TESTING FOR MULTIPLE BUBBLES

By

Peter C.B. Phillips, Shu-Ping Shi, and Jun Yu

January 2012

Testing for Multiple Bubbles[*]

Peter C. B. Phillips

Yale University, University of Auckland,
University of Southampton & Singapore Management University

Shu-Ping Shi   Jun Yu

The Australian National University   *Singapore Management University*

December 23, 2011

Abstract

Identifying and dating explosive bubbles when there is periodically collapsing behavior over time has been a major concern in the economics literature and is of great importance for practitioners. The complexity of the nonlinear structure inherent in multiple bubble phenomena within the same sample period makes econometric analysis particularly difficult. The present paper develops new recursive procedures for practical implementation and surveillance strategies that may be employed by central banks and fiscal regulators. We show how the testing procedure and dating algorithm of Phillips, Wu and Yu (2011, PWY) are affected by multiple bubbles and may fail to be consistent. The present paper proposes a generalized version of the sup ADF test of PWY to address this difficulty, derives its asymptotic distribution, introduces a new date-stamping strategy for the origination and termination of multiple bubbles, and proves consistency of this dating procedure. Simulations show that the test significantly improves discriminatory power and leads to distinct power gains when multiple bubbles occur. Empirical applications are conducted to S&P 500 stock market data over a long historical period from January 1871 to December 2010. The new approach identifies many key historical episodes of exuberance and collapse over this period, whereas the strategy of PWY and the CUSUM procedure locate far fewer episodes in the same sample range.

Keywords: Date-stamping strategy; Generalized sup ADF test; Multiple bubbles, Rational bubble; Periodically collapsing bubbles; Sup ADF test;

JEL classification: C15, C22

Economists have taught us that it is unwise and unnecessary to combat asset price bubbles and excessive credit creation. Even if we were unwise enough to wish to prick an asset price bubble, we are told it is impossible to see the bubble while it is in its inflationary phase. (George Cooper, 2008)

1   Introduction

As financial historians have argued recently (Ahamed, 2009; Ferguson, 2008), financial crises are often preceded by an asset market bubble or rampant credit growth. The global financial crisis of 2007-2009 is no exception. In its aftermath, central bank economists and policy makers are now affirming the recent Basil III accord to work to stabilize the financial system by way of guidelines on capital requirements and related measures to control "excessive credit creation". In this process of control, an important practical issue of market surveillance involves the assessment of what is "excessive". But as Cooper (2008) puts it in the header cited above from his recent bestseller, many economists have declared the task to be impossible and that it is imprudent to seek to combat asset price bubbles. How then can central banks and regulators work to offset a speculative bubble when they are unable to assess whether one exists and are considered unwise to take action if they believe one does exist?

One contribution that econometric techniques can offer in this complex exercise of market surveillance and policy action is the detection of exuberance in financial markets by explicit quantitative measures. These measures are not simply ex post detection techniques but anticipative dating algorithm that can assist regulators in their market monitoring behavior by means of early warning diagnostic tests. If history has a habit of repeating itself and human learning mechanisms do fail, as financial historians such as Ferguson (2008)[1] assert, then quantitative warnings may serve as useful alert mechanisms to both market participants and regulators.

Several attempts to develop econometric tests have been made in the literature going back some decades (see Gurkaynak, 2008, for a recent review). Phillips, Wu and Yu (2011, PWY hereafter) recently proposed a method which can detect exuberance in asset price series during

---

[1] "Nothing illustrates more clearly how hard human beings find it to learn from history than the repetitive history of stock market bubbles." Ferguson (2008).

an inflationary phase. The approach is anticipative as an early warning alert system, so that it meets the needs of central bank surveillance teams and regulators, thereby addressing one of the key concerns articulated by Cooper (2008). The method is especially effective when there is a single bubble episode in the sample data, as in the 1990s Nasdaq episode analyzed in the PWY paper and in the 2000s U.S. house price bubble analyzed in Phillips and Yu (2011).

Just as historical experience confirms the existence of many financial crises (Ahamed reports 60 different financial crises since the 17th century[2]), when the sample period is long enough there will often be evidence of multiple asset price bubbles in the data. The econometric identification of multiple bubbles with periodically collapsing behavior over time is substantially more difficult than identifying a single bubble. The difficulty in practice arises from the complex nonlinear structure involved in multiple bubble phenomena which typically diminishes the discriminatory power of existing test mechanisms such as those given in PWY. These power reductions complicate attempts at econometric dating and enhance the need for new approaches that do not suffer from this problem. If econometric methods are to be useful in practical work conducted by surveillance teams they need to be capable of dealing with multiple bubble phenomena. Of particular concern in financial surveillance is the reliability of a warning alert system that points to inflationary upturns in the market. Such warning systems ideally need to have a low false detection rate to avoid unnecessary policy measures and a high positive detection rate that ensures early and effective policy implementation.

The present paper responds to this need by providing a new framework for testing and dating bubble phenomena when there are potentially multiple bubbles in the data. The mechanisms developed here extend those of PWY by allowing for variable window widths in the recursive regressions on which the test procedures are based. The new mechanisms are shown in simulations to substantially increase discriminatory power in the tests and dating strategies. The paper contributes further by providing a limit theory for the new tests, by proving the consistency of the dating mechanisms, and by showing the inconsistency of certain versions of the

---

[2] "Financial booms and busts were, and continue to be, a feature of the economic landscape. These bubbles and crises seem to be deep-rooted in human nature and inherent to the capitalist system. By one count there have been 60 different crises since the 17th century." Ahamed (2009).

PWY dating strategy when multiple bubbles occur. The final contribution of the paper is to apply the techniques to a long historical series of US stock market data where multiple financial crises and episodes of exuberance and collapse have occurred.

