breve{\beta}} \to +\infty.$$

The results reported in Part (c) of Theorem 5.1 show that, if the data is generated from a trend-stationary model, regardless of the direction of the trend, $n\left(\breve{\beta}-1\right) \xrightarrow{p} 0$ and $t_{\breve{\beta}} \xrightarrow{p} 0$ or $-\infty$. Hence, the right-tailed unit root tests from the regression (16) do not reject the unit root null hypothesis when the critical values for both tests are chosen to be positive constants or to diverge to positive infinity. Whereas, as reported in Part (d) of Theorem 5.1, if the data is generated from a mildly explosive process, $n\left(\breve{\beta}-1\right) \to +\infty$ and $t_{\breve{\beta}} \to +\infty$. Hence, both the coefficient-based test and the t-test should reject the unit root null hypothesis in favor of explosive alternative. As a result, both the two right-tailed unit root tests based on the regression (16) with $k \geq 1$ can distinguish the mildly explosive process from the trend-stationary process.

6 Simulation Studies

6.1 Linear trend model

Tables 1-3 report some simulation results based on 10,000 replications for the LS regressions (6), (7), and (8) with $k = 1$, respectively. For each regression, the true DGP is the linear trend model as in (5) with $\delta = -2$ or -4, $u_t \overset{iid}{\sim} N(0,1)$, and $n = 20$ or 50. Unless explicitly stated, we always report the mean and variance of the LS estimator of the AR(1) coefficient, the minimum, mean and maximum of the coefficient-based statistic (i.e. c-stat), and the minimum, mean and maximum of the t-statistic (i.e. t-stat). Also reported are the 99% finite sample critical values of the c-stat and t-stat which are simulated under the random walk null hypothesis.

Table 1 reports the simulation results for the regression (6), an AR(1) regression without intercept. Some conclusions can be made from Table 1. First, the average values of $\widehat{\beta}$

Table 1: Statistical results for the LS regression (6) when the data is generated from a linear trend model as in (5). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\widehat{\beta}$	1.0749	1.0764	1.0300	1.0302
variance of $\widehat{\beta}$	1.76e-05	4.32e-06	3.90e-07	9.67e-08
minimum of c-stat	1.1426	1.3092	1.3473	1.4195
mean of c-stat	1.4226	1.4517	1.4696	1.4810
maximum of c-stat	1.7351	1.6024	1.5794	1.5378
99% CV of c-stat	2.37	2.37	2.16	2.16
minimum of t-stat	2.1047	4.2141	4.7689	8.0217
mean of t-stat	4.3897	6.3537	7.0046	10.0048
maximum of t-stat	8.4469	8.7309	9.3513	11.7077
99% CV of t-stat	2.24	2.24	2.08	2.08

in all four cases are greater than 1, although getting closer to 1 as n increases. Second, the average values of the c-stat in all four cases are around 1.5, which is the value suggested by the asymptotic theory given in Theorem 2.1. In addition, except for the case where $\delta = -2$ and $n = 20$, the minimum values of the c-stat are all larger than 1.28 which is the 95% asymptotic critical value of the DF distribution. Hence, if this critical value is used, the unit root null hypothesis will always be rejected in favor of the explosive alternative. However, when the 99% finite sample critical values are used, the unit root null hypothesis will not be rejected. Third, except for the case with $\delta = -2$ and $n = 20$ in which the minimum value of the t-stat is slightly smaller than the corresponding 99% finite sample critical value, the minimum values of the t-stat in other cases are all larger than the corresponding 99% finite sample critical values, indicating that the unit root null hypothesis is always rejected in favor of the explosive alternative. Fourth, the simulated average values of the t-stat are consistent with that suggested by the asymptotic theory given in Theorem 2.1. For example, according to the asymptotic theory, when $u_t \overset{iid}{\sim} N(0,1)$ and $\delta = -2$, we should have $t_{\widehat{\beta}}/\sqrt{n} \xrightarrow{p} 2\sqrt{3}/\sqrt{4+8} = 1$, suggesting that $t_{\widehat{\beta}} \approx \sqrt{n}$. Hence, $t_{\widehat{\beta}} \approx 4.47$ when $n = 20$, and $t_{\widehat{\beta}} \approx 7.07$ when $n = 50$. These values are very close to the simulated average values of the t-stat reported in Table 1.

Table 2 reports the simulation results for the regression (7), an AR(1) regression with an intercept. Some conclusions can be made from Table 2. First, $\tilde{\beta}$ converges to one very quickly. Second, $\tilde{\beta}$ has a small downward bias. This explains why the average values of the c-stat and the t-stat are negative. Third, the maximum values of the c-stat in all cases are smaller than the corresponding 99% finite-sample critical values, indicating that the unit root null hypothesis will not be rejected in favor of the explosive alternative. Fourth, although the maximum values of the t-stat in all cases are larger than the respective 99% finite-sample critical values, their average values are much smaller than these critical

Table 2: Statistical results for the LS regression (7) when the data is generated from a linear trend model as in (5). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\tilde{\beta}$	0.9919	0.9980	0.9988	0.9997
variance of $\tilde{\beta}$	1.44e-04	3.47e-05	3.24e-06	7.94e-07
minimum of c-stat	-1.0523	-0.4523	-0.4051	-0.1846
mean of c-stat	-0.1545	-0.0389	-0.0610	-0.0153
maximum of c-stat	0.6914	0.3951	0.2610	0.1425
99% CV of c-stat	1.40	1.40	1.22	1.22
minimum of t-stat	-2.0613	-1.8580	-1.2086	-1.1304
mean of t-stat	-0.2777	-0.1397	-0.1737	-0.0868
maximum of t-stat	1.5702	1.6110	0.8810	0.9641
99% CV of t-stat	0.80	0.80	0.66	0.66

values. Hence, the t-test will not reject the unit root null hypothesis most of the time. However, it can be seen that the coefficient-based test is more powerful than the t-test to distinguish the linear trend process from the explosive process in finite samples. Fifth, the simulation results are consistent with what suggested by the asymptotic theory given in 2.2. For example, when $u_t \overset{iid}{\sim} N(0,1)$, $n = 50$ and $\delta = -4$, the asymptotic theory suggests that $E\left(\tilde{\beta}\right) = 1 - 12/\left(n^2\delta^2\right) + o\left(n^{-2}\right) \approx 0.9997$ and $E\left(t_{\tilde{\beta}}\right) = -\frac{\sqrt{6}}{|\delta|\sqrt{n}} + o\left(1/\sqrt{n}\right) \approx -0.0866$. These values are nearly identical to what we obtained in simulations.

Table 3 reports the simulation results for the regression (8) with $k = 1$. Conclusions similar as those obtained from Table 2 can be made from Table 3. The most important one is that $\check{\beta}$ converges to 1 very quickly making the c-stat taking values around 0 with a small variation. As a result, the right-tailed unit root test based on the c-stat cannot reject the unit root null hypothesis, and hence, has power to distinguish the linear trend process from the explosive process.

6.2 Quadratic trend model

Tables 4-6 report some simulation results based on 10,000 replications for the LS regressions (6), (7), and (8) with $k = 1$, respectively. For each regression, the true DGP is the quadratic trend model as in (10) with $\delta = -2$ or -4, $u_t \overset{iid}{\sim} N(0,1)$, and $n = 20$ or 50.

Table 4 reports the simulation results for the regression (6), an AR(1) model without intercept. Some conclusions can be made from Table 4. First, the average value of $\widehat{\beta}$ is greater than 1 in all four cases, although gets closer to 1 in the cases with larger n. The average value of $\widehat{\beta}$ is very close to $1 + 2.5/n$, a value predicted by our asymptotic theory given in Theorem 3.1. Second, the minimum values of the c-stat in all cases are larger than the respective 99% finite sample critical values. Hence, the coefficient-based test always indicates explosiveness in unit root testing. Third, consistent with the asymptotic theory

Table 3: Statistical results for the LS regression (8) with $k = 1$ when the data is generated from a linear trend model as in (5). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\check{\beta}$	0.9955	0.9988	0.9994	0.9998
variance of $\check{\beta}$	2.15e-04	5.35e-05	4.23e-06	1.05e-06
minimum of c-stat	-0.9786	-0.4748	-0.4051	-0.1939
mean of c-stat	-0.0815	-0.0207	-0.0308	-0.0076
maximum of c-stat	1.0156	0.5005	0.3478	0.1812
99% CV of c-stat	2.29	2.29	1.50	1.50
minimum of t-stat	-1.9943	-1.9186	-1.2923	-1.2605
mean of t-stat	-0.1566	-0.0804	-0.0983	-0.0489
maximum of t-stat	2.1857	2.2014	1.3951	1.4323
99% CV of t-stat	0.92	0.92	0.72	0.72

Table 4: Statistical results for the LS regression (6) when the data is generated from a quadratic trend model as in (10). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\widehat{\beta}$	1.1327	1.1327	1.0512	1.0512
variance of $\widehat{\beta}$	1.32e-07	3.30e-08	4.46e-10	1.11e-10
minimum of c-stat	2.4966	2.5091	2.5042	2.5063
mean of c-stat	2.5212	2.5214	2.5084	2.5084
maximum of c-stat	2.5475	2.5345	2.5124	2.5104
99% CV of c-stat	2.37	2.37	2.16	2.16
minimum of t-stat	15.0876	15.5004	26.2124	26.3187
mean of t-stat	15.7351	15.8186	26.3792	26.4052
maximum of t-stat	16.4100	16.1522	26.5231	26.4794
99% CV of t-stat	2.24	2.24	2.08	2.08

Table 5: Statistical results for the LS regression (7) when the data is generated from a quadratic trend model as in (10). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\tilde{\beta}$	1.0944	1.0944	1.0376	1.0376
variance of $\tilde{\beta}$	3.70e-07	9.26e-08	1.27e-09	3.18e-10
minimum of c-stat	1.7491	1.7712	1.8353	1.8388
mean of c-stat	1.7929	1.7932	1.8422	1.8422
maximum of c-stat	1.8367	1.8151	1.8494	1.8458
99% CV of c-stat	1.40	1.40	1.22	1.22
minimum of t-stat	14.8184	16.6229	26.9877	27.3744
mean of t-stat	17.2173	17.6987	27.5395	27.6500
maximum of t-stat	19.7671	19.0951	28.0039	27.8771
99% CV of t-stat	0.80	0.80	0.66	0.66

given in Theorem 3.1, in all cases, the t-stat takes values around $\sqrt{15n}$ with a very small variation. In addition, the minimum values of the t-stat in all cases are larger than the respective 99% finite sample critical values. Hence, the t-test always rejects the unit root null hypothesis in favor of explosive alternative. All these findings corroborate the large sample theory given in Theorem 3.1.

Table 5 reports the simulation results for the regression (7), an AR(1) model with an intercept. Similar conclusions to those from Table 4 can be made from Table 5. First, the average value of $\tilde{\beta}$ is greater than 1 in all four cases, although it gets closer to 1 as n increases. Second, consistent with the asymptotic theory given in Theorem 3.2, the c-stat takes values around $15/8 \approx 1.8750$ with a very small variation so that its minimum values in all cases are larger than the respective 99% finite sample critical values. Hence, the coefficient-based test always suggests explosiveness in unit root testing. Third, as suggested by the asymptotic theory, the t-stat in every case takes values around $\sqrt{15n}$ with a small variation. As $\sqrt{15n}$ is much larger than the respective 99% finite sample critical values, the t-test always rejects the unit root null hypothesis in favor of the explosive alternative. These findings from Table 5 have important empirical implications as the implementation of the right-tailed unit root testing in the literature has been based on the LS regression (7). When the data is generated from a quadratic trend model, our findings suggest that the unit root tests based on the regression (7) always suggest explosiveness.

Table 6 reports the simulation results for the regression (8) with $k = 1$. Several conclusions are made from Table 6. First, $\breve{\beta}$ converges to 1 very quickly so that the c-stat takes values around 0 with small variations in all cases. The maximum values of the c-stat in all cases are smaller than the respective 99% finite sample critical values. Hence, the right-tailed coefficient-based test from the regression (8) with $k = 1$ does not reject the null hypothesis of unit root. Second, the t-stat slowly goes to zero when n increases,

Table 6: Statistical results for the LS regression (8) with $k = 1$ when the data is generated from a quadratic trend model as in (10). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\breve{\beta}$	1.0113	1.0030	1.0006	1.0002
variance of $\breve{\beta}$	4.22e-05	7.98e-06	1.01e-07	2.030e-08
minimum of c-stat	-0.1857	-0.1301	-0.0273	-0.0189
mean of c-stat	0.2042	0.0548	0.0299	0.0075
maximum of c-stat	0.8005	0.3052	0.0933	0.0373
99% CV of c-stat	2.29	2.29	1.50	1.50
minimum of t-stat	-1.1180	-1.3639	-0.6453	-0.8976
mean of t-stat	0.9370	0.4745	0.5872	0.2936
maximum of t-stat	2.6522	2.1970	1.5824	1.2862
99% CV of t-stat	0.92	0.92	0.72	0.72

as suggested by the asymptotic theory given in Theorem 3.3. In most of the cases, the average value of the t-stat is smaller than the corresponding 99% critical value, indicating that the t-test has a good chance not to reject the unit root hypothesis. However, the probability for the t-test to suggest explosiveness is non-negligible in finite samples because the maximum values of the t-stat in most of the cases are considerably larger than the respective 99% finite sample critical values.

To further understand the behavior of the t-stat, Table 7 reports extra simulation results of the t-stat when the data is generated from the quadratic trend model as in (10) with $\delta = -4$, $n = 100$ or 250. It can be seen that, as n gets larger, the probability for the t-test to find explosiveness becomes smaller. And in the case where $\delta = -4$ and $n = 250$, the maximum value of the t-stat becomes smaller than the corresponding 99% critical value, and hence, the t-test does not find explosiveness any more. Comparing the t-stat with the c-stat, we conclude that the coefficient-based test is more able to distinguish the quadratic trend process from the explosive process in finite samples.[5]

6.3 Cubic trend model

Tables 8-10 report some simulation results based on 10,000 replications for the LS regressions (6), (7), and (8) with $k = 1$ and $k = 2$, respectively. For each regression, the true DGP is the cubic trend model as in (13) with $\delta = -0.02$ or -0.04, $u_t \overset{iid}{\sim} N(0,1)$, and $n = 20$ or 50.

Table 8 reports the simulation results for the regression (6), an AR(1) model without

[5]This conclusion is consistent with that suggested by the large sample theory given in Theorem 3.3, where it is shown that the t-stat $= O_p\left(1/\sqrt{n}\right)$ and the c-stat $= O_p\left(1/n^2\right)$. In fact, when $\delta = -2$, the ratio of the average value of the t-stat for $n = 20$ and for $n = 50$ is $0.9370/0.5872 = 1.5957$ which is very close to $\sqrt{50/20} = 1.5811$, reinforcing the result derived in Theorem 3.3 about the $\sqrt{n}$-convergence of $t_{\breve{\beta}}$.

Table 7: Further statistical results of the t-statistic from the regression (8) with $k = 1$ when data is generated from the quadratic trend model (10).