One starting point in the analysis of financial bubbles is the standard asset pricing equation:

$$P_t = \sum_{i=0}^{\infty} \left(\frac{1}{1+r_f}\right)^i \mathbb{E}_t \left(D_{t+i} + U_{t+i}\right) + B_t, \tag{1}$$

where P_t is the after-dividend price of the asset, D_t is the payoff received from the asset (i.e. dividend), r_f is the risk-free interest rate, U_t represents the unobservable fundamentals and B_t is the bubble component. The quantity $P_t^f = P_t - B_t$ is often called the market fundamental. The pricing equation (1) is not the only model to accommodate bubble phenomena and there is continuing professional debate over how (or even whether) to include bubble components in asset pricing models (see, for example, the discussion in Cochrane, 2005, pp. 402-404) and their relevance in empirical work (notably, Pástor and Veronesi, 2006, but note also the strong critique of that view in Cooper, 2008[3]). There is greater agreement on the existence of market exuberance (which may be rational or irrational depending on possible links to market fundamentals), crises and panics (Kindelberger and Aliber, 2005; Ferguson, 2008). For instance, financial exuberance might originate in pricing errors relative to fundamentals that arise from behavioral factors, or fundamental values may themselves be highly sensitive to changes in the discount rate, which can lead to price run ups that mimic the inflationary phase of a bubble. With regard to the latter, Phillips and Yu (2011) show that in certain dynamic structures a time-varying discount rate can induce temporary explosive behavior in asset prices. Similar considerations may apply in more general stochastic discount factor asset pricing equations. Whatever its origins, explosive or mildly explosive behavior in asset prices is a primary indicator of market exuberance during the inflationary phase of a bubble and this time series manifestation may be subjected to econometric testing.

---

[3] "People outside the world of economics may be amazed to know that a significant body of researchers are still engaged in the task of proving that the pricing of the NASDAQ stock market correctly reflected the market's true value throughout the period commonly known as the NASDAQ bubble.... The intellectual contortions required to rationalize all of these prices beggars belief." (Cooper, 2008, p.9).

Diba and Grossman (1988) argued that the bubble component in (1) has an explosive property characterized by the following submartingale property:

$$\mathbb{E}_t \left(B_{t+1} \right) = \left(1 + r_f \right) B_t. \tag{2}$$

In the absence of bubbles (i.e. $B_t = 0$), the degree of nonstationarity of the asset price is controlled by the character of the dividend series and unobservable fundamentals. For example, if D_t is an $I(1)$ process and U_t is either an $I(0)$ or an $I(1)$ process, then the asset price is at most an $I(1)$ process. On the other hand, given the submartingale behavior (2), asset prices will be explosive in the presence of bubbles. Therefore, when unobservable fundamentals are at most $I(1)$ and D_t is stationary after differencing, empirical evidence of explosive behavior in asset prices may be used to conclude the existence of bubbles.[4] Based on this argument, Diba and Grossman (1988) suggest conducting right-tailed unit root tests (against explosive alternatives) on the asset price and the observable fundamental (i.e. dividend) to detect the existence of bubbles. This method is then referred to as the conventional cointegration-based bubble test.

Evans (1991) demonstrated that this conventional cointegration-based test is not capable of detecting explosive bubbles when they manifest periodically collapsing behavior in the sample (Blanchard, 1979).[5] Collapse signals a break or nonlinearity in the generating mechanism that needs to be accommodated in testing and modeling this type of market phenomena. In partial response to the Evans critique a number of papers have been written proposing extended versions of the conventional cointegration-based test that have some power in detecting periodically collapsing bubbles.

---

[4]This argument also applies to the logarithmic asset price and the logarithmic dividend under certain conditions. This is due to the fact that in the absence of bubbles, equation (1) can be rewritten as

$$(1 - \rho) p_t^f = \kappa + \rho e^{\bar{d} - \bar{p}} d_t + \rho e^{\bar{u} - \bar{p}} u_t + e^{\bar{d} - \bar{p}} \sum_{j=1}^{\infty} \rho^j \mathbb{E}_t \left[\triangle d_{t+j} \right] + e^{\bar{u} - \bar{p}} \sum_{j=1}^{\infty} \rho^j \mathbb{E}_t \left[\triangle u_{t+j} \right],$$

where $p_t^f = \log(P_t^f)$, $d_t = \log(D_t)$, $u_t = \log(U_t)$, $\rho = (1 + r_f)^{-1}$, κ is a constant, $\bar{p}, \bar{d}$ and $\bar{u}$ are the respective sample means of p_t^f, d_t and u_t. The degree of nonstationary of p_t^f is determined by that of d_t and u_t. Lee and Phillips (2011) provide a detailed analysis of the accuracy of this log linear approximation under various conditions.

[5]The failure of the cointegration based test is further studied in Charemza and Deadman (1995) within the setting of bubbles with stochastic explosive roots.

The approach adopted in PWY (2011) uses a sup ADF test (or forward recursive right-tailed ADF test). PWY suggest implementing the right-tailed ADF test repeatedly on a forward expanding sample sequence and performing inference based on the sup value of the corresponding ADF statistic sequence. They show that the sup ADF (SADF) test significantly improves power compared with the conventional cointegration-based test. This test also gives rise to an associated dating strategy which identifies points of origination and termination of a bubble. When there is a single bubble in the data, it is known that this dating strategy is consistent, as first shown in the working paper by Phillips and Yu (2009). Other break testing procedures such as Chow tests, model selection, and CUSUM tests may also be applied as dating mechanisms. Extensive simulations conducted by Homm and Breitung (2011) indicate that the PWY procedure works satisfactorily against other recursive (as distinct from full sample) procedures for structural breaks and is particularly effective as a real time bubble detection algorithm. Importantly, the procedure can detect market exuberance arising from a variety of sources, including mildly explosive behavior that may be induced by changing fundamentals such as a time-varying discount factor.

The present paper demonstrates that when the sample period includes multiple episodes of exuberance and collapse, the SADF test may suffer from reduced power and can be inconsistent, failing to reveal the existence of bubbles. This weakness is a particular drawback in analyzing long time series or rapidly changing market data where more than one episode of exuberance is suspected. To overcome this weakness, we propose an alternative approach named the *generalized sup ADF* (GSADF) test. The GSADF test is also based on the idea of repeatedly implementing a right-tailed ADF test, but the new test extends the sample sequence to a broader and more flexible range. Instead of fixing the starting point of the sample (namely, on the first observation of the sample), the GSADF test extends the sample sequence by changing both the starting point and the ending point of the sample over a feasible range of flexible windows.

The sample sequences used in the SADF and GSADF tests are designed to capture any explosive behavior manifested within the overall sample and ensure that there are sufficient observations to initiate the recursion. Since the GSADF test covers more subsamples of the

data and has greater window flexibility, it is expected to outperform the SADF test in detecting explosive behavior in multiple episodes. This enhancement in performance by the GSADF test is demonstrated in simulations which compare the two tests in terms of their size and power in bubble detection. The paper also derives the asymptotic distribution of the GSADF statistic in comparison with that of the SADF statistic.

A further contribution of the paper is to develop a new dating strategy and provide a limit theory that confirms the consistency of the dating mechanism. The recursive ADF test is used in PWY to date stamp the origination and termination of a bubble. More specifically, the recursive procedure compares the ADF statistic sequence against critical values for the standard right-tailed ADF statistic and uses a first crossing time occurrence to date origination and collapse. For the generalized sup ADF test, we recommend a new date-stamping strategy, which compares the backward sup ADF (BSADF) statistic sequence with critical values for the sup ADF statistic, where the BSADF statistics are obtained from implementing the right-tailed ADF test on backward expanding sample sequences.

For a data generating process with only one bubble episode in the sample period, we show that both date-stamping strategies successfully estimate the origination and termination of a single bubble consistently. We then consider a situation in which there are two bubbles in the sample period and allow the duration of the first bubble to be longer or shorter than the second one. We demonstrate that the date-stamping strategy of PWY cannot consistently estimate the origination and termination of a (shorter) second bubble, whereas the strategy proposed in this paper can consistently estimate the origination and termination of each bubble. The same technology is applicable and similar results apply in multiple bubble scenarios.

The organization of the paper is as follows. The new test and its limit theory are given in Section 2. Section 3 discusses the implementation of the new test and investigates size and power. Section 4 proposes a date-stamping strategy based on the new test and derives the consistency properties of this strategy and the PWY strategy under both single bubble and twin bubble alternatives. An alternative sequential implementation of the PWY procedure is developed which is shown to be capable of consistent date estimation in a twin bubble scenario.

In Section 5, these procedures and the CUSUM test are applied to S&P 500 price-dividend ratio data over a long historical period from January 1871 to December 2010. Section 6 concludes. Two appendices contain supporting lemmas and derivations for the limit theory presented in the paper covering both single and multiple bubble scenarios. A technical supplement to the paper (Phillips, Shi and Yu, 2011b)[6] provides a complete set of mathematical derivations of the limit theory presented here.

2   Identifying Bubbles: A New Test

A common issue that arises in unit root testing is the specification of the model used for estimation purposes, not least because of its impact on the appropriate asymptotic theory and the critical values that are used in testing. Related issues arise in right-tailed unit root tests of the type used in bubble detection. The impact of hypothesis formulation and model specification on right-tailed unit root tests has been studied recently in Phillips, Shi and Yu (2011a). Their analysis allowed for a null random walk process with an asymptotically negligible drift, namely

$$y_t = dT^{-\eta} + \theta y_{t-1} + \varepsilon_t, \ \varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2\right), \ \ \theta = 1 \tag{3}$$

where d is a constant, T is the sample size and $\eta > 1/2$, and their recommended empirical regression model for bubble detection follows (3) and therefore includes an intercept but no fitted time trend in the regression. Suppose a regression sample starts from the r_1^{th} fraction of the total sample and ends at the r_2^{th} fraction of the sample, where $r_2 = r_1 + r_w$ and r_w is the (fractional) window size of the regression. The empirical regression model is

$$\Delta y_t = \alpha_{r_1,r_2} + \beta_{r_1,r_2} y_{t-1} + \sum_{i=1}^{k} \psi^i_{r_1,r_2} \Delta y_{t-i} + \varepsilon_t, \tag{4}$$

where k is the lag order and $\varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2_{r_1,r_2}\right)$. The number of observations in the regression is $T_w = \lfloor Tr_w \rfloor$, where $\lfloor . \rfloor$ signifies the integer part of the argument. The ADF statistic (t-ratio) based on this regression is denoted by $ADF^{r_2}_{r_1}$.

---

[6]It is downloadable from `https://sites.google.com/site/shupingshi/TN_GSADFtest.pdf?attredirects=0&d=1`.

The SADF test estimates the ADF model repeatedly on a forward expanding sample sequence and conducts a hypothesis test based on the sup value of the corresponding ADF statistic sequence. The window size r_w expands from r_0 to 1, where r_0 is the smallest sample window (selected to ensure estimation efficiency) and 1 is the largest sample window (the total sample size). The starting point r_1 of the sample sequence is fixed at 0, so the ending point of each sample r_2 is equal to r_w, changing from r_0 to 1. The ADF statistic for a sample that runs from 0 to r_2 is denoted by $ADF_0^{r_2}$. The SADF statistic is defined as $\sup_{r_2 \in [r_0, 1]} ADF_0^{r_2}$, and is denoted by $SADF(r_0)$.