δ	-4	-4
n	100	250
minimum of t-stat	-0.4716	-0.2782
mean of t-stat	0.2079	0.1298
maximum of t-stat	0.9289	0.5717
99% CV of t-stat	0.66	0.63

Table 8: Statistical results for the LS regression (6) when the data is generated from a cubic trend model as in (13). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-0.02	-0.04	-0.02	-0.04
n	20	20	50	50
mean of $\widehat{\beta}$	1.1907	1.1909	1.0724	1.0725
variance of $\widehat{\beta}$	7.18e-06	1.79e-06	3.64e-09	9.10e-10
minimum of c-stat	3.4441	3.5376	3.5379	3.5440
mean of c-stat	3.6227	3.6278	3.5500	3.5501
maximum of c-stat	3.8245	3.7275	3.5614	3.5558
99% CV of c-stat	2.37	2.37	2.16	2.16
minimum of t-stat	12.3593	18.1769	38.8708	39.5090
mean of t-stat	18.8805	22.1239	39.6705	39.9027
maximum of t-stat	25.1421	25.6325	40.3300	40.2198
99% CV of t-stat	2.24	2.24	2.08	2.08

Table 9: Statistical results for the LS regression (7) when the data is generated from a cubic trend model as in (13). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-0.02	-0.04	-0.02	-0.04
n	20	20	50	50
mean of $\tilde{\beta}$	1.1654	1.1659	1.0639	1.0639
variance of $\tilde{\beta}$	1.41e-05	3.52e-06	7.26e-09	1.81e-09
minimum of c-stat	2.8777	3.0176	3.1126	3.1208
mean of c-stat	3.1428	3.1519	3.1288	3.1289
maximum of c-stat	3.4192	3.2883	3.1458	3.1374
99% CV of c-stat	1.40	1.40	1.22	1.22
minimum of t-stat	8.7165	14.6320	38.2528	39.6795
mean of t-stat	15.8692	21.1515	39.8576	40.4008
maximum of t-stat	27.4502	27.7903	41.1378	41.0256
99% CV of t-stat	0.80	0.80	0.66	0.66

intercept. Some conclusions can be made from Table 8. First, the average value of $\widehat{\beta}$ is greater than 1 in all four cases, and gets closer to 1 as n increases. Second, in all cases, the c-stat takes values around 3.5 with a very small variation, a value predicted by our asymptotic theory given in Theorem 4.1. Moreover, the minimum values of the c-stat in all cases are larger than the respective 99% finite sample critical values. Hence, the coefficient-based test always indicates explosiveness in unit root testing. Third, consistent with the asymptotic theory given in Theorem 4.1, in all cases, the t-stat takes values around $\sqrt{35n}$ with a small variation. Consequently, the minimum values of the t-stat in all cases are larger than the respective 99% finite sample critical values. Hence, the t-test always rejects the unit root null hypothesis in favor of explosive alternative. All these findings corroborate the large sample theory given in Theorem 4.1.

Table 9 reports the simulation results for the regression (7), an AR(1) model with an intercept. Similar conclusions to those from Table 7 can be made from Table 9. First, the average value of $\tilde{\beta}$ is greater than 1 in all four cases, and get closer to 1 as n increases. Second, consistent with the asymptotic theory given in Theorem 4.2, the c-stat takes values around $28/9 \approx 3.1111$ with a very small variation so that its minimum values in all cases are larger than the respective 99% finite sample critical values. Hence, the coefficient-based test always suggests explosiveness in unit root testing. Third, as suggested by the asymptotic theory, the t-stat in every case takes values around $\sqrt{35n}$ with a small variation. As $\sqrt{35n}$ is much larger than the respective 99% finite sample critical values, the t-test also always find explosiveness in unit root testing. All these findings corroborate the large sample theory given in Theorem 4.2.

Table 10 reports the simulation results for the regression (8) with $k = 1$. Some conclusions different with those from Tables 7-8 can be made from Table 9. First, as predicted by Part (1) of Theorem 4.3, $\check{\beta}$ converges to 1 very quickly. When n and the absolute value

Table 10: Statistical results for the LS regression (8) with $k = 1$ when the data is generated from a cubic trend model as in (13). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-0.02	-0.04	-0.02	-0.04	-0.02	-0.04
n	20	20	50	50	250	250
mean of $\check{\beta}$	1.1491	1.0626	1.0005	0.9976	0.9999	0.9999
variance of $\check{\beta}$	0.0016	8.83e-04	1.67e-06	2.43e-07	1.39e-12	3.44e-13
minimum of c-stat	-0.4477	-0.6767	-0.1719	-0.1955	-0.0341	-0.0338
mean of c-stat	2.6846	1.1275	0.0222	-0.1141	-0.0331	-0.0333
maximum of c-stat	5.2222	3.3061	0.2875	-0.0136	-0.0320	-0.0327
99% CV of c-stat	2.29	2.29	1.50	1.50	1.02	1.02
minimum of t-stat	-0.6116	-1.5677	-1.5055	-3.3262	-14.6752	-24.1792
mean of t-stat	3.2143	1.5686	0.1023	-1.4358	-11.6925	-20.4927
maximum of t-stat	7.2216	4.0850	1.3016	-0.1418	-9.4362	-17.4079
99% CV of t-stat	0.92	0.92	0.72	0.72	0.63	0.63

of δ are reasonably large, $\check{\beta}$ has downward bias. Second, the c-stat gets closes to zero as n increases whereas the t-stat diverges as n increases. When $n = 50$, the maximum values of the c-stat are smaller than the respective 99% finite sample critical values, indicating that the coefficient-based test will not suggest explosiveness in unit root testing regardless of $\delta = -0.02$ or -0.04. Third, in the cases where $n = 50$ and $\delta = -0.04$, the maximum values of the t-stat are smaller than the respective 99% finite sample critical values, indicating that the t-test will not reject unit root null hypothesis. However, when $n = 20$ or when $n = 50$ and $\delta = -0.02$, the t-test has to reject the null hypothesis of unit root in favor of explosiveness for some replications. Hence, the coefficient-based test is more able to distinguish the cubic trend process from the explosive process than the t-test in finite samples.

To further understand the behavior of the two statistics under the regression (8) with $k = 1$, we have done extra calculations for $n = 250$ and report the results in the last two columns of Table 10. The average value of $\check{\beta}$ becomes smaller than one. Furthermore, the average values of the two statistics are close to what Part (1) of Theorem 4.3 predicts. For example, the average value of the coefficient-based statistics is close to $-8.4/n = -0.0336$. The average value of the t-statistics is closer to $-\sqrt{21 \times n}/2 \approx -36$. Both tests cannot reject the null hypothesis of unit root in all cases.

Table 11 reports the simulation results for the regression (8) with $k = 2$. It can be seen that $\check{\beta} - 1$, the c-stat, and the t-stat converge to zero very fast. As a result, both the coefficient-based test and the t-test will not find evidence for explosiveness. These findings corroborate the large sample theory reported in Part (2) of Theorem 4.3.

Table 11: Statistical results for the LS regression (8) with $k = 2$ when the data is generated from a cubic trend model as in (13). The 99% finite sample critical values of the test statistics are simulated under the random walk null hypothesis.

δ	-2	-4	-2	-4
n	20	20	50	50
mean of $\check{\beta}$	0.9914	0.9973	0.9998	1.0000
variance of $\check{\beta}$	1.41e-05	3.24e-06	4.61e-09	7.79e-10
minimum of c-stat	-0.4015	-0.1927	-0.0225	-0.0071
mean of c-stat	-0.1464	-0.0465	-0.0075	-0.0019
maximum of c-stat	0.0610	0.0626	0.0024	-0.0031
99% CV of c-stat	3.7344	3.7344	1.9928	1.9928
minimum of t-stat	-4.0344	-2.7740	–2.1171	-1.6387
mean of t-stat	-1.6916	-0.8830	-0.9371	-0.4696
maximum of t-stat	1.0553	1.3609	0.4687	0.7518
99% CV of t-stat	1.1982	1.1982	0.8499	0.8499

Table 12: The percentiles of the coefficient-based statistic and the t-statistic based on the LS regression (16) with $k = 1$ when the true DGP is $y_t = y_{t-1} + u_t$.

n	90		95		99	
	c-stat	t-stat	c-stat	t-stat	c-stat	t-stat
20	-0.5299	-0.2422	0.3415	0.1435	2.2865	0.9240
50	-0.7389	-0.3743	0.0342	0.0175	1.4955	0.7208
100	-0.8221	-0.4180	-0.0520	-0.0337	1.2574	0.6592
250	-0.8144	-0.4058	-0.0957	-0.0556	1.0203	0.6283
500	-0.8222	-0.4128	-0.1448	-0.0844	1.0641	0.5921
∞	-0.85	-0.44	-0.13	-0.07	1.04	0.60

6.4 Proposed regression model

In this section, we fit simulated data to the LS regression (16) with $k = 1$. First, we obtain the critical values of the coefficient-based test and the t-test when the true DGP is $y_t = y_{t-1} + u_t$ with $y_0 = 10$ and $u_t \overset{iid}{\sim} N(0, 1)$. Both the c-stat and the t-stat are calculated for 10,000 replications and we report the 90%, 95% and 99% critical values for $n = 20, 50, 100, 250, 500$ in Table 11. Also reported are the 90%, 95% and 99% critical values of the asymptotic distributions obtained from Table B.5 and Table B.6 in Hamilton (1994). These critical values can be used to test the null hypothesis of unit root against the explosive alternative in the proposed regression (16) with $k = 1$. The asymptotic critical values can be used to test the null hypothesis of unit root against the explosive alternative in the proposed regression (16) with any k. When the test statistics take values larger than the corresponding critical values, the evidence of explosiveness is found.

Second, in Table 12 we report the proportions of replications, out of 10,000 replications, where the coefficient-based test and the t-test reject the unit root null hypothesis in favor

Table 13: Proportion of replications for unit root tests to reject the null hypothesis of unit root in favor of explosiveness based on the proposed regression

	linear trend		quadratic trend		explosive ($\beta = 1.03$)		explosive ($\beta = 1.05$)	
n	c-stat	t-stat	c-stat	t-stat	c-stat	t-stat	c-stat	t-stat
50	0	0.007	0	0.043	0.268	0.700	0.984	0.995
100	0	0.001	0	0.006	0.987	0.987	1.000	1.000
250	0	0	0	0.001	1.000	1.000	1.000	1.000

of explosiveness when the true DGP is $y_t = -4t + u_t$, or $y_t = -4t^2 + u_t$, or $y_t = \beta y_{t-1} + u_t$ with $y_0 = 10$ and $\beta = 1.03$ or $\beta = 1.05$, respectively. The two values of β for the explosive process are empirically reasonable. In all DGPs, $u_t \overset{iid}{\sim} N(0, 1)$ and three different sample sizes are used, i.e., $n = 50, 100, 250$. When the c-stat (t-stat) is larger than the corresponding 99% finite sample critical value, the unit root test finds evidence of explosiveness in the simulated data. The overall conclusion from this simulation study is that our proposed procedure can effectively distinguish the trend-stationary processes from the explosive process. In particular, when data is generated from the trend-stationary models, in no replication the coefficient-based test based on the proposed regression finds evidence of explosiveness in all six cases considered. In a small number of replications, the t-test based on the proposed regression finds the evidence of explosiveness in unit root testing. The proportion becomes smaller when the sample size increases. When data is generated from an explosive process with a stronger explosive behavior ($\beta = 1.05$), the two tests almost always find evidence of explosiveness. When the explosive behavior is not so strong ($\beta = 1.03$) and the sample size is small ($n = 50$), the two statistics, especially the coefficient-based statistic, has difficulty in rejecting the unit root hypothesis. When $n = 100$ or 200, the two unit root tests almost always find the evidence of explosiveness regardless of $\beta = 1.03$ or 1.05.

From the above simulations, it is clear that there is a trade-off between the coefficient-based test and the t-test. While the c-statistic is more robust against trend-stationarity in data, it is less powerful in identifying a mildly explosive behavior in small samples. For conservative users whose primary concern is on the robustness property of the right-tailed unit root testing against the trend stationary behavior, our recommendation is to use the coefficient-based test.

7 Empirical Studies

The empirical usefulness of the right-tailed unit root tests has been made clear in PSY (2015a) for testing for the presence of bubbles and for dating each bubble. The PSY procedure relies on repeated calculations of the t-statistic in autoregression in a recursive manner where the end point r_2 (fraction) of each sample takes a value between r_0 to 1 and the starting point r_1 (fraction) of the sample takes a value between 0 to $r_2 - r_0$ with r_0 (fraction) being the smallest sample window. So $\lfloor nr_0 \rfloor$, the integer part of nr_0, is the minimum window size in the calculations. PSY (2015a) proposed the GSADF statistic to be the largest t-statistic in the double recursion over all possible combinations of r_1 and r_2, namely

$$GSADF(r_0) = \sup_{r_2 \in [r_0,1], r_1 \in [0, r_2 - r_0]} \left\{ t_{r_1}^{r_2} \right\}, \tag{19}$$

where $t_{r_1}^{r_2}$ is the t-statistic based on the sample from r_1 to r_2. PSY (2015a) derived the asymptotic distribution of $GSADF(r_0)$ when the null hypothesis is a unit root process, from which the right-tailed critical values can be obtained. The intuition why the test is reasonable is that if there is a subsample of data corresponding to an explosive bubble period, the t-statistic calculated from this subsample should take a large value. The proposed test is done in a recursive way to find such a subsample by seeking the largest value of the t-statistic in different subsamples. Clearly, the GSADF test can also be done based on the coefficient-based statistics, and the corresponding limiting distribution under the unit root null hypothesis can be easily obtained based on the results given in Part (a) of Theorem 5.1.

After the presence of bubbles has been detected, one can estimate the origination date and the termination date of each bubble by

$$\hat{r}_e = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSADF_{r_2}(r_0) > cv_{r_0} \right\}, \tag{20}$$

$$\hat{r}_f = \inf_{r_2 \in [\hat{r}_e,1]} \left\{ r_2 : BSADF_{r_2}(r_0) < cv_{r_0} \right\}, \tag{21}$$

where

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \left\{ t_{r_1}^{r_2} \right\}, \tag{22}$$

and cv_{r_0} is the critical value of the sup t-statistic. Phillips, Wu and Yu (2011) developed the asymptotic distribution of the sup t-statistic as $\sup_{r \in [r_0,1]} \int_0^r \widetilde{W}_r dW / \left(\int_0^r \widetilde{W}_r^2 \right)^{1/2}$, where $\widetilde{W}_r(s) = W(s) - \frac{1}{r}\int_0^r W$ is the demeaned Brownian motion. The 90%, 95%, 99% asymptotic and finite sample critical values of the sup t-statistic were reported in Table 1 of PSY for various values of r_0. Interestingly, in all cases the 95% critical value is much larger than zero. The intuition for the two estimators is that, $\hat{r}_e$ is the first time when the evidence of explosive behavior is found while $\hat{r}_f$ is, given an explosive subsample of data having been found, the first time when the evidence of explosive behavior disappears.

The t-statistics in (22) can be replaced with the coefficient-based statistics. Phillips and Yu (2011) obtained the asymptotic distribution of the corresponding sup c-statistic as $\sup_{r \in [r_0,1]} r \int_0^r \widetilde{W}_r dW / \int_0^r \widetilde{W}_r^2$. The 90%, 95%, 99% asymptotic and finite sample critical values of the sup c-statistic can similarly be obtained. As for the sup t-statistic, the 95% critical values of the sup c-statistic (both asymptotic and finite sample) are greater than zero. For this reason, we recommend the use of 95% critical values in empirical studies. This is because, according to Part (c) of Theorem 5.1, if the data is generated from a trend-stationary model, the coefficient-based statistic, when obtained from the regression (16) with $k \geq 1$, converges in probability to zero and hence is less than the critical value.

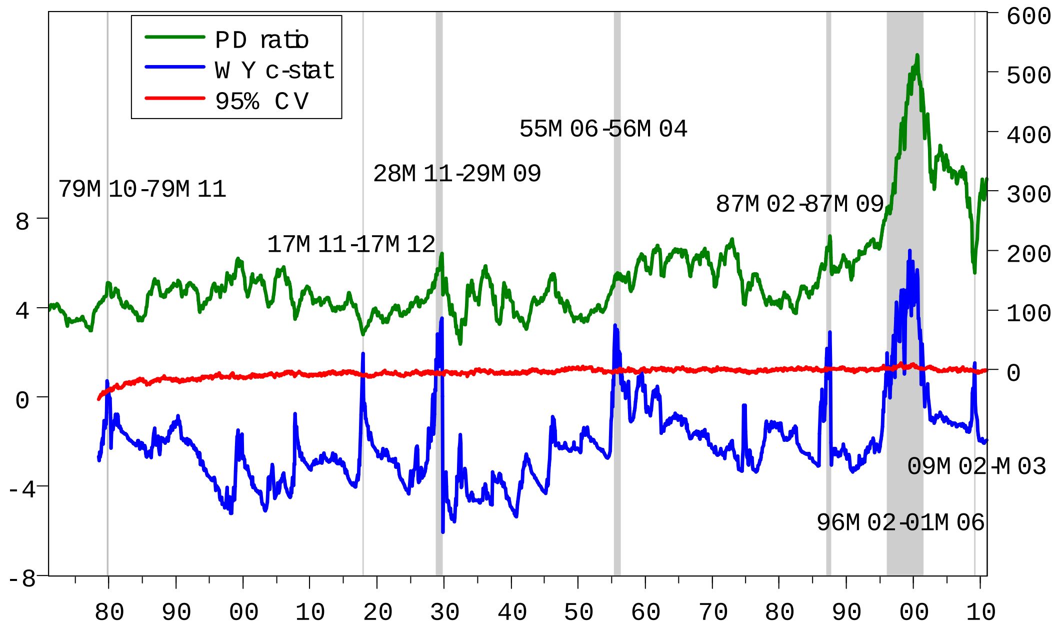

Figure 3: Date-stamping bubble periods in the S&P 500 price-dividend ratio based on the c-statistic from the proposed regression (16) with $k = 1$.