The SADF test and other right-sided unit root tests are not the only method of detecting explosive behavior. An alternative approach is the two-regime Markov-switching unit root test of Hall, Psaradakis and Sola (1999). While this procedure offers some appealing features like regime probability estimation, recent simulation work by Shi (2011) reveals that the Markov switching model is susceptible to false detection or spurious explosiveness. In addition, when allowance is made for a regime-dependent error variance as in Funke, Hall and Sola (1994) and van Norden and Vigfusson (1998), filtering algorithms can find it difficult to distinguish periods which may appear spuriously explosive due to high variance and periods when there is genuine explosive behavior. Furthermore, the bootstrapping procedure embedded in the Markov switching unit root test is computationally burdensome as Psaradakis, Sola and Spagnolo (2001) pointed out. These pitfalls make the Markov switching unit root test a difficult and somewhat unreliable tool of financial surveillance.

Other econometric approaches may be adapted to use the same recursive feature of the SADF test, such as the modified Bhargava statistic (Bhargava, 1986), the modified Busetti-Taylor statistic (Busetti and Taylor, 2004), and the modified Kim statistic (Kim, 2000). These tests are considered in Homm and Breitung (2011) for bubble detection and all share the spirit of the SADF test of PWY. That is, the statistic is calculated recursively and then the sup functional of the recursive statistics is calculated for testing. Since all these tests are similar in character to the SADF test and since Homm and Breitung (2011) found in their simulations that the PWY test was the most powerful in detecting multiple bubbles, we focus attention in
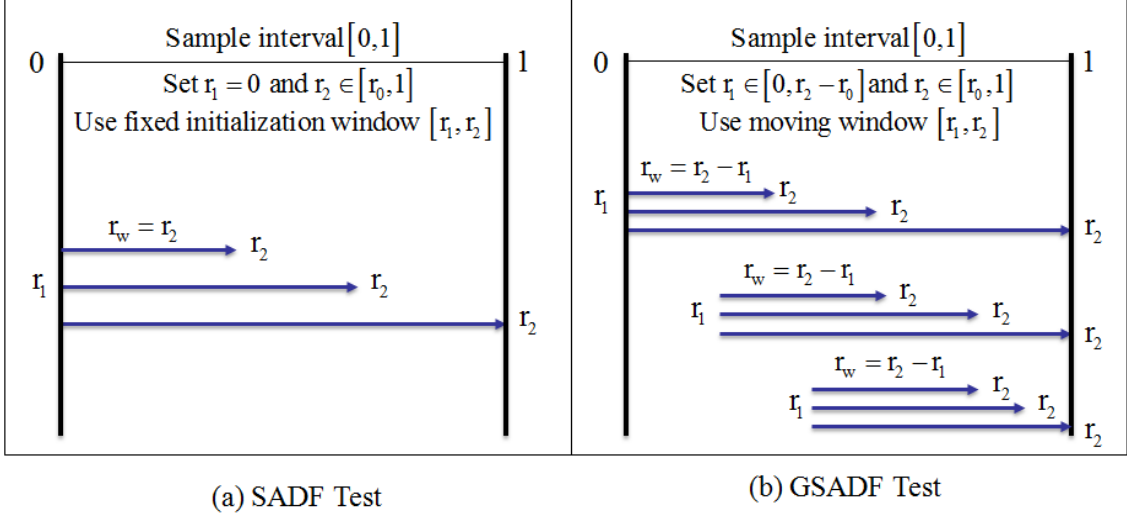
Figure 1: The sample sequences and window widths of the SADF test and the GSADF test

this paper on the SADF test. However, in our empirical application we provide comparative results with the CUSUM procedure in view of its good overall performance in the Homm and Breitung simulations.

The GSADF test deployed in the current paper continues the idea of repeatedly running the ADF test regression (4) on a sample sequence. However, the sample sequence is broader than that of the SADF test. Besides varying the end point of the regression r_2 from r_0 to 1, the GSADF test allows the starting points r_1 to change within a feasible range, which is from 0 to $r_2 - r_0$. Figure 1 illustrates the sample sequences of the SADF test and the GSADF test. We define the GSADF statistic to be the largest ADF statistic over the feasible ranges of r_1 and r_2, and we denote this statistic by $GSADF(r_0)$. That is,

$$GSADF(r_0) = \sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \left\{ ADF_{r_1}^{r_2} \right\}.$$

Proposition 1 *When the regression model includes an intercept and the null hypothesis is a random walk with an asymptotically negligible drift (i.e. $dT^{-\eta}$ with $\eta > 1/2$ and constant d),*

10

the limit distribution of the GSADF test statistic is:

$$\sup_{\substack{r_2 \in [r_0, 1] \\ r_1 \in [0, r_2 - r_0]}} \left\{ \frac{\frac{1}{2} r_w \left[W(r_2)^2 - W(r_1)^2 - r_w \right] - \int_{r_1}^{r_2} W(r) \, dr \left[W(r_2) - W(r_1) \right]}{r_w^{1/2} \left\{ r_w \int_{r_1}^{r_2} W(r)^2 \, dr - \left[\int_{r_1}^{r_2} W(r) \, dr \right]^2 \right\}^{1/2}} \right\}, \tag{5}$$

where $r_w = r_2 - r_1$ and W is a standard Wiener process.

The proof of Proposition 1 is similar to that of PWY and is given in the technical supplement to the paper (Phillips, Shi and Yu, 2011b). The limit distribution of the GSADF statistic is identical to the case when the regression model includes an intercept and the null hypothesis is a random walk without drift. The usual limit distribution of the ADF statistic is a special case of equation (5) with $r_1 = 0$ and $r_2 = r_w = 1$ while the limit distribution of the SADF statistic is a further special case of equation (5) with $r_1 = 0$ and $r_2 = r_w \in [r_0, 1]$ (see Phillips, Shi and Yu, 2011a).

Similar to the SADF statistic, the asymptotic GSADF distribution depends on the smallest window size r_0. In practice, r_0 needs to be chosen according to the total number of observations T. If T is small, r_0 needs to be large enough to ensure there are enough observations for adequate initial estimation. If T is large, r_0 can be set to be a smaller number so that the test does not miss any opportunity to detect an early explosive episode. In our empirical application we use $r_0 = 36/1680$, corresponding to around 2% of the data.