In addition, we recommend the use of the sup c-statistic because the coefficient-based statistic has better robustness properties under deterministic trends than the t-statistic in finite samples as argued earlier.

7.1 Stock market

In the first empirical study, we analyze the monthly S&P 500 stock price index-dividend ratio over the period from January 1871 to December 2010. Figure 3 plots the data, the sequence of the sup c-statistics from the regression (16) with $k = 1$, and the sequence of the 95% finite sample critical values of the sup c-statistic obtained from Monte Carlo simulations. Comparing Figure 3 with Figure 1 which is based on the sup t-statistic obtained by PSY from the regression (16) with $k = 0$, we can see that the empirical findings from the two procedure are qualitatively identical. That is, the same seven bubble episodes have been identified, namely the post long-depression period in 1880s, the great crash episode in 1920s, the postwar boom in 1950s, black Monday in 1987, the dot-com bubble in 1990s, the 1917 stock market crash in 1910s, and the subprime mortgage crisis in 2008. The last two periods, namely the 1917 stock market crash in 1910s, and the subprime mortgage crisis in 2008, although experiencing market downturns, continue to be identified as bubble episodes.

Given that the proposed testing procedure in our paper is able to distinguish the ex-

Table 14: Empirical results based on the PSY procedure and the proposed procedure for the S&P 500 stock price index-dividend ratio between October 2006 and March 2009.

	PSY		Proposed Procedure	
	c-test	t-test	c-test	t-test
	1.3439	1.0714	1.0282	0.73773
95% Finite Sample CV	-0.0498	-0.0264	0.0809	0.0377

plosive process from trend-stationary processes, the results from Figure 3 suggest that all 7 bubbles found by PSY (both the positive and negative bubbles) are robust to deterministic trends, and the deterministic trend models are probably irrelevant for describing the movement of the price-dividend time series.

Although the PSY procedure and the proposed procedure flag the same bubble episodes, the identified bubble durations are not identical. In particular, the estimated bubble termination date by the proposed method is earlier than that by the PSY method in 5 out of 7 cases. For example, according to our method, the bubble episode in 1920s ends in September 1929, one month earlier than that based on the PSY method. When checking the S&P 500 stock price index-dividend ratio, we find that after a substantial period of upward movements, the time series reaches 194.99 in September 1929 and drops to 172.53 (or 13%) in October 1929, suggesting that September 1929 is a turning point. Clearly, the termination date estimated by our method matches the turning point of the time series better.

To examine the empirical results for the data in downturn periods more closely, we apply the proposed regression (16) with $k = 1$ to the data over the period from October 2006 to March 2009 whose time series plot is shown in Figure 2. Both the coefficient-based statistic and the t-statistic are reported in Table 14, together with 95% finite sample critical values obtained under the null hypothesis of unit root. Clearly, both the PSY and our proposed procedure have to reject the null hypothesis of unit root in favor of explosiveness. Again, since our empirical method is robust to deterministic trends, the negative bubble identified by PSY is not likely due to a deterministic trend. Our result reinforces the empirical conclusion drawn in PSY.

7.2 Real estate market

In the second empirical study, we analyze the data of the US real estate market, which contains the monthly S&P/Case-Shiller US National Home Price Index and the monthly rent of primary residence, both over the period from January 1981 to June 2017 (438 monthly observations in the full sample).[6] The price-rent ratio is calculated for the sample period.

For the purpose of comparison, we first apply the PSY method (based on the sup c-statistic) with $k = 0$ to the price-rent ratio. Figure 4 plots the price-rent ratio, the

[6] The data is downloaded from Federal Reserve Bank of St. Louis. The series code for the home price index is CSUSHPISA while the code for the rent is CUSR0000SEHA.

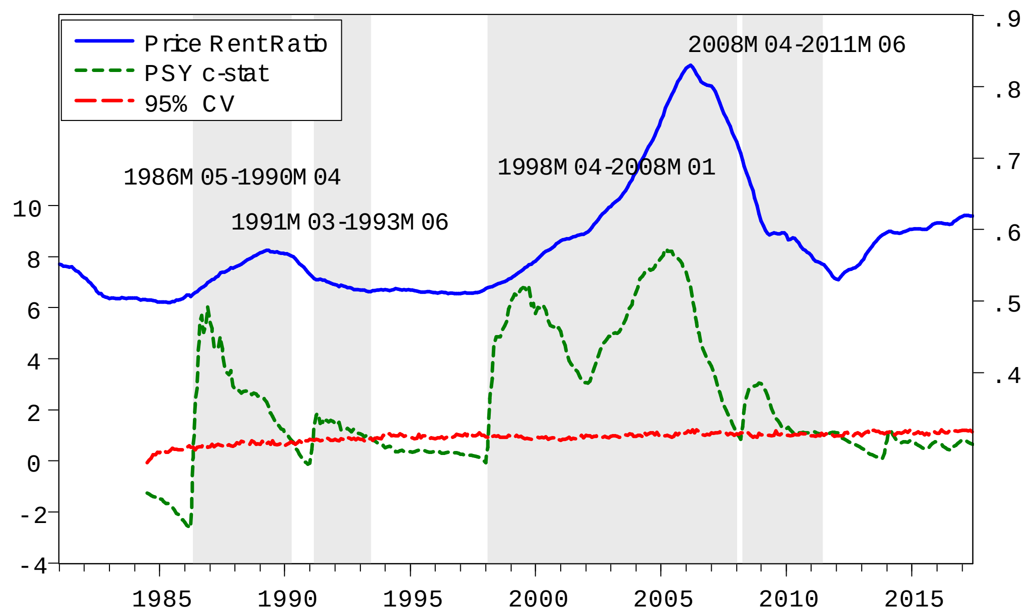

Figure 4: Date-stamping bubble periods in the US home price rent ratio based on the PSY method.

sequence of PSY c-statistics which are calculated in a recursive backward manner with the minimum window size chosen as $0.01 + 1.8/\sqrt{438} \approx 42$, and the sequence of the 95% finite sample critical values of the sup c-statistic obtained from Monte Carlo simulations. It can be seen that four bubble episodes are identified by the PSY method, namely, May 1986 to April 1990, March 1991 to June 1993, April 1998 to January 2008, and April 2008 to June 2011. The first and third episodes correspond to the well-known periods of real estate market expansions in the US. However, some post-peak periods are included as a part of bubble expansion. The price-rent ratio reached a peak in May 1989, whereas, the estimated termination date of the first bubble by PSY method is April 1990. The price-rent ratio reached another peak in March 2006, whereas, the PSY method estimates January 2008 to be the termination date of the third bubble. During the second and last detected bubble periods the real estate market experienced market downturns. Hence, according to PSY, these two periods must have negative bubbles.

We then apply the proposed regression (16) with $k = 1$ to the price-rent ratio of the US real estate market. Figure 4 plots the price-rent ratio, the sequence of the c-statistics from the proposed procedure calculated in a recursive backward manner with the same minimum window size as in PSY method, and the sequence of the 95% finite sample critical values obtained from Monte Carlo simulations. Only two bubble periods have been identified by the proposed method, namely, May 1986 to May 1989 and July 1998 to March 2006. Although these two bubble periods correspond to the first and third bubble periods

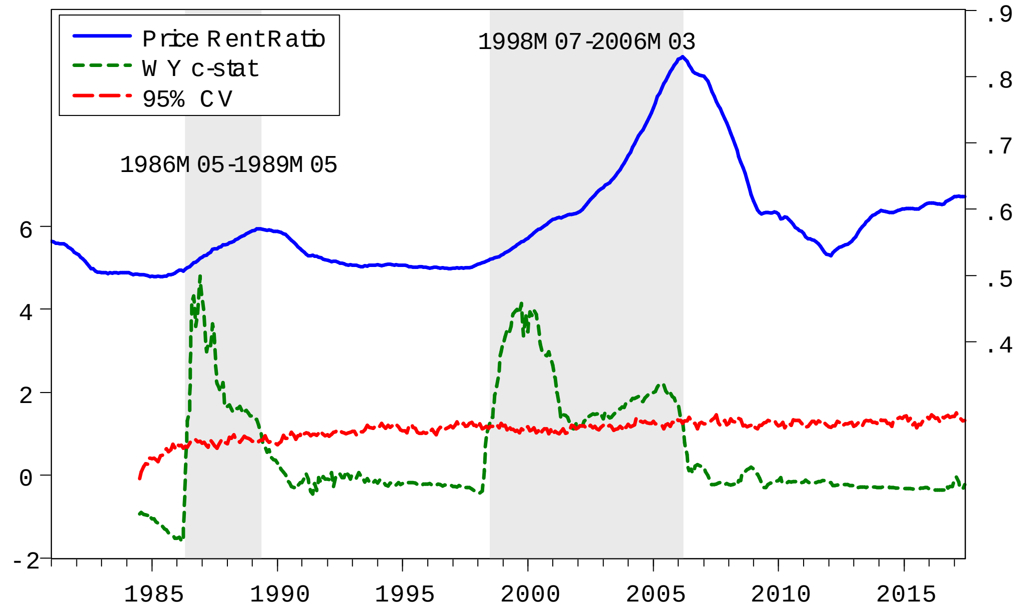

Figure 5: Date-stamping bubble periods in the US home price rent ratio based on the proposed regression method.

identified by the PSY method, they are shorter in the sense that both bubbles ended much earlier. The first bubble ended in May 1989 according to the proposed method, eleven months earlier than that identified by the PSY method. The second bubble ended in March 2006 according to the proposed method, twenty-two months earlier than that identified by the PSY method. As noted earlier, both May 1989 and March 2006 are two peaks of the time series. Clearly, the estimated termination dates synchronize the turning points by the proposed procedure. Moreover, the third and the last bubble periods identified by PSY method are not flagged as a negative bubble episode by the proposed method. This observation indicates that the market downturns in 1991-1993 and in 2008-2011 may be better explained by deterministic trend models than pure AR models.

8 Conclusion

This paper is concerned about the performance of the right-tailed unit root tests against explosive alternative when the true DGP has a deterministic trend. It is shown that when there is a linear trend in DGP (upward or downward), the unit root tests based on the AR(1) regression without intercept tend to reject the null hypothesis of unit root in favor of the explosive alternative. Similarly, when there is a quadratic trend or a cubic trend in DGP (upward or downward), the unit root tests based on the AR(1) regressions with or without an intercept also tend to reject the null hypothesis of unit root in favor of the

Table 15: The asymptotic properties of the two unit root test statistics obtained from different regression models when data come from different DGPs.

	Model (6)		Model (7)		Model (8) $k=1$		Model (8) $k=2$	
DGP	c-stat	t-stat	c-stat	t-stat	c-stat	t-stat	c-stat	t-stat
Linear Trend	1.5	∞	0	0	0	0	0	0
Quadratic Trend	2.5	∞	15/8	∞	0	0	0	0
Cubic Trend	3.5	∞	28/9	∞	0	$-\infty$	0	0
Explosive	∞	∞	∞	∞	∞	∞	∞	∞

explosive alternative. Extensive simulation studies reinforce the analytical findings.

In order for the right-tailed unit root tests to be able to distinguish the trend stationary process from the explosive AR process, we propose a new autoregressive procedure. Asymptotic distributions for the coefficient-based statistic and the t-statistic under the null hypothesis of unit root are derived. When data is generated from a trend-stationary model, we show that the null hypothesis of unit root will not be rejected. Whereas, when data is generated from an explosive autoregressive process, we show that both test statistics go to positive infinity, suggesting that the null hypothesis of unit root has to be rejected in favor of explosiveness. Interestingly, our proposed procedure is nearly identical to the augmented DF procedure discussed in PSY (2015a, Equation (4)) with an important distinction. That is our method requires k to be at least as large as 1, whereas in Equation (4) of PSY k is allowed to be zero. However, in the empirical implementation, PSY and many other empirical studies that applied the PSY method have always fixed k to be zero for simplicity.

We have applied our proposed method to real data. For the S&P 500 stock price index-dividend ratio over January 1871 to December 2010, our testing results are qualitatively identical to those found in PSY, hence, reinforcing the empirical conclusion of negative bubbles made in PSY, and suggesting that deterministic trend models are not relevant for describing the movement of the price-dividend time series. However, for the price-rent ratio of the US real estate market from January 1981 to June 2017, our testing results are different from those obtained by using PSY method. Two negative bubbles identified by PSY method are not flagged as bubbles any more. In addition, the termination dates of two identified positive bubble periods are estimated to be much earlier than those estimated by PSY method.

APPENDIX

Before we prove the asymptotic results reported in the paper, we first summarize them. Table 15 summarizes the asymptotic properties of the two unit root test statistics obtained from different regression models, (6)-(8) with $k = 1, 2$, and when data come from different DGPs (the linear trend model (5), the quadratic trend model (10), the cubic trend model (13), and the mildly explosive model).

A Proof of theorems in Section 2

Lemma A.1 *Let* $u_t = C(L)\varepsilon_t = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$, *where* $\sum_{j=0}^{\infty} j|c_j| < \infty$ *and* $\{\varepsilon_t\} \equiv$ *i.i.d.*$\left(0, \sigma^2\right)$ *with finite fourth moment. Define* $\gamma_j = E(u_t u_{t-j})$ *for* $j = 0, 1, 2, \ldots$, $\lambda = \sigma \sum_{j=0}^{\infty} c_j = \sigma C(1)$, *and* $W(\cdot)$ *being a standard Brownian motion. Then,*

(a) $n^{-1/2} \sum_{t=1}^{n} u_t \Rightarrow \lambda W(1) \overset{d}{=} N\left(0, \lambda^2\right)$;

(b) $n^{-1} \sum_{t=1}^{T} u_{t-j} u_t \xrightarrow{p} \gamma_j$ *for* $j = 0, 1, 2, \ldots$;

(c) $n^{-3/2} \sum_{t=1}^{n} t u_{t-j} \Rightarrow \lambda \left(W(1) - \int_0^1 W(r)\,dr\right)$ *for* $j = 0, 1, 2, \ldots$;

(d) $n^{-5/2} \sum_{t=1}^{n} t^2 u_{t-j} \Rightarrow \lambda \left(W(1) - 2\int_0^1 rW(r)\,dr\right)$ *for* $j = 0, 1, 2, \ldots$;

(e) $n^{-(s+1)} \sum_{t=1}^{n} t^s = 1/(s+1) + O(1/n)$ *for* $s = 1, 2, \ldots$

Proof of Lemma A.1: Proofs of these standard results are omitted.