Critical values of the SADF and GSADF statistics are displayed in Table 1. The asymptotic critical values are obtained by numerical simulations, where the Wiener process is approximated by partial sums of $2,000$ independent $N(0,1)$ variates and the number of replications is $2,000$. The finite sample critical values are obtained from $5,000$ Monte Carlo replications. The lag order k is set to zero. The parameters, d and η, in the null hypothesis are set to unity.[7]

We observe the following phenomena. First, as the minimum window size r_0 decreases, critical values of the test statistic (including the SADF statistic and the GSADF statistic) increase. For instance, when r_0 decreases from 0.4 to 0.1, the 95% asymptotic critical value of

---

[7]From Phillips, Shi and Yu (2011a), we know that when $d = 1$ and $\eta > 1/2$, the finite sample distribution of the SADF statistic is almost invariant to the value of η.

Table 1: Critical values of the SADF and GSADF tests against an explosive alternative

(a) The asymptotic critical values

	$r_0 = 0.4$		$r_0 = 0.2$		$r_0 = 0.1$	
	SADF	GSADF	SADF	GSADF	SADF	GSADF
90%	0.86	1.25	1.04	1.66	1.18	1.89
95%	1.18	1.56	1.38	1.92	1.49	2.14
99%	1.79	2.18	1.91	2.44	2.01	2.57

(b) The finite sample critical values

	$T = 100$ and $r_0 = 0.4$		$T = 200$ and $r_0 = 0.4$		$T = 400$ and $r_0 = 0.4$	
	SADF	GSADF	SADF	GSADF	SADF	GSADF
90%	0.72	1.16	0.75	1.21	0.78	1.27
95%	1.05	1.48	1.08	1.52	1.10	1.55
99%	1.66	2.08	1.75	2.18	1.75	2.12

(c) The finite sample critical values

	$T = 100$ and $r_0 = 0.4$		$T = 200$ and $r_0 = 0.2$		$T = 400$ and $r_0 = 0.1$	
	SADF	GSADF	SADF	GSADF	SADF	GSADF
90%	0.72	1.16	0.97	1.64	1.19	1.97
95%	1.05	1.48	1.30	1.88	1.50	2.21
99%	1.66	2.08	1.86	2.46	1.98	2.71

Note: the asymptotic critical values are obtained by numerical simulations with 2,000 iterations. The Wiener process is approximated by partial sums of $N(0,1)$ with 2,000 steps. The finite sample critical values are obtained from the 5,000 Monte Carlo simulations. The parameters, d and η, are set to unity.

the GSADF statistic rises from 1.56 to 2.14 and the 95% finite sample critical value of the test statistic with sample size 400 increases from 1.48 to 2.21. Second, for a given r_0, the finite sample critical values of the test statistic are almost invariant. Third, critical values for the GSADF statistic are larger than those of the SADF statistic. As a case in point, when $T = 400$ and $r_0 = 0.1$, the 95% critical value of the GSADF statistic is 2.21 while that of the SADF statistic is 1.50. Figure 2 shows the asymptotic distribution of the ADF, $SADF(0.1)$ and $GSADF(0.1)$ statistics. The distributions move sequentially to the right and have greater concentration in the order ADF, $SADF(0.1)$ and $GSADF(0.1)$.
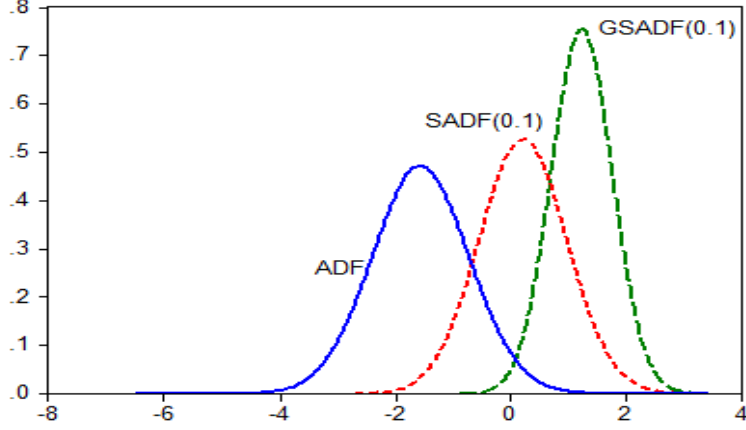
Figure 2: Asymptotic distributions of the ADF and supADF statistics ($r_0 = 0.1$)

3 Simulations

3.1 Generating the test sample

We first simulate an asset price series based on the Lucas asset pricing model and the Evans (1991) bubble model. The simulated asset prices consist of a market fundamental component P_t^f, which combines a random walk dividend process and equation (1) with $U_t = 0$ and $B_t = 0$ for all t to obtain[8]

$$D_t = \mu + D_{t-1} + \varepsilon_{Dt}, \quad \varepsilon_{Dt} \sim N\left(0, \sigma_D^2\right) \tag{6}$$

$$P_t^f = \frac{\mu\rho}{(1-\rho)^2} + \frac{\rho}{1-\rho}D_t, \tag{7}$$

and the Evans bubble component

$$B_{t+1} = \rho^{-1}B_t\varepsilon_{B,t+1}, \quad \text{if } B_t < b \tag{8}$$

---

[8]An alternative data generating process, which assumes that the logarithmic dividend is a random walk with drift, is as follows:

$$\ln D_t = \mu + \ln D_{t-1} + \varepsilon_t, \varepsilon_t \sim N\left(0, \sigma_d^2\right)$$

$$P_t^f = \frac{\rho \exp\left(\mu + \frac{1}{2}\sigma_d^2\right)}{1 - \rho \exp\left(\mu + \frac{1}{2}\sigma_d^2\right)}D_t.$$