Proof of Theorem 2.1: (a) From Model (5), we have $\Delta y_t = \delta + \Delta u_t$, which leads to

$$\widehat{\beta} = \frac{\sum_{t=2}^{n} y_{t-1} y_t}{\sum_{t=2}^{n} y_{t-1}^2} = 1 + \frac{\sum_{t=2}^{n} y_{t-1} \Delta y_t}{\sum_{t=2}^{n} y_{t-1}^2} = 1 + \frac{\sum_{t=2}^{n} y_{t-1}(\delta + \Delta u_t)}{\sum_{t=2}^{n} y_{t-1}^2}.$$

Based on the results in Lemma A.1, it can be obtained that

$$n^{-2} \sum_{t=2}^{n} \delta y_{t-1} = \frac{\delta}{n^2} \sum_{t=2}^{n} \left[\delta(t-1) + u_{t-1}\right] = \delta^2/2 + O_p\left(n^{-1}\right),$$

$$\begin{aligned} n^{-1} \sum_{t=2}^{n} (\Delta u_t) y_{t-1} &= \frac{\delta}{n} \sum_{t=2}^{n} (\Delta u_t)(t-1) + \frac{1}{n} \sum_{t=2}^{n} (\Delta u_t) u_{t-1} \\ &= \frac{\delta}{n} \left(n u_n - \sum_{t=2}^{n} u_t\right) + \frac{1}{n} \left(\sum_{t=2}^{n} u_t u_{t-1} - \sum_{t=2}^{n} u_{t-1}^2\right) \\ &= \delta u_n + \gamma_1 - \gamma_0 + o_p(1) \end{aligned}$$

and

$$\begin{aligned} n^{-3} \sum_{t=2}^{n} y_{t-1}^2 &= n^{-3} \left(\delta^2 \sum_{t=2}^{n} (t-1)^2 + \sum_{t=2}^{n} u_{t-1}^2 + 2\delta \sum_{t=2}^{n} (t-1) u_{t-1}\right) \\ &= \frac{\delta^2}{n^3} \frac{2n^3 - 3n^2 + n}{6} + O_p\left(n^{-3/2}\right) = \delta^2/3 + O_p\left(n^{-1}\right). \end{aligned}$$

Consequently, we have

$$n\left(\hat{\beta} - 1\right) = \frac{n^{-2} \sum_{t=2}^{n} (\delta + \Delta u_t) y_{t-1}}{n^{-3} \sum_{t=2}^{n} y_{t-1}^2} = \frac{\delta^2/2 + O_p\left(n^{-1}\right)}{\delta^2/3 + O_p\left(n^{-1}\right)} = \frac{3}{2} + O_p\left(\frac{1}{n}\right).$$

(b) Note that $\left[se\left(\widehat{\beta}\right)\right]^2 = \left(\sum_{t=2}^{n} y_{t-1}^2\right)^{-1} \left(\frac{1}{n-2} \sum_{t=2}^{n} \left(y_t - \widehat{\beta} y_{t-1}\right)^2\right)$. We have

$$\sum_{t=2}^{n} \left(y_t - \widehat{\beta} y_{t-1}\right)^2 = \sum_{t=2}^{n} \left[\Delta y_t - \left(\widehat{\beta} - 1\right) y_{t-1}\right]^2 = \sum_{t=2}^{n} (\Delta y_t)^2 - \left(\widehat{\beta} - 1\right)^2 \sum_{t=2}^{n} y_{t-1}^2,$$

which leads to

$$\left[se\left(\widehat{\beta}\right)\right]^2 = \frac{1}{n-2}\left[\sum_{t=2}^{n}\left(\Delta y_t\right)^2 / \sum_{t=2}^{n} y_{t-1}^2 - \left(\widehat{\beta}-1\right)^2\right].$$

Note that, as $n \to \infty$,

$$\begin{aligned}\frac{1}{n-2}\sum_{t=2}^{n}\left(\Delta y_t\right)^2 &= \frac{1}{n-2}\sum_{t=2}^{n}\left(\delta+\Delta u_t\right)^2 = \frac{(n-1)\,\delta^2 + \sum_{t=2}^{n}\left(\Delta u_t\right)^2 + 2\delta\sum_{t=2}^{n}\Delta u_t}{n-2} \\ &= \delta^2 + 2\left(\gamma_0-\gamma_1\right) + o_p\left(1\right).\end{aligned}$$

Together with the limits of $n^{-3}\sum_{t=2}^{n} y_{t-1}^2$ and $n\left(\widehat{\beta}-1\right)$ derived above, we have

$$\begin{aligned}n^3\left[se\left(\widehat{\beta}\right)\right]^2 &= \frac{\frac{1}{n-1}\sum_{t=2}^{n}\left(\Delta y_t\right)^2}{n^{-3}\sum_{t=2}^{n} y_{t-1}^2} - \frac{n^3}{n-1}\left(\widehat{\beta}-1\right)^2 \\ &\xrightarrow{p} \frac{\delta^2+2\left(\gamma_0-\gamma_1\right)}{\delta^2/3} - \left(\frac{3}{2}\right)^2 = \frac{3\delta^2+24\left(\gamma_0-\gamma_1\right)}{4\delta^2},\end{aligned}$$

which leads to the final result as

$$\frac{t_{\widehat{\beta}}}{\sqrt{n}} = \frac{n\left(\widehat{\beta}-1\right)}{n^{3/2}se\left(\widehat{\beta}\right)} \xrightarrow{p} \frac{(3/2)\sqrt{4\delta^2}}{\sqrt{3\delta^2+24\left(\gamma_0-\gamma_1\right)}} = \frac{\sqrt{3}\,|\delta|}{\sqrt{\delta^2+8\left(\gamma_0-\gamma_1\right)}}.$$

Proof of Theorem 2.2: (a) As $y_t = \delta t + u_t = \delta + y_{t-1} + \Delta u_t$, the centered LS estimator takes the form of

$$\begin{pmatrix}\tilde{\alpha}-\delta \\ \tilde{\beta}-1\end{pmatrix} = \begin{pmatrix}\sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum_{t=2}^{n}\Delta u_t \\ \sum_{t=2}^{n} y_{t-1}\Delta u_t\end{pmatrix}.$$

Based on the limits of $n^{-2}\sum_{t=2}^{n} y_{t-1}$, $n^{-3}\sum_{t=2}^{n} y_{t-1}^2$ and $n^{-1}\sum_{t=2}^{n} y_{t-1}\Delta u_t$ derived in the proof of Theorem 2.1, we have

$$\begin{pmatrix}1 & 0 \\ 0 & n^{-1}\end{pmatrix}\begin{pmatrix}\sum_{t=2}^{n}\Delta u_t \\ \sum_{t=2}^{n} y_{t-1}\Delta u_t\end{pmatrix} = \begin{bmatrix}u_n - u_1 \\ \delta u_n + \gamma_1 - \gamma_0 + o_p\left(1\right)\end{bmatrix},$$

and

$$\begin{aligned}&\begin{pmatrix}n & 0 \\ 0 & n^2\end{pmatrix}\begin{pmatrix}\sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2\end{pmatrix}^{-1}\begin{pmatrix}1 & 0 \\ 0 & n\end{pmatrix} \\ =\; &\begin{pmatrix}1 & \delta/2 \\ \delta/2 & \delta^2/3\end{pmatrix}^{-1} + O_p\left(n^{-1}\right) = \begin{pmatrix}4 & -6/\delta \\ -6/\delta & 12/\delta^2\end{pmatrix} + O_p\left(n^{-1}\right).\end{aligned}$$

Hence,

$$\begin{pmatrix}n & 0 \\ 0 & n^2\end{pmatrix}\begin{pmatrix}\tilde{\alpha}-\delta \\ \tilde{\beta}-1\end{pmatrix} = \begin{pmatrix}4 & -6/\delta \\ -6/\delta & 12/\delta^2\end{pmatrix}\begin{pmatrix}u_n - u_1 \\ \delta u_n + \left(\gamma_1-\gamma_0\right)\end{pmatrix} + o_p\left(1\right)$$

which leads to the result in (a).

(b) Note that $se\left(\tilde{\beta}\right)$ takes the form of

$$se\left(\tilde{\beta}\right) = \left[\begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \left(\frac{\sum_{t=2}^{n} \left(y_t - \tilde{\alpha} - \tilde{\beta} y_{t-1}\right)^2}{n-3} \right)\right]^{1/2}.$$

As $y_t = \delta t + u_t = \delta + y_{t-1} + \Delta u_t$ and $\left(\tilde{\alpha}, \tilde{\beta}\right) \xrightarrow{p} (\delta, 1)$, it can be proved that

$$n^{-1} \sum_{t=2}^{n} \left(y_t - \tilde{\alpha} - \tilde{\beta} y_{t-1}\right)^2 = n^{-1} \sum_{t=2}^{n} (\Delta u_t)^2 + o_p(1) = 2(\gamma_0 - \gamma_1) + o_p(1).$$

Note that

$$\begin{aligned} & \begin{pmatrix} 0 & n^2 \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ n \end{pmatrix} \\ = & \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} n & 0 \\ 0 & n^2 \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 \\ 0 & n \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ = & \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 4 & -6/\delta \\ -6/\delta & 12/\delta^2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} + O_p\left(n^{-1}\right). \end{aligned}$$

We then have

$$\begin{aligned} n^3 \left[se\left(\tilde{\beta}\right)\right]^2 &= \begin{pmatrix} 0 & n^2 \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ n \end{pmatrix} \left(\frac{\sum_{t=2}^{n} \left(y_t - \tilde{\alpha} - \tilde{\beta} y_{t-1}\right)^2}{n-1} \right) \\ &= 24(\gamma_0 - \gamma_1)/\delta^2 + o_p(1). \end{aligned}$$

Together with the limit of $n^2\left(\tilde{\beta} - 1\right)$, it is obtained that

$$\sqrt{n} t_{\tilde{\beta}} = n^2 \left(\tilde{\beta} - 1\right) / \left[n^{3/2} se\left(\tilde{\beta}\right)\right] = \frac{-12(\gamma_0 - \gamma_1)/\delta^2 + 6(u_n + u_1)/\delta}{\sqrt{24(\gamma_0 - \gamma_1)/\delta^2}} + o_p(1)$$

which leads to the result in (b) directly.

Proof of Theorem 2.3: For simplicity, we give only the proof for the regression with $k = 1$. The same approach can be applied straightforwardly to prove the results in general cases where $k > 1$. When $k = 1$, the regression (8) becomes

$$y_t = \breve{\alpha} + \breve{\beta} y_{t-1} + \breve{\psi}_1 \Delta y_{t-1} + \breve{e}_t.$$

The fact of $\Delta y_{t-1} = \delta + \Delta u_{t-1}$ makes $\breve{\alpha}$ and $\breve{\psi}_1$ not consistent, which causes difficulty to derive the limits of $\breve{\beta}$ and $t_{\breve{\beta}}$. Hence, we turn to study an alternative LS regression:

$$y_t = \breve{\alpha}^* + \breve{\beta}^* y_{t-1} + \breve{\psi}_1^* \Delta u_{t-1} + \breve{e}_t^*, \tag{23}$$

where $\left(\check{\alpha}^*, \check{\beta}^*, \check{\psi}_1^*\right)$ are the LS coefficients. It is easy to see that

$$\begin{pmatrix} \check{\alpha} \\ \check{\beta} \\ \check{\psi}_1 \end{pmatrix} = D' \begin{pmatrix} \check{\alpha}^* \\ \check{\beta}^* \\ \check{\psi}_1^* \end{pmatrix} \quad \text{and} \quad se\left(\check{\beta}\right) = se\left(\check{\beta}^*\right) \quad \text{with} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\delta & 0 & 1 \end{bmatrix}$$

where $se\left(\check{\beta}\right)$ and $se\left(\check{\beta}^*\right)$ are the standard errors of $\check{\beta}$ and $\check{\beta}^*$, respectively. Therefore,

$$\check{\beta} = \check{\beta}^* \quad \text{and} \quad t_{\check{\beta}} = \frac{\check{\beta} - 1}{se\left(\check{\beta}\right)} = \frac{\check{\beta}^* - 1}{se\left(\check{\beta}^*\right)} = t_{\check{\beta}^*},$$

where $t_{\check{\beta}^*}$ is the t statistic associated with $\check{\beta}^*$. We now focus on the regression (23) to study the limits of $\check{\beta}^*$ and $t_{\check{\beta}^*}$ as $n \to \infty$.

(a) As $y_t = \delta + y_{t-1} + \Delta u_t$, the centered LS estimator of the regression (23) is

$$\begin{bmatrix} \check{\alpha}^* - \delta \\ \check{\beta}^* - 1 \\ \check{\psi}_1^* - 0 \end{bmatrix} = \begin{bmatrix} \sum_{t=3}^n 1 & \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n \Delta u_{t-1} \\ \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n y_{t-1}^2 & \sum_{t=3}^n y_{t-1}\Delta u_{t-1} \\ \sum_{t=3}^n \Delta u_{t-1} & \sum_{t=3}^n y_{t-1}\Delta u_{t-1} & \sum_{t=3}^n (\Delta u_{t-1})^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_{t=3}^n \Delta u_t \\ \sum_{t=3}^n y_{t-1}\Delta u_t \\ \sum_{t=3}^n \Delta u_{t-1}\Delta u_t \end{bmatrix}.$$

With the limit of $n^{-1}\sum_{t=3}^n y_{t-1}\Delta u_t$ obtained in the proof of Theorem (2.1), we have

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1} \end{pmatrix} \begin{bmatrix} \sum_{t=3}^n \Delta u_t \\ \sum_{t=3}^n y_{t-1}\Delta u_t \\ \sum_{t=3}^n \Delta u_{t-1}\Delta u_t \end{bmatrix} = \begin{bmatrix} u_n - u_2 \\ \delta u_n + \gamma_1 - \gamma_0 \\ 2\gamma_1 - \gamma_0 - \gamma_2 \end{bmatrix} + o_p(1).$$

Note that

$$\begin{aligned} \frac{1}{n}\sum_{t=3}^n y_{t-1}\Delta u_{t-1} &= \frac{1}{n}\sum_{t=3}^n \left[\delta(t-1) + u_{t-1}\right]\Delta u_{t-1} \\ &= \frac{\delta}{n}\left[(n-1)u_{T-1} - \sum_{t=2}^{n-2} u_t - 2u_1\right] + \frac{1}{n}\sum_{t=3}^n u_{t-1}\Delta u_{t-1} \\ &= \delta u_{T-1} + \gamma_0 - \gamma_1 + o_p(1). \end{aligned}$$

We now have

$$\begin{pmatrix} n & 0 & 0 \\ 0 & n^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} \sum_{t=3}^n 1 & \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n \Delta u_{t-1} \\ \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n y_{t-1}^2 & \sum_{t=3}^n y_{t-1}\Delta u_{t-1} \\ \sum_{t=3}^n \Delta u_{t-1} & \sum_{t=3}^n y_{t-1}\Delta u_{t-1} & \sum_{t=3}^n (\Delta u_{t-1})^2 \end{bmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & n \end{pmatrix}$$

$$= \begin{bmatrix} 1 & \delta/2 & u_{n-1} - u_2 \\ \delta/2 & \delta^2/3 & \delta u_{T-1} + \gamma_0 - \gamma_1 \\ 0 & 0 & 2(\gamma_0 - \gamma_1) \end{bmatrix}^{-1} + o_p(1).$$

Consequently, we have

$$\begin{pmatrix} n & 0 & 0 \\ 0 & n^2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} \check{\alpha}^* - \delta \\ \check{\beta}^* - 1 \\ \check{\psi}_1^* - 0 \end{bmatrix} = \begin{bmatrix} 1 & \delta/2 & u_{n-1} - u_2 \\ \delta/2 & \delta^2/3 & \delta u_{T-1} + \gamma_0 - \gamma_1 \\ 0 & 0 & 2(\gamma_0 - \gamma_1) \end{bmatrix}^{-1} \begin{bmatrix} u_n - u_2 \\ \delta u_n + \gamma_1 - \gamma_0 \\ 2\gamma_1 - \gamma_0 - \gamma_2 \end{bmatrix} + o_p(1)$$

which shows that $n^2 \left(\check{\beta}^* - 1\right) = O_p(1)$. Hence, $n^2 \left(\check{\beta} - 1\right) = O_p(1)$.

(b) Note that

$$\begin{aligned} \frac{1}{n}\sum\nolimits_{t=3}^{n} (\check{e}_t^*)^2 &= \frac{1}{n}\sum\nolimits_{t=3}^{n} \left(y_t - \check{\alpha}^* - \check{\beta}^* y_{t-1} - \check{\psi}_1^* \Delta u_{t-1}\right)^2 \\ &= \frac{1}{n}\sum\nolimits_{t=3}^{n} \left(\Delta u_t - (\check{\alpha}^* - \delta) - \left(\check{\beta}^* - 1\right) y_{t-1} - \check{\psi}_1^* \Delta u_{t-1}\right)^2 \\ &= \frac{1}{n}\sum\nolimits_{t=3}^{n} \left(\Delta u_t - (\check{\alpha}^* - \delta) - \left(\check{\beta}^* - 1\right) y_{t-1} - \check{\psi}_1^* \Delta u_{t-1}\right) \Delta u_t \\ &= \frac{1}{n}\sum\nolimits_{t=3}^{n} (\Delta u_t)^2 - \check{\psi}_1^* \frac{1}{n}\sum\nolimits_{t=3}^{n} \Delta u_{t-1}\Delta u_t + O_p\left(n^{-2}\right) \\ &= \frac{4(\gamma_0 - \gamma_1)^2 - (2\gamma_1 - \gamma_0 - \gamma_2)^2}{2(\gamma_0 - \gamma_1)} + o_p(1) \end{aligned}$$

where the second equation is from $y_t = \delta + y_{t-1} + \Delta u_t$, the third equation comes from first-order conditions of LS regression, and the fourth equation is based on the asymptotic results obtained in the proof of (a). Hence,

$$\begin{aligned} & n^3 \left[se\left(\check{\beta}^*\right)\right]^2 \\ = & \begin{pmatrix} 0 & n^2 & 0 \end{pmatrix} \begin{bmatrix} \sum_{t=3}^{n} 1 & \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} \Delta u_{t-1} \\ \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} y_{t-1}^2 & \sum_{t=3}^{n} y_{t-1}\Delta u_{t-1} \\ \sum_{t=3}^{n} \Delta u_{t-1} & \sum_{t=3}^{n} y_{t-1}\Delta u_{t-1} & \sum_{t=3}^{n} (\Delta u_{t-1})^2 \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ n \\ 0 \end{pmatrix} \frac{\sum_{t=3}^{n} (\check{e}_t^*)^2}{n-5} \\ = & \frac{12}{\delta^2}\left[\frac{4(\gamma_0 - \gamma_1)^2 - (2\gamma_1 - \gamma_0 - \gamma_2)^2}{2(\gamma_0 - \gamma_1)}\right] + o_p(1). \end{aligned}$$

Finally, we have

$$\sqrt{n} t_{\check{\beta}} = \sqrt{n} t_{\check{\beta}^*} = \frac{n^2\left(\check{\beta}^* - 1\right)}{n^{3/2} se\left(\check{\beta}^*\right)} = O_p(1).$$

B Proof of theorems in Section 3

Proof of Theorem 3.1: (a) From Model (10), we have $\Delta y_t = -\delta + 2\delta t + \Delta u_t$. Hence,

$$\widehat{\beta} - 1 = \frac{\sum_{t=2}^{n} y_{t-1}\Delta y_t}{\sum_{t=2}^{n} y_{t-1}^2} = \frac{\sum_{t=2}^{n} y_{t-1}\left(-\delta + 2\delta t + \Delta u_t\right)}{\sum_{t=2}^{n} y_{t-1}^2}.$$

Based on the results in Lemma A.1, it can be obtained that

$$n^{-3}\sum_{t=2}^{n} y_{t-1} = \frac{1}{n^3}\sum_{t=2}^{n}\left[\delta(t-1)^2 + u_{t-1}\right] = \delta/3 + O_p\left(n^{-1}\right),$$

$$n^{-4}\sum_{t=2}^{n} ty_{t-1} = \frac{1}{n^4}\sum_{t=2}^{n} t\left[\delta\left(t-1\right)^2 + u_{t-1}\right] = \delta/4 + O_p\left(n^{-1}\right),$$

$$\begin{aligned}
n^{-2}\sum_{t=2}^{n}\left(\Delta u_t\right) y_{t-1} &= \frac{\delta}{n^2}\sum_{t=2}^{n}\left(\Delta u_t\right)\left(t-1\right)^2 + \frac{1}{n^2}\sum_{t=2}^{n}\left(\Delta u_t\right) u_{t-1} \\
&= \frac{\delta}{n^2}\left(\left(n-1\right)^2 u_n - 2\sum_{t=2}^{n}\left(t-1\right) u_{t-1} + \sum_{t=2}^{n} u_{t-1}\right) + O_p\left(n^{-1}\right) \\
&= \delta u_n + O_p\left(n^{-1/2}\right)
\end{aligned}$$

and

$$n^{-5}\sum_{t=2}^{n} y_{t-1}^2 = n^{-5}\sum_{t=2}^{n}\left(\delta^2\left(t-1\right)^4 + u_{t-1}^2 + 2\delta\left(t-1\right)^2 u_{t-1}\right) = \delta^2/5 + O_p\left(n^{-1}\right).$$

Consequently, we have

$$n\left(\hat{\beta}-1\right) = \frac{n^{-4}\sum_{t=2}^{n} y_{t-1}\left(-\delta + 2\delta t + \Delta u_t\right)}{n^{-5}\sum_{t=2}^{n} y_{t-1}^2} = \frac{2\delta\left(\delta/4\right) + O_p\left(n^{-1}\right)}{\delta^2/5 + O_p\left(n^{-1}\right)} = \frac{5}{2} + O_p\left(\frac{1}{n}\right).$$

(b) Note that

$$n^{-3}\sum_{t=2}^{n}\left(\Delta y_t\right)^2 = n^{-3}\sum_{t=2}^{n}\left(-\delta + 2\delta t + \Delta u_t\right)^2 = \frac{4\delta^2}{n^3}\sum_{t=2}^{n} t^2 + O_p\left(n^{-1}\right) = \frac{4\delta^2}{3} + O_p\left(n^{-1}\right).$$

Then, we have

$$\begin{aligned}
n^3\left[se\left(\widehat{\beta}\right)\right]^2 &= \frac{n^3}{n-2}\frac{\sum_{t=2}^{n}\left(y_t - \widehat{\beta} y_{t-1}\right)^2}{\sum_{t=2}^{n} y_{t-1}^2} = \frac{n^3}{n-2}\frac{\sum_{t=2}^{n}\left[\Delta y_t - \left(\widehat{\beta}-1\right) y_{t-1}\right]^2}{\sum_{t=2}^{n} y_{t-1}^2} \\
&= \frac{n^3}{n-2}\left[\frac{\sum_{t=2}^{n}\left(\Delta y_t\right)^2}{\sum_{t=2}^{n} y_{t-1}^2} - \left(\widehat{\beta}-1\right)^2\right] \\
&= \frac{4\delta^2/3}{\delta^2/5} - \left(\frac{5}{2}\right)^2 + O_p\left(n^{-1}\right) = \frac{5}{12} + O_p\left(n^{-1}\right).
\end{aligned}$$

As a result,

$$\frac{t_{\widehat{\beta}}}{\sqrt{n}} = \frac{n\left(\widehat{\beta}-1\right)}{n^{3/2} se\left(\widehat{\beta}\right)} \xrightarrow{p} \frac{5/2}{\sqrt{5/12}} = \sqrt{15}.$$

Proof of Theorem 3.2: (a) From Model (10), we have $y_t = -\delta + y_{t-1} + 2\delta t + \Delta u_t$. Hence, the centered LS estimator takes the form of

$$\begin{pmatrix}\tilde{\alpha}+\delta \\ \tilde{\beta}-1\end{pmatrix} = \begin{pmatrix}\sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2\end{pmatrix}^{-1}\begin{pmatrix}\sum_{t=2}^{n}\left(2\delta t + \Delta u_t\right) \\ \sum_{t=2}^{n} y_{t-1}\left(2\delta t + \Delta u_t\right)\end{pmatrix}.$$

Based on the asymptotic results obtained in the proof of Theorem 3.1, we have

$$\begin{pmatrix} n^{-2} & 0 \\ 0 & n^{-4} \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} (2\delta t + \Delta u_t) \\ \sum_{t=2}^{n} y_{t-1} (2\delta t + \Delta u_t) \end{pmatrix} = \begin{bmatrix} \delta \\ \delta^2/2 \end{bmatrix} + O_p\left(n^{-1}\right),$$

and

$$\begin{pmatrix} n^{-1} & 0 \\ 0 & n \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} n^2 & 0 \\ 0 & n^4 \end{pmatrix} = \begin{pmatrix} 1 & \delta/3 \\ \delta/3 & \delta^2/5 \end{pmatrix}^{-1} + O_p\left(n^{-1}\right).$$

Hence,

$$\begin{pmatrix} n^{-1} & 0 \\ 0 & n \end{pmatrix} \begin{pmatrix} \tilde{\alpha} + \delta \\ \tilde{\beta} - 1 \end{pmatrix} = \begin{pmatrix} 1 & \delta/3 \\ \delta/3 & \delta^2/5 \end{pmatrix}^{-1} \begin{bmatrix} \delta \\ \delta^2/2 \end{bmatrix} + O_p\left(n^{-1}\right) = \begin{bmatrix} 3\delta/8 \\ 15/8 \end{bmatrix} + O_p\left(n^{-1}\right).$$

(b) Let $\eta_t = 2\delta t + \Delta u_t = \Delta y_t + \delta$. We have

$$\frac{1}{n^3} \sum_{t=2}^{n} (\eta_t)^2 = 4\delta^2/3 + O_p\left(n^{-1}\right), \quad \frac{1}{n^2} \sum_{t=2}^{n} \eta_t = \delta + O_p\left(n^{-1}\right), \text{ and}$$

$$\frac{1}{n^4} \sum_{t=2}^{n} y_{t-1} \eta_t = 2\delta \frac{1}{n^4} \sum_{t=2}^{n} t y_{t-1} + O_p\left(n^{-2}\right) = \delta^2/2 + O_p\left(n^{-1}\right),$$

Together with the limits of $(\tilde{\alpha} + \delta)/n$ and $n\left(\tilde{\beta} - 1\right)$, we have

$$\begin{aligned} \frac{1}{n^3} \sum_{t=2}^{n} \tilde{e}_t^2 &= \frac{1}{n^3} \sum_{t=2}^{n} \left(y_t - \tilde{\alpha} - \tilde{\beta} y_{t-1}\right)^2 = \frac{1}{n^3} \sum_{t=2}^{n} \left[\eta_t - (\tilde{\alpha} + \delta) - \left(\tilde{\beta} - 1\right) y_{t-1}\right]^2 \\ &= \frac{1}{n^3} \sum_{t=2}^{n} \left[\eta_t - (\tilde{\alpha} + \delta) - \left(\tilde{\beta} - 1\right) y_{t-1}\right] \eta_t = \delta^2/48 + O_p\left(n^{-1}\right), \end{aligned}$$

where the third equation is from the first-order conditions of LS regression. Therefore,

$$\begin{aligned} n^3 \left[se\left(\tilde{\beta}\right)\right]^2 &= \begin{pmatrix} 0 & n \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ n^4 \end{pmatrix} \left(\frac{\sum_{t=2}^{n} \tilde{e}_t^2}{n^2 (n-3)} \right) \\ &= \begin{pmatrix} 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & \delta/3 \\ \delta/3 & \delta^2/5 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{\delta^2}{48} + O_p\left(n^{-1}\right) = \frac{15}{64} + O_p\left(n^{-1}\right). \end{aligned}$$

Consequently, we have

$$t_{\tilde{\beta}}/\sqrt{n} = \frac{n\left(\tilde{\beta} - 1\right)}{n^{3/2} se\left(\tilde{\beta}\right)} = \frac{15/8}{\sqrt{15/64}} + O_p\left(n^{-1}\right) \xrightarrow{p} \sqrt{15}.$$

Proof of Theorem 3.3: We first prove the results for the regression (8) with $k = 1$ in details. Then, for simplicity, we give the outline of the proof for general case with $k > 1$.

(a) For Model (10), we have $\Delta^2 y_t = 2\delta + \Delta^2 u_t$ which leads to

$$y_t = 2\delta + y_{t-1} + \Delta y_{t-1} + \Delta^2 u_t$$

Hence, the centered LS estimator of the regression (8) with $k=1$ is

$$\begin{bmatrix}\breve{\alpha}-2\delta\\ \breve{\beta}-1\\ \breve{\psi}_1-1\end{bmatrix}=\begin{bmatrix}\sum_{t=3}^n 1 & \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n \Delta y_{t-1}\\ \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n y_{t-1}^2 & \sum_{t=3}^n y_{t-1}\Delta y_{t-1}\\ \sum_{t=3}^n \Delta y_{t-1} & \sum_{t=3}^n y_{t-1}\Delta y_{t-1} & \sum_{t=3}^n (\Delta y_{t-1})^2\end{bmatrix}^{-1}\begin{bmatrix}\sum_{t=3}^n \Delta^2 u_t\\ \sum_{t=3}^n y_{t-1}\Delta^2 u_t\\ \sum_{t=3}^n \Delta y_{t-1}\Delta^2 u_t\end{bmatrix}.$$

Note that

$$\begin{aligned}
n^{-2}\sum_{t=3}^n y_{t-1}\Delta^2 u_t &= n^{-2}\sum_{t=3}^n \left[\delta(t-1)^2+u_{t-1}\right]\Delta^2 u_t\\
&= n^{-2}\sum_{t=3}^n \left[\delta(t-1)^2\right][\Delta u_t-\Delta u_{t-1}]+O_p\left(n^{-1}\right)\\
&= \delta n^{-2}\left[(n-1)^2\Delta u_n-2\sum_{t=2}^{n-1} t\Delta u_t+\sum_{t=3}^{n-1}\Delta u_t\right]+O_p\left(n^{-1}\right)\\
&= \delta\Delta u_n+O_p\left(n^{-1}\right)
\end{aligned}$$

and

$$\begin{aligned}
n^{-1}\sum_{t=3}^n \Delta y_{t-1}\Delta^2 u_t &= n^{-1}\sum_{t=3}^n (2\delta t-3\delta+\Delta u_{t-1})\Delta^2 u_t\\
&= 2\delta n^{-1}\sum_{t=3}^n t\Delta^2 u_t+n^{-1}\sum_{t=3}^n \Delta u_{t-1}\Delta^2 u_t+O_p\left(n^{-1}\right)\\
&= 2\delta n^{-1}(n\Delta u_n)+n^{-1}\sum_{t=3}^n \Delta u_{t-1}\Delta^2 u_t+O_p\left(n^{-1}\right)\\
&= 2\delta\Delta u_n+(4\gamma_1-3\gamma_0-\gamma_2)+o_p(1).
\end{aligned}$$

Hence,

$$\begin{pmatrix}1&0&0\\0&n^{-2}&0\\0&0&n^{-1}\end{pmatrix}\begin{bmatrix}\sum_{t=3}^n \Delta^2 u_t\\ \sum_{t=3}^n y_{t-1}\Delta^2 u_t\\ \sum_{t=3}^n \Delta y_{t-1}\Delta^2 u_t\end{bmatrix}=\begin{bmatrix}\Delta u_n-\Delta u_2\\ \delta\Delta u_n\\ 2\delta\Delta u_n+(4\gamma_1-3\gamma_0-\gamma_2)\end{bmatrix}+o_p(1).$$

Next, note that $n^{-2}\sum_{t=3}^n \Delta y_{t-1}=n^{-2}(y_{n-1}-y_2)=\delta+O_p\left(n^{-1}\right)$,

$$\begin{aligned}
n^{-4}\sum_{t=3}^n y_{t-1}\Delta y_{t-1} &= n^{-4}\sum_{t=3}^n \left[\delta(t-1)^2+u_{t-1}\right][2\delta t-3\delta+\Delta u_{t-1}]\\
&= 2\delta^2 n^{-4}\sum_{t=3}^n t^3+O_p\left(n^{-1}\right)=\delta^2/2+O_p\left(n^{-1}\right)
\end{aligned}$$

and

$$\begin{aligned}
n^{-3}\sum_{t=3}^n (\Delta y_{t-1})^2 &= n^{-3}\sum_{t=3}^n (2\delta t-3\delta+\Delta u_{t-1})^2\\
&= \frac{4\delta^2}{n^3}\sum_{t=2}^n t^2+O_p\left(n^{-1}\right)=\frac{4\delta^2}{3}+O_p\left(n^{-1}\right).
\end{aligned}$$

Together with the limits of $n^{-3}\sum_{t=3}^n y_{t-1}$ and $n^{-5}\sum_{t=3}^n y_{t-1}^2$ obtained in the proof of Theorem (3.1), we have

$$\begin{pmatrix}n&0&0\\0&n^3&0\\0&0&n^2\end{pmatrix}\begin{bmatrix}\sum_{t=3}^n 1 & \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n \Delta y_{t-1}\\ \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n y_{t-1}^2 & \sum_{t=3}^n y_{t-1}\Delta y_{t-1}\\ \sum_{t=3}^n \Delta y_{t-1} & \sum_{t=3}^n y_{t-1}\Delta y_{t-1} & \sum_{t=3}^n (\Delta y_{t-1})^2\end{bmatrix}^{-1}\begin{pmatrix}1&0&0\\0&n^2&0\\0&0&n\end{pmatrix}$$

$$= \begin{bmatrix} 1 & \delta/3 & \delta \\ \delta/3 & \delta^2/5 & \delta^2/2 \\ \delta & \delta^2/2 & 4\delta^2/3 \end{bmatrix}^{-1} + O_p\left(n^{-1}\right)$$

As a result, we have

$$\begin{pmatrix} n & 0 & 0 \\ 0 & n^3 & 0 \\ 0 & 0 & n^2 \end{pmatrix} \begin{bmatrix} \breve{\alpha} - 2\delta \\ \breve{\beta} - 1 \\ \breve{\psi}_1 - 1 \end{bmatrix} = \begin{bmatrix} 1 & \delta/3 & \delta \\ \delta/3 & \delta^2/5 & \delta^2/2 \\ \delta & \delta^2/2 & 4\delta^2/3 \end{bmatrix}^{-1} \begin{bmatrix} \Delta u_n - \Delta u_2 \\ \delta \Delta u_n \\ 2\delta\Delta u_n + (4\gamma_1 - 3\gamma_0 - \gamma_2) \end{bmatrix} + o_p(1)$$

which leads to $n^3\left(\breve{\beta} - 1\right) = O_p(1)$.