13

$$B_{t+1} = \left[\zeta + (\pi\rho)^{-1} \theta_{t+1} (B_t - \rho\zeta) \right] \varepsilon_{B,t+1}, \quad \text{if } B_t \geq b . \tag{9}$$

This series has the submartingale property $\mathbb{E}_t (B_{t+1}) = (1 + r_f) B_t$. Parameter μ is the drift of the dividend process, σ_D^2 is the variance of the dividend, ρ is a discount factor with $\rho^{-1} = 1 + r_f > 1$ and $\varepsilon_{B,t} = \exp \left(y_t - \tau^2/2 \right)$ with $y_t \sim NID \left(0, \tau^2 \right)$. The quantity ζ is the re-initializing value after the bubble collapse. The series θ_t follows a Bernoulli process which takes the value 1 with probability π and 0 with probability $1 - \pi$. Equations (8) - (9) state that a bubble grows explosively at rate ρ^{-1} when its size is less than b while if the size is greater than b, the bubble grows at a faster rate $(\pi\rho)^{-1}$ but with a $1 - \pi$ probability of collapsing. The asset price is the sum of the market fundamental and the bubble component, namely $P_t = P_t^f + \kappa B_t$, where $\kappa > 0$ controls the relative magnitudes of these two components.

The parameter settings used by Evans (1991) are displayed in the top line of Table 2 and labeled *yearly*. The parameter values for μ and σ_D^2 were originally obtained by West (1988), by matching the sample mean and sample variance of first differenced real S&P 500 stock price index dividends from 1871 to 1980. The value for the discount factor ρ is equivalent to a 5% yearly interest rate.

Table 2: Parameter settings

	μ	σ_D^2	D_0	ρ	b	B_0	π	ζ	τ	κ
Yearly	0.0373	0.1574	1.3	0.952	1	0.50	0.85	0.50	0.05	20
Monthly	0.0024	0.0010	1.0	0.985	1	0.50	0.85	0.50	0.05	50

Due to the availability of higher frequency data, we apply the SADF test and the GSADF test to monthly data. The parameters μ and σ_D^2 are set to correspond to the sample mean and sample variance of the first differenced monthly real S&P 500 stock price index dividend described in the application section below, so that the settings are in accordance with our empirical application. The discount value ρ equals 0.985 (we allow ρ to vary from 0.975 to 0.999 in the size and power comparisons section). The new setting is labeled *monthly* in Table 2.

Figure 3 depicts one realization of the data generating process with the monthly parameter settings. As we observe in this graph, there are several obvious collapsing episodes of different magnitudes within this particular sample trajectory.
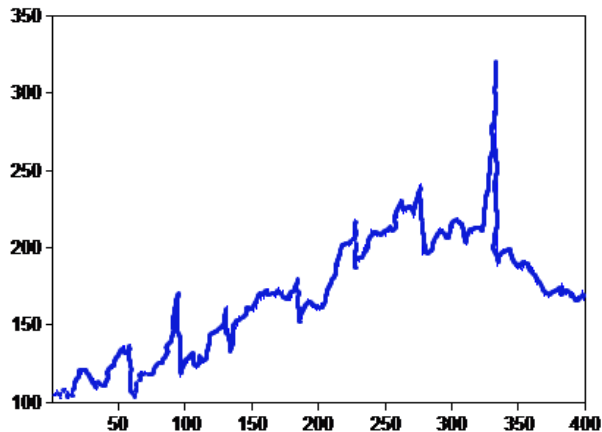
Figure 3: Simulated time series of $P_t = P_t^f + \kappa B_t$ using the Evans collapsing bubble model with sample size 400 and monthly parameter settings.

3.2   Implementation of the new test

We first implement the SADF test on the whole sample range. We repeat the test on a sub-sample which contains fewer collapsing episodes to illustrate the instability of the SADF test. Furthermore, we conduct the test on the same simulated data series (over the whole sample range) to show the advantage of the GSADF test.

The lag order k is set to zero for all tests in this paper.[9] The smallest window size considered in the SADF test for the whole sample contains 40 observations ($r_0 = 0.1$). The SADF statistic for the simulated data series is 0.71, which is smaller than the 10% finite sample critical value 1.19 (see Table 1). Therefore, we conclude that there are no bubbles in this sample. Now suppose that the SADF test starts from the 201st observation, and the smallest regression window also contains 40 observations ($r_0 = 0.2$). The SADF statistic obtained from this sample is 1.39 and it is greater than 1.30 (Table 1). In this case, we reject the null hypothesis of no bubble at the 5% significance level.

---

[9]In PWY, the lag order was determined by significance testing, as in Campbell and Perron (1991). However, we demonstrate in the size and power comparison section below that this lag selection criteria results in significant size distortion and reduces the power of both the SADF and GSADF tests.

15

Evidently the SADF test fails to find bubbles when the whole sample is utilized, whereas by re-selecting the starting point of the sample to exclude some of the collapse episodes, it succeeds in finding evidence of bubbles. Each of the above experiments can be viewed as special cases of the GSADF test in which the sample starting points are fixed. In the first experiment, the sample starting point of the GSADF test r_1 is set to 0. The sample starting point r_1 of the second experiment is fixed at 0.502. The conflicting results obtained from these two experiments demonstrates the importance of using variable starting points, as is done in the GSADF test.