(b) Note that $y_t = 2\delta + y_{t-1} + \Delta y_{t-1} + \Delta^2 u_t$. Together with the fact that $\breve{\alpha}$, $\breve{\beta}$, and $\breve{\psi}_1$ are all consistent, it can be proved that

$$\begin{aligned} \frac{1}{n}\sum\nolimits_{t=3}^{n}(\breve{e}_t)^2 &= \frac{1}{n}\sum\nolimits_{t=3}^{n}\left(y_t - \breve{\alpha} - \breve{\beta}y_{t-1} - \breve{\psi}_1\Delta y_{t-1}\right)^2 = \frac{1}{n}\sum\nolimits_{t=3}^{n}\left(\Delta^2 u_t\right)^2 + o_p(1) \\ &= 6\gamma_0 - 8\gamma_1 + 2\gamma_2 + o_p(1). \end{aligned}$$

As a result, we have

$$\begin{aligned} & n^5\left[se\left(\breve{\beta}\right)\right]^2 \\ =\ & n^5\begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} \sum_{t=3}^{n} 1 & \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} \Delta y_{t-1} \\ \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} y_{t-1}^2 & \sum_{t=3}^{n} y_{t-1}\Delta y_{t-1} \\ \sum_{t=3}^{n} \Delta y_{t-1} & \sum_{t=3}^{n} y_{t-1}\Delta y_{t-1} & \sum_{t=3}^{n} (\Delta y_{t-1})^2 \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \frac{\sum_{t=3}^{n}(\breve{e}_t)^2}{n-5} \\ =\ & \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} 1 & \delta/3 & \delta \\ \delta/3 & \delta^2/5 & \delta^2/2 \\ \delta & \delta^2/2 & 4\delta^2/3 \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} [6\gamma_0 - 8\gamma_1 + 2\gamma_2] + o_p(1). \end{aligned}$$

Therefore,

$$\sqrt{n}t_{\breve{\beta}} = \frac{n^3\left(\breve{\beta} - 1\right)}{n^{5/2}se\left(\breve{\beta}\right)} = O_p(1).$$

For the regression (8) with $k = 2$:

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

it is confronted with the problem of perfect multi-collinearity as $\Delta y_{t-1} = 2\delta t - 3\delta + \Delta u_{t-1}$ and $\Delta y_{t-2} = 2\delta t - 5\delta + \Delta u_{t-2}$. Note that $\Delta y_{t-2} - \Delta y_{t-1} + 2\delta = -\Delta^2 u_{t-1}$. We now consider the regression

$$y_t = \breve{\alpha}^* + \breve{\beta}^* y_{t-1} + \breve{\psi}_1^*\Delta y_{t-1} + \breve{\psi}_2^*\left(-\Delta^2 u_{t-1}\right) + \breve{e}_t^*. \tag{24}$$

It can be proved that

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

which leads to $\breve{\beta} = \breve{\beta}^*$. It is also easy to prove that $t_{\breve{\beta}} = t_{\breve{\beta}^*}$. Note that the regression (24) does not face the problem of perfect multi-collinearity, and includes the true DGP of y_t when $\begin{pmatrix} \breve{\alpha}^* & \breve{\beta}^* & \breve{\psi}_1^* & \breve{\psi}_2^* \end{pmatrix} = \begin{pmatrix} 2\delta & 1 & 1 & 0 \end{pmatrix}$. It can be proved that $n^3 \left(\breve{\beta}^* - 1 \right) = O_p(1)$ and $\sqrt{n} t_{\breve{\beta}^*} = O_p(1)$.

The same method can be extended to prove the results in the regression (8) with $k > 2$. We will still have $n^3 \left(\breve{\beta}^* - 1 \right) = O_p(1)$ and $\sqrt{n} t_{\breve{\beta}^*} = O_p(1)$, but the form of the limiting distributions may change as k varies.

C Proof of theorems in Section 4

Proof of Theorem 4.1: (a) From Model (13), we have $\Delta y_t = \delta - 3\delta t + 3\delta t^2 + \Delta u_t$. Hence,

$$\widehat{\beta} - 1 = \frac{\sum_{t=2}^n y_{t-1} \Delta y_t}{\sum_{t=2}^n y_{t-1}^2} = \frac{\sum_{t=2}^n y_{t-1} \left(\delta - 3\delta t + 3\delta t^2 + \Delta u_t \right)}{\sum_{t=2}^n y_{t-1}^2}.$$

Based on the results in Lemma A.1, it is obtained that

$$n^{-4} \sum_{t=2}^n y_{t-1} = \frac{1}{n^4} \sum_{t=2}^n \left[\delta (t-1)^3 + u_{t-1} \right] = \delta/4 + O_p\left(n^{-1}\right),$$

$$n^{-5} \sum_{t=2}^n t y_{t-1} = \frac{1}{n^5} \sum_{t=2}^n t \left[\delta (t-1)^3 + u_{t-1} \right] = \delta/5 + O_p\left(n^{-1}\right),$$

$$n^{-6} \sum_{t=2}^n t^2 y_{t-1} = \frac{1}{n^6} \sum_{t=2}^n t^2 \left[\delta (t-1)^3 + u_{t-1} \right] = \delta/6 + O_p\left(n^{-1}\right),$$

$$n^{-3} \sum_{t=2}^n (\Delta u_t) y_{t-1} = n^{-3} \sum_{t=2}^n (\Delta u_t) \left[\delta (t-1)^3 + u_{t-1} \right] = \delta u_n + O_p\left(n^{-1/2}\right)$$

and

$$n^{-7} \sum_{t=2}^n y_{t-1}^2 = n^{-7} \sum_{t=2}^n \left[\delta (t-1)^3 + u_{t-1} \right]^2 = \delta^2/7 + O_p\left(n^{-1}\right).$$

Consequently, we have

$$n\left(\hat{\beta} - 1\right) = \frac{(3\delta) n^{-6} \sum_{t=2}^n t^2 y_{t-1}}{n^{-7} \sum_{t=2}^n y_{t-1}^2} = \frac{3\delta (\delta/6) + O_p\left(n^{-1}\right)}{\delta^2/7 + O_p\left(n^{-1}\right)} = \frac{7}{2} + O_p\left(\frac{1}{n}\right).$$

(b) Note that

$$n^{-5} \sum_{t=2}^n (\Delta y_t)^2 = \frac{(3\delta)^2}{n^5} \sum_{t=2}^n t^4 + O_p\left(n^{-1}\right) = \frac{9\delta^2}{5} + O_p\left(n^{-1}\right).$$

Then, we have

$$n^3 \left[se\left(\widehat{\beta}\right) \right]^2 = \frac{n^3}{n-2} \frac{\sum_{t=2}^n \left(y_t - \widehat{\beta} y_{t-1} \right)^2}{\sum_{t=2}^n y_{t-1}^2} = \frac{n^3}{n-2} \frac{\sum_{t=2}^n \left[\Delta y_t - \left(\widehat{\beta} - 1 \right) y_{t-1} \right]^2}{\sum_{t=2}^n y_{t-1}^2}$$

$$\begin{aligned}
&= \frac{n^3}{n-2}\left[\frac{\sum_{t=2}^{n}(\Delta y_t)^2}{\sum_{t=2}^{n} y_{t-1}^2} - \left(\widehat{\beta}-1\right)^2\right] \\
&= \frac{9\delta^2/5}{\delta^2/7} - \left(\frac{7}{2}\right)^2 + O_p\left(n^{-1}\right) = \frac{7}{20} + O_p\left(n^{-1}\right).
\end{aligned}$$

As a result,

$$\frac{t_{\widehat{\beta}}}{\sqrt{n}} = \frac{n\left(\widehat{\beta}-1\right)}{n^{3/2} se\left(\widehat{\beta}\right)} \xrightarrow{p} \frac{7/2}{\sqrt{7/20}} = \sqrt{35}.$$

Proof of Theorem 4.2: (a) From Model (13), we have $y_t = \delta + y_{t-1} + f(t) + \Delta u_t$, where $f(t) = 3\delta t^2 - 3\delta t$. Hence, the centered LS estimator takes the form of

$$\begin{pmatrix} \tilde{\alpha}-\delta \\ \tilde{\beta}-1 \end{pmatrix} = \begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=2}^{n}\left[f(t)+\Delta u_t\right] \\ \sum_{t=2}^{n} y_{t-1}\left[f(t)+\Delta u_t\right] \end{pmatrix}.$$

Based on the asymptotic results obtained in the proof of Theorem 4.1, we have

$$\begin{pmatrix} n^{-3} & 0 \\ 0 & n^{-6} \end{pmatrix}\begin{pmatrix} \sum_{t=2}^{n}\left[f(t)+\Delta u_t\right] \\ \sum_{t=2}^{n} y_{t-1}\left[f(t)+\Delta u_t\right] \end{pmatrix} = \begin{bmatrix} \delta \\ \delta^2/2 \end{bmatrix} + O_p\left(n^{-1}\right),$$

and

$$\begin{pmatrix} n^{-2} & 0 \\ 0 & n \end{pmatrix}\begin{pmatrix} \sum_{t=2}^{n} 1 & \sum_{t=2}^{n} y_{t-1} \\ \sum_{t=2}^{n} y_{t-1} & \sum_{t=2}^{n} y_{t-1}^2 \end{pmatrix}^{-1}\begin{pmatrix} n^3 & 0 \\ 0 & n^6 \end{pmatrix} = \begin{pmatrix} 1 & \delta/4 \\ \delta/4 & \delta^2/7 \end{pmatrix}^{-1} + O_p\left(n^{-1}\right).$$

Hence,

$$\begin{pmatrix} n^{-2} & 0 \\ 0 & n \end{pmatrix}\begin{pmatrix} \tilde{\alpha}-\delta \\ \tilde{\beta}-1 \end{pmatrix} = \begin{pmatrix} 1 & \delta/4 \\ \delta/4 & \delta^2/7 \end{pmatrix}^{-1}\begin{bmatrix} \delta \\ \delta^2/2 \end{bmatrix} + O_p\left(n^{-1}\right) = \begin{bmatrix} 2\delta/9 \\ 28/9 \end{bmatrix} + O_p\left(n^{-1}\right).$$

(b) Let $\xi_t = 3\delta t^2 - 3\delta t + \Delta u_t = \Delta y_t - \delta$. We have

$$\frac{1}{n^5}\sum_{t=2}^{n}(\xi_t)^2 = 9\delta^2/5 + O_p\left(n^{-1}\right), \quad \frac{1}{n^3}\sum_{t=2}^{n}\xi_t = \delta + O_p\left(n^{-1}\right), \text{ and}$$

$$\frac{1}{n^6}\sum_{t=2}^{n} y_{t-1}\xi_t = 3\delta\frac{1}{n^6}\sum_{t=2}^{n} t^2 y_{t-1} + O_p\left(n^{-1}\right) = \delta^2/2 + O_p\left(n^{-1}\right).$$

Together with the limits of $n^{-2}(\tilde{\alpha}-\delta)$ and $n\left(\tilde{\beta}-1\right)$ derived above, we have

$$\begin{aligned}
\frac{1}{n^5}\sum_{t=2}^{n}\tilde{e}_t^2 &= \frac{1}{n^5}\sum_{t=2}^{n}\left(y_t - \tilde{\alpha} - \tilde{\beta}y_{t-1}\right)^2 = \frac{1}{n^5}\sum_{t=2}^{n}\left[\xi_t - (\tilde{\alpha}-\delta) - \left(\tilde{\beta}-1\right)y_{t-1}\right]^2 \\
&= \frac{1}{n^5}\sum_{t=2}^{n}\left[\xi_t - (\tilde{\alpha}-\delta) - \left(\tilde{\beta}-1\right)y_{t-1}\right]\xi_t = \delta^2/45 + O_p\left(n^{-1}\right),
\end{aligned}$$

where the third equation is from the first-order conditions of LS regression. Therefore,

$$\begin{aligned}
n^3\left[se\left(\tilde{\beta}\right)\right]^2 &= \begin{pmatrix}0 & n\end{pmatrix}\begin{pmatrix}\sum_{t=2}^n 1 & \sum_{t=2}^n y_{t-1}\\ \sum_{t=2}^n y_{t-1} & \sum_{t=2}^n y_{t-1}^2\end{pmatrix}^{-1}\begin{pmatrix}0\\ n^6\end{pmatrix}\left(\frac{\sum_{t=2}^n \tilde{e}_t^2}{n^4(n-3)}\right)\\
&= \begin{pmatrix}0 & 1\end{pmatrix}\begin{pmatrix}1 & \delta/4\\ \delta/4 & \delta^2/7\end{pmatrix}^{-1}\begin{pmatrix}0\\ 1\end{pmatrix}\frac{\delta^2}{45}+O_p\left(n^{-1}\right)=\frac{112}{405}+O_p\left(n^{-1}\right).
\end{aligned}$$

Consequently, we have

$$t_{\tilde{\beta}}/\sqrt{n}=\frac{n\left(\tilde{\beta}-1\right)}{n^{3/2}se\left(\tilde{\beta}\right)}=\frac{28/9}{\sqrt{112/405}}+O_p\left(n^{-1}\right)\xrightarrow{p}\sqrt{35}.$$

Proof of Theorem 4.3: (a) From Model (13), we have $\Delta^2 y_t = 6\delta(t-1)+\Delta^2 u_t$, which leads to

$$y_t=-6\delta+y_{t-1}+\Delta y_{t-1}+\omega_t \quad \text{with} \quad \omega_t=6\delta t+\Delta^2 u_t.$$

Hence, the centered LS estimator of the regression (8) with $k=1$ is

$$\begin{bmatrix}\breve{\alpha}+6\delta\\ \breve{\beta}-1\\ \breve{\psi}_1-1\end{bmatrix}=\begin{bmatrix}\sum_{t=3}^n 1 & \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n \Delta y_{t-1}\\ \sum_{t=3}^n y_{t-1} & \sum_{t=3}^n y_{t-1}^2 & \sum_{t=3}^n y_{t-1}\Delta y_{t-1}\\ \sum_{t=3}^n \Delta y_{t-1} & \sum_{t=3}^n y_{t-1}\Delta y_{t-1} & \sum_{t=3}^n \left(\Delta y_{t-1}\right)^2\end{bmatrix}^{-1}\begin{bmatrix}\sum_{t=3}^n \omega_t\\ \sum_{t=3}^n y_{t-1}\omega_t\\ \sum_{t=3}^n \Delta y_{t-1}\omega_t\end{bmatrix}.$$

Note that

$$n^{-2}\sum_{t=3}^n \omega_t=n^{-2}\sum_{t=3}^n\left(6\delta t+\Delta^2 u_t\right)=3\delta+O_p\left(n^{-1}\right)$$

$$\begin{aligned}
n^{-5}\sum_{t=3}^n y_{t-1}\omega_t &= n^{-5}\sum_{t=3}^n y_{t-1}\left(6\delta t+\Delta^2 u_t\right)=6\delta n^{-5}\sum_{t=3}^n y_{t-1}t+O_p\left(n^{-2}\right)\\
&= 6\delta^2/5+O_p\left(n^{-1}\right)
\end{aligned}$$

and

$$\begin{aligned}
n^{-4}\sum_{t=3}^n \Delta y_{t-1}\omega_t &= n^{-4}\sum_{t=3}^n\left[4\delta-3\delta t+3\delta(t-1)^2+\Delta u_{t-1}\right]\left[6\delta t+\Delta^2 u_t\right]\\
&= 18\delta^2 n^{-4}\sum_{t=3}^n t(t-1)^2+O_p\left(n^{-1}\right)=9\delta^2/2+O_p\left(n^{-1}\right)
\end{aligned}$$

Hence,

$$\begin{pmatrix}n^{-2} & 0 & 0\\ 0 & n^{-5} & 0\\ 0 & 0 & n^{-4}\end{pmatrix}\begin{bmatrix}\sum_{t=3}^n \omega_t\\ \sum_{t=3}^n y_{t-1}\omega_t\\ \sum_{t=3}^n \Delta y_{t-1}\omega_t\end{bmatrix}=\begin{bmatrix}3\delta\\ 6\delta^2/5\\ 9\delta^2/2\end{bmatrix}+O_p(1).$$

It can also be proved that $n^{-3}\sum_{t=3}^n \Delta y_{t-1}=\delta+O_p\left(n^{-1}\right)$, $n^{-6}\sum_{t=3}^n y_{t-1}\Delta y_{t-1}=\delta^2/2+O_p\left(n^{-1}\right)$, and $n^{-5}\sum_{t=3}^n\left(\Delta y_{t-1}\right)^2=9\delta^2/5+O_p\left(n^{-1}\right)$. Together with the limits of $n^{-4}\sum_{t=3}^{n} y_{t-1}$ and $n^{-7}\sum_{t=3}^{n} y_{t-1}^2$ obtained in the proof of Theorem (4.1), we have

$$
\begin{aligned}
&\begin{pmatrix} n^{-1} & 0 & 0 \\ 0 & n^2 & 0 \\ 0 & 0 & n \end{pmatrix}
\begin{bmatrix} \sum_{t=3}^{n} 1 & \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} \Delta y_{t-1} \\ \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} y_{t-1}^2 & \sum_{t=3}^{n} y_{t-1}\Delta y_{t-1} \\ \sum_{t=3}^{n} \Delta y_{t-1} & \sum_{t=3}^{n} y_{t-1}\Delta y_{t-1} & \sum_{t=3}^{n} \left(\Delta y_{t-1}\right)^2 \end{bmatrix}^{-1}
\begin{pmatrix} n^2 & 0 & 0 \\ 0 & n^5 & 0 \\ 0 & 0 & n^4 \end{pmatrix} \\
= &\begin{bmatrix} 1 & \delta/4 & \delta \\ \delta/4 & \delta^2/7 & \delta^2/2 \\ \delta & \delta^2/2 & 9\delta^2/5 \end{bmatrix}^{-1} + O_p\left(n^{-1}\right)
\end{aligned}
$$

As a result, we have

$$
\begin{pmatrix} n^{-1} & 0 & 0 \\ 0 & n^2 & 0 \\ 0 & 0 & n \end{pmatrix}
\begin{bmatrix} \check{\alpha}+6\delta \\ \check{\beta}-1 \\ \check{\psi}_1-1 \end{bmatrix}
= \begin{bmatrix} 1 & \delta/4 & \delta \\ \delta/4 & \delta^2/7 & \delta^2/2 \\ \delta & \delta^2/2 & 9\delta^2/5 \end{bmatrix}^{-1}
\begin{bmatrix} 3\delta \\ 6\delta^2/5 \\ 9\delta^2/2 \end{bmatrix} + O_p\left(n^{-1}\right)
$$

which leads to the result of $n^2\left(\check{\beta}-1\right) = -8.4 + O_p\left(n^{-1}\right)$.

To obtain the limit of $t_{\check{\beta}}$, we first have

$$
\begin{aligned}
& n^{-3}\sum_{t=3}^{n} \left(\check{e}_t\right)^2 \\
= \;& n^{-3}\sum_{t=3}^{n} \left(y_t - \check{\alpha} - \check{\beta} y_{t-1} - \check{\psi}_1 \Delta y_{t-1}\right)^2 \\
= \;& n^{-3}\sum_{t=3}^{n} \left(\omega_t - \left(\check{\alpha}+6\delta\right) - \left(\check{\beta}-1\right) y_{t-1} - \left(\check{\psi}_1-1\right)\Delta y_{t-1}\right)^2 \\
= \;& n^{-3}\sum_{t=3}^{n} \left(\omega_t - \left(\check{\alpha}+6\delta\right) - \left(\check{\beta}-1\right) y_{t-1} - \left(\check{\psi}_1-1\right)\Delta y_{t-1}\right)\omega_t \\
= \;& n^{-3}\sum_{t=3}^{n} \omega_t^2 - \begin{pmatrix} n^{-1}\left(\check{\alpha}+6\delta\right) & n^2\left(\check{\beta}-1\right) & n\left(\check{\psi}_1-1\right) \end{pmatrix}
\begin{bmatrix} n^{-2}\sum_{t=3}^{n} \omega_t \\ n^{-5}\sum_{t=3}^{n} y_{t-1}\omega_t \\ n^{-4}\sum_{t=3}^{n} \Delta y_{t-1}\omega_t \end{bmatrix} \\
= \;& 12\delta^2 - \begin{bmatrix} 3\delta \\ 6\delta^2/5 \\ 9\delta^2/2 \end{bmatrix}' \begin{bmatrix} 1 & \delta/4 & \delta \\ \delta/4 & \delta^2/7 & \delta^2/2 \\ \delta & \delta^2/2 & 9\delta^2/5 \end{bmatrix}^{\prime -1} \begin{bmatrix} 3\delta \\ 6\delta^2/5 \\ 9\delta^2/2 \end{bmatrix} + O_p\left(n^{-1}\right) \\
= \;& 0.03\delta^2 + O_p\left(n^{-1}\right)
\end{aligned}
$$

where the third equation is from the first-order conditions of LS regression, and the five equation comes from $n^{-3}\sum_{t=3}^{n} \omega_t^2 = n^{-3}\sum_{t=3}^{n} \left(6\delta t + \Delta^2 u_t\right)^2 = 12\delta^2 + O_p\left(n^{-1}\right)$. As a result, we have

$$
\begin{aligned}
& n^5\left[se\left(\check{\beta}\right)\right]^2 \\
= \;& \begin{pmatrix} 0 & n^2 & 0 \end{pmatrix}
\begin{bmatrix} \sum_{t=3}^{n} 1 & \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} \Delta y_{t-1} \\ \sum_{t=3}^{n} y_{t-1} & \sum_{t=3}^{n} y_{t-1}^2 & \sum_{t=3}^{n} y_{t-1}\Delta y_{t-1} \\ \sum_{t=3}^{n} \Delta y_{t-1} & \sum_{t=3}^{n} y_{t-1}\Delta y_{t-1} & \sum_{t=3}^{n} \left(\Delta y_{t-1}\right)^2 \end{bmatrix}^{-1}
\begin{pmatrix} 0 \\ n^5 \\ 0 \end{pmatrix} \frac{\sum_{t=3}^{n} \left(\check{e}_t\right)^2}{n^2\left(n-5\right)}
\end{aligned}
$$

$$= \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{bmatrix} 1 & \delta/4 & \delta \\ \delta/4 & \delta^2/7 & \delta^2/2 \\ \delta & \delta^2/2 & 9\delta^2/5 \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \left(0.03\delta^2\right) + O_p\left(n^{-1}\right)$$

$$= \frac{21 \times 64}{100} + O_p\left(n^{-1}\right).$$

Therefore,

$$t_{\breve{\beta}}/\sqrt{n} = \frac{n^2\left(\breve{\beta}-1\right)}{n^{5/2}se\left(\breve{\beta}\right)} \xrightarrow{p} -\sqrt{21}/2.$$

(b) Considering the regression (8) with $k = 2$:

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

it is confronted with the problem of perfect multi-collinearity as $\Delta y_{t-j} = \delta\left[3\left(t-j\right)^2 - 3\left(t-j\right) + 1\right] + \Delta u_{t-j}$, for $j = 1, 2$. From Model (13), it can be seen that $\Delta^2 y_t = 6\delta\left(t-1\right) + \Delta^2 u_t$ which leads to

$$\begin{aligned} y_t &= y_{t-1} + \Delta y_{t-1} + 6\delta\left(t-1\right) + \Delta^2 u_t \\ &= 6\delta + y_{t-1} + \Delta y_{t-1} + 6\delta\left(t-2\right) + \Delta^2 u_{t-1} + \Delta^2 u_t - \Delta^2 u_{t-1} \\ &= 6\delta + y_{t-1} + \Delta y_{t-1} + \Delta^2 y_{t-1} + \Delta^3 u_t \end{aligned}$$

where $\Delta^3 u_t = \Delta^2 u_t - \Delta^2 u_{t-1}$. We now consider the regression

$$y_t = \breve{\alpha}^* + \breve{\beta}^* y_{t-1} + \breve{\psi}_1^* \Delta y_{t-1} + \breve{\psi}_2^* \Delta^2 y_{t-1} + \breve{e}_t^*. \tag{25}$$

It can be proved that

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

which leads to $\breve{\beta} = \breve{\beta}^*$. It is also easy to prove that $t_{\breve{\beta}} = t_{\breve{\beta}^*}$. Note that the regression (25) does not face the problem of perfect multi-collinearity, and includes the true DGP of y_t when $\begin{pmatrix} \breve{\alpha}^* & \breve{\beta}^* & \breve{\psi}_1^* & \breve{\psi}_2^* \end{pmatrix} = \begin{pmatrix} 6\delta & 1 & 1 & 1 \end{pmatrix}$. It can be proved that $n^4\left(\breve{\beta}^* - 1\right) = O_p\left(1\right)$ and $\sqrt{n}t_{\breve{\beta}^*} = O_p\left(1\right)$. We omit the details for simplicity.

The same method can be extended to prove the results in the regression (8) with $k > 2$. We will still have $n^4\left(\breve{\beta}^* - 1\right) = O_p\left(1\right)$ and $\sqrt{n}t_{\breve{\beta}^*} = O_p\left(1\right)$, but the form of the limiting distributions may change as k varies.

D Proof of theorems in Section 5

Proof of Theorem 5.1: (a) We only give the proof for the regression (16) with $k = 1$. It can be extended straightforwardly to the regression with $k \geq 1$.

When the true DGP is $y_t = y_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim iid\left(0, \sigma^2\right)$, and the regression (16) with $k = 1$ is done, the centered LS estimator is

$$\begin{pmatrix} \check{\alpha} - 0 \\ \check{\beta} - 1 \\ \check{\psi}_1 - 0 \end{pmatrix} = \left[\sum_{t=2}^{n} \begin{pmatrix} 1 & y_{t-1} & \varepsilon_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\varepsilon_{t-1} \\ \varepsilon_{t-1} & y_{t-1}\varepsilon_{t-1} & \varepsilon_{t-1}^2 \end{pmatrix} \right]^{-1} \begin{pmatrix} \sum_{t=2}^{n} \varepsilon_t \\ \sum_{t=2}^{n} y_{t-1}\varepsilon_t \\ \sum_{t=2}^{n} \varepsilon_{t-1}\varepsilon_t \end{pmatrix}$$

Based on the results in Lemma (A.1) and the large sample theory for unit root process developed in the literature (see, for example, Phillips (1987) and Phillips and Perron (1989)), it can be proved that

$$\begin{pmatrix} n^{-1/2} & 0 & 0 \\ 0 & n^{-1} & 0 \\ 0 & 0 & n^{-1} \end{pmatrix} \begin{pmatrix} \sum_{t=2}^{n} \varepsilon_t \\ \sum_{t=2}^{n} y_{t-1}\varepsilon_t \\ \sum_{t=2}^{n} \varepsilon_{t-1}\varepsilon_t \end{pmatrix} \Rightarrow \begin{pmatrix} \sigma W(1) \\ \sigma^2 \int_0^1 W(r)\, dW(r) \\ 0 \end{pmatrix}$$

and

$$\begin{pmatrix} n^{1/2} & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & 1 \end{pmatrix} \left[\sum_{t=2}^{n} \begin{pmatrix} 1 & y_{t-1} & \varepsilon_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\varepsilon_{t-1} \\ \varepsilon_{t-1} & y_{t-1}\varepsilon_{t-1} & \varepsilon_{t-1}^2 \end{pmatrix} \right]^{-1} \begin{pmatrix} n^{1/2} & 0 & 0 \\ 0 & n & 0 \\ 0 & 0 & n \end{pmatrix}$$
$$\Rightarrow \begin{pmatrix} 1 & \sigma \int_0^1 W(r)\, dr & \sigma W(1) \\ \sigma \int_0^1 W(r)\, dr & \sigma^2 \int_0^1 [W(r)]^2\, dr & \sigma^2 \int_0^1 W(r)\, dW(r) + \sigma^2 \\ 0 & 0 & \sigma^2 \end{pmatrix}^{-1} \equiv \Pi^{-1}.$$

Hence,

$$\begin{pmatrix} \sqrt{n}\,(\check{\alpha} - 0) \\ n\left(\check{\beta} - 1\right) \\ \left(\check{\psi}_1 - 0\right) \end{pmatrix} \Rightarrow \Pi^{-1} \begin{pmatrix} \sigma W(1) \\ \sigma^2 \int_0^1 W(r)\, dW(r) \\ 0 \end{pmatrix},$$

which leads to

$$\begin{pmatrix} \sqrt{n}\,(\check{\alpha} - 0) \\ n\left(\check{\beta} - 1\right) \end{pmatrix} \Rightarrow \begin{pmatrix} 1 & \sigma \int_0^1 W(r)\, dr \\ \sigma \int_0^1 W(r)\, dr & \sigma^2 \int_0^1 [W(r)]^2\, dr \end{pmatrix}^{-1} \begin{pmatrix} \sigma W(1) \\ \sigma^2 \int_0^1 W(r)\, dW(r) \end{pmatrix}$$

and

$$n\left(\check{\beta} - 1\right) \Rightarrow \frac{\int_0^1 W(r)\, dW(r) - W(1) \int_0^1 W(r)\, dr}{\int_0^1 [W(r)]^2\, dr - \left(\int_0^1 W(r)\, dr\right)^2}.$$

Considering that the regression (16) with $k = 1$ covers the true DGP and all parameters are consistently estimated, it can be shown that

$$\begin{aligned} \frac{1}{n-4} \sum_{t=2}^{n} \check{e}_t^2 &= \frac{1}{n-4} \sum_{t=2}^{n} \left(y_t - \check{\alpha} - \check{\beta} y_{t-1} - \check{\psi}_1 \Delta y_{t-1}\right)^2 \\ &= \frac{1}{n-4} \sum_{t=2}^{n} \varepsilon_t^2 + o_p(1) \xrightarrow{p} \sigma^2. \end{aligned}$$

Consequently,

$$n^2 \left[se\left(\check{\beta}\right)\right]^2 = \begin{bmatrix} 0 & n & 0 \end{bmatrix} \left[\sum_{t=2}^{n} \begin{pmatrix} 1 & y_{t-1} & \varepsilon_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\varepsilon_{t-1} \\ \varepsilon_{t-1} & y_{t-1}\varepsilon_{t-1} & \varepsilon_{t-1}^2 \end{pmatrix}\right]^{-1} \begin{bmatrix} 0 \\ n \\ 0 \end{bmatrix} \frac{\sum_{t=2}^{n} \check{e}_t^2}{n-4}$$

$$\Rightarrow \left[\int_0^1 \left[W(r)\right]^2 dr - \left(\int_0^1 W(r)\, dr\right)^2\right]^{-1}.$$

Together with the limit of $n\left(\check{\beta}-1\right)$, we have

$$t_{\check{\beta}} = \frac{n\left(\check{\beta}-1\right)}{n\left[se\left(\check{\beta}\right)\right]} \Rightarrow \frac{\int_0^1 W(r)\, dW(r) - W(1)\int_0^1 W(r)\, dr}{\left\{\int_0^1 \left[W(r)\right]^2 dr - \left(\int_0^1 W(r)\, dr\right)^2\right\}^{1/2}}.$$

The results in Part (b) are the standard results of augmented Dickey-Fuller tests as given in Dickey and Fuller (1979). The results in Part (c) have already been obtained in Theorem (2.3), Theorem (3.3) and Theorem (4.3).

(d) For simplicity, we only prove the limits of $n\left(\check{\beta}-1\right)$ and $t_{\check{\beta}} = \left(\check{\beta}-1\right)/se\left(\check{\beta}\right)$ based on the regression (16) with $k=1$. The results for the regression with $k>1$ can be proved similarly. We also assume $u_t = \varepsilon_t \sim iid\left(0,\sigma^2\right)$. With the use of the large sample theory of mildly explosive process with serially dependent errors developed in Phillips and Magdalinos (2005) and Magdalinos (2012), the same approach can be straightforwardly applied to prove the results when u_t is a weakly stationary process.