We then apply the GSADF test to the simulated asset prices. The GSADF statistic of the simulated data is 8.59, which is substantially greater than the 1% finite sample critical value 2.71 (Table 1). Thus, the GSADF test finds strong evidence of bubbles. Compared to the SADF test, the GSADF identifies bubbles without re-selecting the sample starting point, giving an obvious improvement that is particularly useful in empirical applications.[10]

3.3 Size and Power Comparisons

This section compares the sizes and powers of the SADF and GSADF tests. The data generating process for the size comparison is the null hypothesis in equation (3) with $d = \eta = 1$. We calculate size based on the asymptotic critical values displayed in Table 1. The nominal size is 5%. The number of replications is 5,000. We observe from Table 3 that the size distortion of the GSADF test is smaller than that of the SADF test. For example, when $T = 400$ and $r_0 = 0.1$, the size distortion of the GSADF test is 0.9% whereas that of the SADF test is 1.6%.[11]

Powers in Table 4 and 5 are calculated with the 95% quantiles of the finite sample distributions (Table 1), and the number of iterations for the calculation is 5,000. The smallest window size for both the SADF test and the GSADF test has 40 observations. The data generating process of the power comparison is the periodically collapsing explosive process, equations (6)

---

[10]We observe similar phenomena from the alternative data generating process where the logarithmic dividend is a random walk with drift. Parameters in the alternative data generating process (monthly) are set as follows: $B_0 = 0.5, b = 1, \pi = 0.85, \zeta = 0.5, \rho = 0.985, \tau = 0.05, \mu = 0.001, \ln D_0 = 1, \sigma^2_{\ln D} = 0.0001$, and $P_t = P_t^f + 500B_t$.

[11]Suppose the lag order is determined by significance testing as in Campbell and Perron (1991) with a maximum lag order of 12. When $T = 400$ and $r_0 = 0.1$, the sizes of the SADF test and the GSADF test are 0.130 and 0.790 (the nominal size is 5%), indicating size distortion in both tests and a particularly large size distortion for the GSADF test.

Table 3: Sizes of the SADF and GSADF tests with asymptotic critical values. The data generating process is equation (3) with $d = \eta = 1$. The nominal size is 5%.

| | $T = 100$ | $T = 200$ | | $T = 400$ | |
	$r_0 = 0.4$	$r_0 = 0.4$	$r_0 = 0.2$	$r_0 = 0.4$	$r_0 = 0.1$
SADF	0.043	0.040	0.038	0.041	0.034
GSADF	0.048	0.041	0.044	0.045	0.059

Note: size calculations are based on 5000 replications..

- (9). For comparison with the literature, we first set the parameters in the DGP as in Evans (1991) with sample sizes of 100 and 200. From the left panel of Table 4 (labeled *yearly*), the power of the GSADF test is 7% and 15.2% higher than those of the SADF test when the sample size is 100 and 200.[12]

Table 4: Powers of the SADF and GSADF tests. The data generating process is equation (6)-(9).

| | Yearly | | Monthly | |
	SADF	GSADF	SADF	GSADF
$T = 100$ and $r_0 = 0.4$	0.408	0.478	0.509	0.556
$T = 200$ and $r_0 = 0.2$	0.634	0.786	0.699	0.833
$T = 400$ and $r_0 = 0.1$	-	-	0.832	0.977

Note: power calculations are based on 5000 replications.

Table 4 also displays powers of the SADF and GSADF tests under the DGP with monthly parameter settings and sample sizes 100, 200 and 400. From the right panel of the table, when the sample size $T = 400$, the GSADF test raises test power from 83.2% to 97.7%, giving a 14.5% improvement. The power improvement of the GSADF test is 4.7% when $T = 100$ and 13.4% when $T = 200$. For any given bubble collapsing probability π in the Evans model, the sample period is more likely to include multiple collapsing episodes the larger the sample size. Hence, the advantages of the GSADF test are more evident for large T.

In Table 5, we compare powers of the SADF and GSADF tests with the discount factor ρ varying from 0.975 to 0.990, under the DGP with the monthly parameter settings. First, due

[12]Suppose the lag order is determined by significance testing as in Campbell and Perron (1991) with a maximum lag order of 12. When $T = 200$ and $r_0 = 0.2$, the powers of the SADF test and the GSADF test are 0.565 and 0.661, which are smaller than those in Table 4.

to the fact that the rate of bubble expansion is inversely related to the discount factor, powers of both SADF test and GSADF tests are expected to decrease as ρ increases. The power of the SADF (GSADF) test declines from 84.5% to 76.9% (from 99.3% to 91.0%) as the discount factor rises from 0.975 to 0.990 (see Table 5). Second, we observe from Table 5 that the GSADF test has greater discriminatory power for detecting bubbles than the SADF test. The power improvement is 14.8%, 14.8%, 14.5% and 14.1% for $\rho = \{0.975, 0.980, 0.985, 0.990\}$.

Table 5: Powers of the SADF and GSADF tests. The data generating process is equations (6)-(9) with the *monthly* parameter settings and sample size 400 ($r_0 = 0.1$).

ρ	0.975	0.980	0.985	0.990	
SADF	0.845	0.840	0.832	0.769	
GSADF	0.993	0.988	0.977	0.910	

Note: power calculations are based on 5000 replications.

4   Date-stamping Strategies

Suppose that one is interested in knowing whether any particular observation, such as the point $\lfloor Tr_2 \rfloor$, belongs to a bubble phase in the trajectory. PWY suggest conducting a right-tailed ADF test recursively using information up to this observation (i.e. $I_{\lfloor Tr_2 \rfloor} = \{y_1, y_2, \cdots, y_{\lfloor Tr_2 \rfloor}\}$). Since it is possible that $I_{\lfloor Tr_2 \rfloor}$ includes one or more collapsing bubble episodes, the ADF test, like the cointegration-based test for bubbles, may result in finding *pseudo stationary* behavior. We therefore recommend performing a *backward sup ADF test* on $I_{\lfloor Tr_2 \rfloor}$ to improve identification accuracy.