Given $y_{t-1} = \rho_n y_{t-2} + u_{t-1}$ with $\rho_n = 1 + c/n^\theta$, Phillips and Magdalinos (2007) showed $y_n = O_p\left(\rho_n^n n^{\theta/2}\right)$. We then have

$$\Delta y_{t-1} = y_{t-1} - \rho_n^{-1}\left(y_{t-1} - u_{t-1}\right) = \frac{\rho_n - 1}{\rho_n} y_{t-1} + \rho_n^{-1} u_{t-1} = \frac{c}{n^\theta \rho_n} y_{t-1} + \rho_n^{-1} u_{t-1}$$

where the first term dominates the second term when t is large. Hence, the regression (16) with $k=1$ faces the problem of perfect multi-collinearity. We turn to consider the transformed regression as

$$y_t = \check{\alpha}^* + \check{\beta}^* y_{t-1} + \check{\psi}_1^*\left(\rho_n^{-1} u_{t-1}\right) + \check{e}_t^*, \tag{26}$$

whose centered LS estimators have the following relationship with the centered LS estimators of the regression (16) with $k=1$:

$$\begin{pmatrix} \check{\alpha} - 0 \\ \check{\beta} - \rho_n \\ \check{\psi}_1 - 0 \end{pmatrix} = D' \begin{pmatrix} \check{\alpha}^* - 0 \\ \check{\beta}^* - \rho_n \\ \check{\psi}_1^* - 0 \end{pmatrix} \quad \text{with} \quad D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{c}{n^\theta \rho_n} & 1 \end{pmatrix},$$

which leads to

$$\check{\beta} - \rho_n = \left(\check{\beta}^* - \rho_n\right) - \frac{c}{n^\theta \rho_n}\left(\check{\psi}_1^* - 0\right).$$

Note that

$$\begin{pmatrix}\check{\alpha}^*-0\\ \check{\beta}^*-\rho_n\\ \check{\psi}_1^*-0\end{pmatrix}=\left[\sum_{t=2}^{n}\begin{pmatrix}1 & y_{t-1} & \rho_n^{-1}u_{t-1}\\ y_{t-1} & y_{t-1}^2 & y_{t-1}\left(\rho_n^{-1}u_{t-1}\right)\\ \left(\rho_n^{-1}u_{t-1}\right) & y_{t-1}\left(\rho_n^{-1}u_{t-1}\right) & \left(\rho_n^{-1}u_{t-1}\right)^2\end{pmatrix}\right]^{-1}\begin{pmatrix}\sum_{t=2}^{n}u_t\\ \sum_{t=2}^{n}y_{t-1}u_t\\ \sum_{t=2}^{n}\left(\rho_n^{-1}u_{t-1}\right)u_t\end{pmatrix}.$$

With the assumption of $u_t=\varepsilon_t\sim iid\left(0,\sigma^2\right)$, Phillips and Magdalinos (2007) proved that

$$n^{-2\theta}\rho_n^{-2n}\sum_{t=2}^{n}y_{t-1}^2\Rightarrow\frac{\eta^2}{4c^2}\quad\text{and}\quad n^{-\theta}\rho_n^{-n}\sum_{t=2}^{n}y_{t-1}u_t\Rightarrow\frac{\eta\xi}{2c}$$

where η and ξ are two independent and standard normally distributed random variables. We then have

$$\begin{aligned}n^{-\theta}\rho_n^{-n}\sum_{t=2}^{n}y_{t-1}\left(\rho_n^{-1}u_{t-1}\right) &= n^{-\theta}\rho_n^{-n}\left(\sum_{t=2}^{n}y_{t-2}u_{t-1}+\rho_n^{-1}\sum_{t=2}^{n}u_{t-2}u_{t-1}\right)\\ &= n^{-\theta}\rho_n^{-n}\left(\sum_{t=2}^{n}y_{t-1}u_t+y_0u_1-y_{n-1}u_n+\rho_n^{-1}\sum_{t=2}^{n}u_{t-2}u_{t-1}\right)\\ &= n^{-\theta}\rho_n^{-n}\sum_{t=2}^{n}y_{t-1}u_t+o_p(1)\Rightarrow\frac{\eta\xi}{2c}\end{aligned}$$

Together with the result in Wang and Yu (2016) that

$$n^{-3\theta/2}\rho_n^{-n}\sum_{t=2}^{n}y_{t-1}\Rightarrow\frac{\eta}{c\sqrt{2c}},$$

we now have

$$\begin{pmatrix}n^{-1/2} & 0 & 0\\ 0 & n^{-\theta}\rho_n^{-n} & 0\\ 0 & 0 & n^{-1}\end{pmatrix}\begin{pmatrix}\sum_{t=2}^{n}u_t\\ \sum_{t=2}^{n}y_{t-1}u_t\\ \sum_{t=2}^{n}\left(\rho_n^{-1}u_{t-1}\right)u_t\end{pmatrix}\Rightarrow\begin{pmatrix}\sigma W(1)\\ \eta\xi/(2c)\\ 0\end{pmatrix},$$

and

$$\begin{aligned}&\begin{pmatrix}n^{1/2} & 0 & 0\\ 0 & n^{\theta}\rho_n^{n} & 0\\ 0 & 0 & 1\end{pmatrix}\left[\sum_{t=2}^{n}\begin{pmatrix}1 & y_{t-1} & \rho_n^{-1}u_{t-1}\\ y_{t-1} & y_{t-1}^2 & y_{t-1}\left(\rho_n^{-1}u_{t-1}\right)\\ \left(\rho_n^{-1}u_{t-1}\right) & y_{t-1}\left(\rho_n^{-1}u_{t-1}\right) & \left(\rho_n^{-1}u_{t-1}\right)^2\end{pmatrix}\right]^{-1}\begin{pmatrix}n^{1/2} & 0 & 0\\ 0 & n^{\theta}\rho_n^{n} & 0\\ 0 & 0 & n\end{pmatrix}\\ \Rightarrow\ &\begin{pmatrix}1 & 0 & \sigma W(1)\\ 0 & \eta^2/\left(4c^2\right) & \eta\xi/(2c)\\ 0 & 0 & \sigma^2\end{pmatrix}^{-1}\end{aligned}$$

As a result,

$$\begin{pmatrix}n^{1/2} & 0 & 0\\ 0 & n^{\theta}\rho_n^{n} & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}\check{\alpha}^*-0\\ \check{\beta}^*-\rho_n\\ \check{\psi}_1^*-0\end{pmatrix}\Rightarrow\begin{pmatrix}1 & 0 & \sigma W(1)\\ 0 & \eta^2/\left(4c^2\right) & \eta\xi/(2c)\\ 0 & 0 & \sigma^2\end{pmatrix}^{-1}\begin{pmatrix}\sigma W(1)\\ \eta\xi/(2c)\\ 0\end{pmatrix}$$

which leads to

$$n^{\theta}\rho_n^n\left(\check{\beta}^* - \rho_n\right) \Rightarrow 2c\frac{\xi}{\eta} \text{ and } \check{\psi}_1^* \xrightarrow{p} 0.$$

Consequently, we have

$$\begin{aligned}
n\left(\check{\beta}-1\right) &= n\left(\check{\beta}-\rho_n\right) + n\left(\rho_n - 1\right) \\
&= n\left(\check{\beta}^*-\rho_n\right) - \frac{nc}{n^{\theta}\rho_n}\left(\check{\psi}_1^* - 0\right) + n\left(\rho_n - 1\right) \\
&= n\left(\check{\beta}^*-\rho_n\right) - \frac{nc}{n^{\theta}\rho_n}\left(\check{\psi}_1^* - 0\right) + n\frac{c}{n^{\theta}} \\
&= n\left(\check{\beta}^*-\rho_n\right) + \frac{nc}{n^{\theta}}\left(1 - \rho_n^{-1}\check{\psi}_1^*\right) \longrightarrow +\infty
\end{aligned}$$

where the last limit comes from the fact that

$$n\left(\check{\beta}^* - \rho_n\right) = O_p\left(\frac{n}{n^{\theta}\rho_n^n}\right) = o_p(1) \text{ and } \left(1 - \rho_n^{-1}\check{\psi}_1^*\right) \xrightarrow{p} 1.$$

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

$$\begin{aligned}
\frac{1}{n}\sum_{t=2}^{n}\left(\check{e}_t\right)^2 &= \frac{1}{n}\sum_{t=2}^{n}\left[y_t - \check{\alpha} - \check{\beta}y_{t-1} - \check{\psi}_1\Delta y_{t-1}\right]^2 \\
&= \frac{1}{n}\sum_{t=2}^{n}\left[y_t - \check{\alpha}^* - \check{\beta}^* y_{t-1} - \check{\psi}_1^*\left(\rho_n^{-1}u_{t-1}\right)\right]^2 = \frac{1}{n}\sum_{t=2}^{n}\left(\check{e}_t^*\right)^2
\end{aligned}$$

Given that the true DGP of y_t is covered by the transformed regression (26) and the LS estimates of the parameters are consistent, it can be shown that

$$\frac{1}{n}\sum_{t=2}^{n}\left(\check{e}_t^*\right)^2 = \frac{1}{n}\sum_{t=2}^{n}\left(u_t\right)^2 + o_p(1) \xrightarrow{p} \sigma^2.$$

Note that

$$\begin{aligned}
& n^{2\theta}\begin{bmatrix}0 & 1 & 0\end{bmatrix}\left[\sum_{t=2}^{n}\begin{pmatrix}1 & y_{t-1} & \Delta y_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\Delta y_{t-1} \\ \Delta y_{t-1} & y_{t-1}\Delta y_{t-1} & \left(\Delta y_{t-1}\right)^2\end{pmatrix}\right]^{-1}\begin{bmatrix}0\\1\\0\end{bmatrix} \\
= \; & n^{2\theta}\begin{bmatrix}0 & 1 & 0\end{bmatrix}D'\left[\sum_{t=2}^{n}D\begin{pmatrix}1 & y_{t-1} & \Delta y_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\Delta y_{t-1} \\ \Delta y_{t-1} & y_{t-1}\Delta y_{t-1} & \left(\Delta y_{t-1}\right)^2\end{pmatrix}D'\right]^{-1}D\begin{bmatrix}0\\1\\0\end{bmatrix} \\
= \; & n^{2\theta}\begin{bmatrix}0 & 1 & -\frac{c}{n^{\theta}\rho_n}\end{bmatrix}\left[\sum_{t=2}^{n}\begin{pmatrix}1 & y_{t-1} & \rho_n^{-1}u_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\left(\rho_n^{-1}u_{t-1}\right) \\ \left(\rho_n^{-1}u_{t-1}\right) & y_{t-1}\left(\rho_n^{-1}u_{t-1}\right) & \left(\rho_n^{-1}u_{t-1}\right)^2\end{pmatrix}\right]^{-1}\begin{bmatrix}0\\1\\-\frac{c}{n^{\theta}\rho_n}\end{bmatrix} \\
= \; & n^{2\theta}\begin{bmatrix}0 & 1 & -\frac{c}{n^{\theta}\rho_n}\end{bmatrix}\begin{pmatrix}n^{-1/2} & 0 & 0 \\ 0 & n^{-\theta}\rho_n^{-n} & 0 \\ 0 & 0 & 1\end{pmatrix}\left\{\begin{pmatrix}1 & 0 & \sigma W(1) \\ 0 & \eta^2/\left(4c^2\right) & \eta\xi/(2c) \\ 0 & 0 & \sigma^2\end{pmatrix}^{-1} + o_p(1)\right\}
\end{aligned}$$

$$
\times \begin{pmatrix} n^{-1/2} & 0 & 0 \\ 0 & n^{-\theta}\rho_n^{-n} & 0 \\ 0 & 0 & n^{-1} \end{pmatrix} \begin{bmatrix} 0 \\ 1 \\ -\frac{c}{n^\theta \rho_n} \end{bmatrix}
$$

$$
\begin{aligned}
&= n^{2\theta} \begin{bmatrix} 0 & n^{-\theta}\rho_n^{-n} & -\frac{c}{n^\theta \rho_n} \end{bmatrix} \left\{ \begin{pmatrix} 1 & 0 & \sigma W(1) \\ 0 & \eta^2/\left(4c^2\right) & \eta\xi/(2c) \\ 0 & 0 & \sigma^2 \end{pmatrix}^{-1} + o_p(1) \right\} \begin{bmatrix} 0 \\ n^{-\theta}\rho_n^{-n} \\ -\frac{c}{n^{1+\theta}\rho_n} \end{bmatrix} \\
&= \begin{bmatrix} 0 & \rho_n^{-n} & -\frac{c}{\rho_n} \end{bmatrix} \left\{ \begin{pmatrix} 1 & 0 & \sigma W(1) \\ 0 & \eta^2/\left(4c^2\right) & \eta\xi/(2c) \\ 0 & 0 & \sigma^2 \end{pmatrix}^{-1} + o_p(1) \right\} \begin{bmatrix} 0 \\ \rho_n^{-n} \\ -\frac{c}{n\rho_n} \end{bmatrix} \\
&\xrightarrow{p} \begin{bmatrix} 0 & 0 & -c \end{bmatrix} \begin{pmatrix} 1 & 0 & \sigma W(1) \\ 0 & \eta^2/\left(4c^2\right) & \eta\xi/(2c) \\ 0 & 0 & \sigma^2 \end{pmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
n^{2\theta}\left[se\left(\check{\beta}\right)\right]^2 &= n^{2\theta} \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \left[\sum_{t=2}^{n} \begin{pmatrix} 1 & y_{t-1} & \Delta y_{t-1} \\ y_{t-1} & y_{t-1}^2 & y_{t-1}\Delta y_{t-1} \\ \Delta y_{t-1} & y_{t-1}\Delta y_{t-1} & \left(\Delta y_{t-1}\right)^2 \end{pmatrix} \right]^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \frac{\sum_{t=2}^{n} \left(\check{e}\right)^2}{n-4} \\
&\xrightarrow{p} 0.
\end{aligned}
$$

Therefore, as $n \to \infty$,

$$
\begin{aligned}
t_{\check{\beta}} &= \frac{\check{\beta}-1}{se\left(\check{\beta}\right)} = \frac{n^\theta\left(\check{\beta}-1\right)}{n^\theta se\left(\check{\beta}\right)} = \frac{n^\theta\left(\check{\beta}-\rho_n\right) + n^\theta\left(\rho_n - 1\right)}{n^\theta se\left(\check{\beta}\right)} \\
&= \frac{o_p(1)+c}{o_p(1)} \xrightarrow{p} +\infty.
\end{aligned}
$$

References

Dickey, D. A., and W. A. Fuller, 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 421-431.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press.

Homm, U., and J. Breitung, 2012. Testing for speculative bubbles in stock markets: A comparison of alternative methods. *Journal of Financial Econometrics*, 10, 198-231.

Harvey, D. I., Leybourne, S. J. and Zu, Y., 2017a. Testing explosive bubbles with time-varying volatility. University of Nottingham, Working Paper.

Harvey, D. I., Leybourne, S. J. and Zu, Y., 2017b. Sign-based unit root tests for explosive financial bubbles in the presence of nonstationary volatility. University of Nottingham, Working Paper.

Harvey, D., S. J. Leybourne, R. Sollis and A. M. R. Taylor, 2016. Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance* 38, 548–574.

Magdalinos, T., 2012. Mildly explosive autoregression under weak and strong dependence. *Journal of Econometrics* 169, 179-187.

Phillips, P.C.B., 1987a. Time series regression with a unit root. *Econometrica* 55, 277-301.

Phillips, P.C.B. and T. Magdalinos, 2005. Limit theory for moderate deviations from a unit root under weak dependence. Working paper.

Phillips, P.C.B. and T. Magdalinos, 2007. Limit theory for moderate deviations from a unit root. *Journal of Econometrics* 136, 115-130.

Phillips, P.C.B., and S.P. Shi, 2017. Financial bubble implosion and reverse regression. *Econometric Theory,* forthcoming.

Phillips, P.C.B., S.P. Shi and J. Yu, 2015a. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P500. *International Economic Review*, 56(4), 1043-1078.

Phillips, P.C.B., S.P. Shi and J. Yu, 2015b. Testing for multiple bubbles: Limit theory of real time detector. *International Economic Review*, 56(4), 1079-1134.

Phillips, P.C.B., Y. Wu and J. Yu, 2011. Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review* 52, 201-226.

Phillips, P.C.B., and J. Yu, 2011. Dating the timeline of financial bubbles during the subprime crisis. *Quantitative Economics* 2, 455-491.

Wang X. and J. Yu, 2016. Double asymptotics for explosive continuous time models. *Journal of Econometrics*, 193, 35-53.