The backward SADF test performs a sup ADF test on a backward expanding sample sequence, where the ending points of the samples are fixed at r_2 and the starting point varies from 0 to $r_2 - r_0$. Suppose we label the ADF statistic for each regression using its starting point r_1 and ending point r_2 to obtain $BADF_{r_1}^{r_2}$. The corresponding ADF statistic sequence is $\left\{BADF_{r_1}^{r_2}\right\}_{r_1 \in [0, r_2 - r_0]}$. The backward SADF statistic is defined as the sup value of the ADF statistic sequence, denoted by

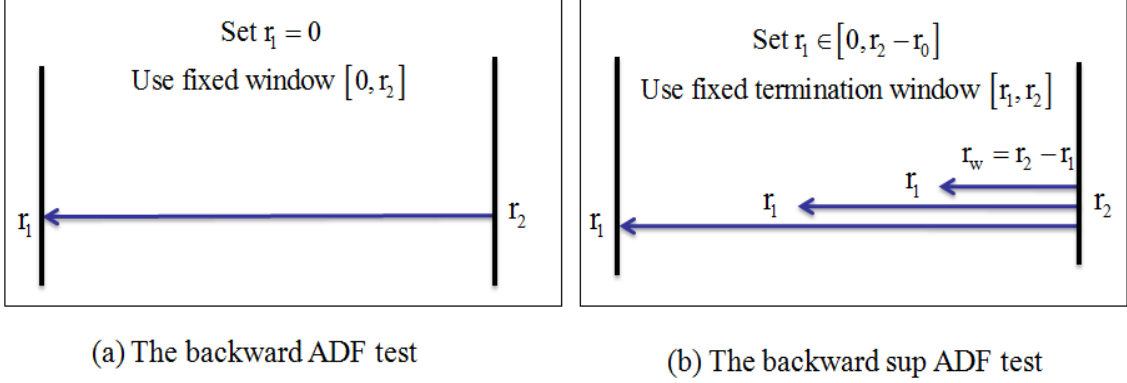$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \left\{BADF_{r_1}^{r_2}\right\}.$$

Figure 4: The sample sequences of the backward ADF test and the backward SADF test

The backward ADF test is a special case of the backward sup ADF test with $r_1 = 0$. We denote the backward ADF statistic by $BADF_{r_2}$. Figure 4 illustrates the difference between the backward ADF test and the backward SADF test. PWY propose comparing $BADF_{r_2}$ with the (right-tail) critical values of the standard ADF statistic to identify explosiveness at observation $\lfloor Tr_2 \rfloor$. The feasible range of r_2 runs from r_0 to 1. The origination date of a bubble $\lfloor Tr_e \rfloor$ is calculated as the first chronological observation whose backward ADF statistic exceeds the critical value. We denote the calculated origination date by $\lfloor T\hat{r}_e \rfloor$. The estimated termination date of a bubble $\lfloor T\hat{r}_f \rfloor$ is the first chronological observation after $\lfloor T\hat{r}_e \rfloor + \log(T)$ whose backward ADF statistic goes below the critical value. PWY impose a condition that for a bubble to exist its duration must exceed $\log(T)$. This requirement helps to exclude short lived blips in the fitted autoregressive coefficient. The dating estimates are then

$$\hat{r}_e = \inf_{r_2 \in [r_0, 1]} \left\{ r_2 : BADF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_f = \inf_{r_2 \in [\hat{r}_e + \log(T)/T, 1]} \left\{ r_2 : BADF_{r_2} < cv_{r_2}^{\beta_T} \right\}, \quad (10)$$

where $cv_{r_2}^{\beta_T}$ is the $100\beta_T\%$ critical value of the backward ADF statistic based on $\lfloor Tr_2 \rfloor$ observations. The significance level β_T depends on the sample size T and we assume that $\beta_T \to 0$ as $T \to \infty$.

Instead of using the backward ADF statistic, the new strategy suggests making inferences on

19

the explosiveness of observation $\lfloor Tr_2 \rfloor$ based on the backward sup ADF statistic, $BSADF_{r_2}(r_0)$. We define the origination date of a bubble as the first observation whose backward sup ADF statistic exceeds the critical value of the backward sup ADF statistic. The termination date of a bubble is calculated as the first observation after $\lfloor T\hat{r}_e \rfloor + \delta \log(T)$ whose backward sup ADF statistic falls below the critical value of the backward sup ADF statistic. Here it is assumed that the duration of the bubble exceeds $\delta \log(T)$, where δ is a frequency dependent parameter.[13] The (fractional) origination and termination points of a bubble (i.e. r_e and r_f) are calculated according to the following first crossing time equations:

$$\hat{r}_e = \inf_{r_2 \in [r_0, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\}, \tag{11}$$

$$\hat{r}_f = \inf_{r_2 \in [\hat{r}_e + \delta \log(T)/T, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}, \tag{12}$$

where $scv_{r_2}^{\beta_T}$ is the $100\beta_T\%$ critical value of the sup ADF statistic based on $\lfloor Tr_2 \rfloor$ observations. Analogously, the significance level β_T depends on the sample size T and it goes to zero as the sample size approaches infinity.

In addition, the SADF test can be viewed as a repeated implementation of the backward ADF test for each $r_2 \in [r_0, 1]$. The GSADF test is equivalent to a test which implements the backward sup ADF test repeatedly for each $r_2 \in [r_0, 1]$ and makes inferences based on the sup value of the backward sup ADF statistic sequence, $\{BSADF_{r_2}(r_0)\}_{r_2 \in [r_0, 1]}$. Hence, the SADF and GSADF statistics can respectively be rewritten as

$$SADF(r_0) = \sup_{r_2 \in [r_0, 1]} \{BADF_{r_2}\},$$

$$GSADF(r_0) = \sup_{r_2 \in [r_0, 1]} \{BSADF_{r_2}(r_0)\}.$$

Thus, the PWY date-stamping strategy corresponds to the SADF test and the new strategy corresponds to the GSADF test. The essential features of the two tests are shown in stylized form in the diagrams of Figure 5.

---

[13] For instance, one might wish to impose a minimal condition that to be classified as a bubble its duration should exceed a certain period such as one year (which is inevitably arbitrary). Then, when the sample size is 30 years (360 months), δ is 0.7 for yearly data and 5 for monthly data.

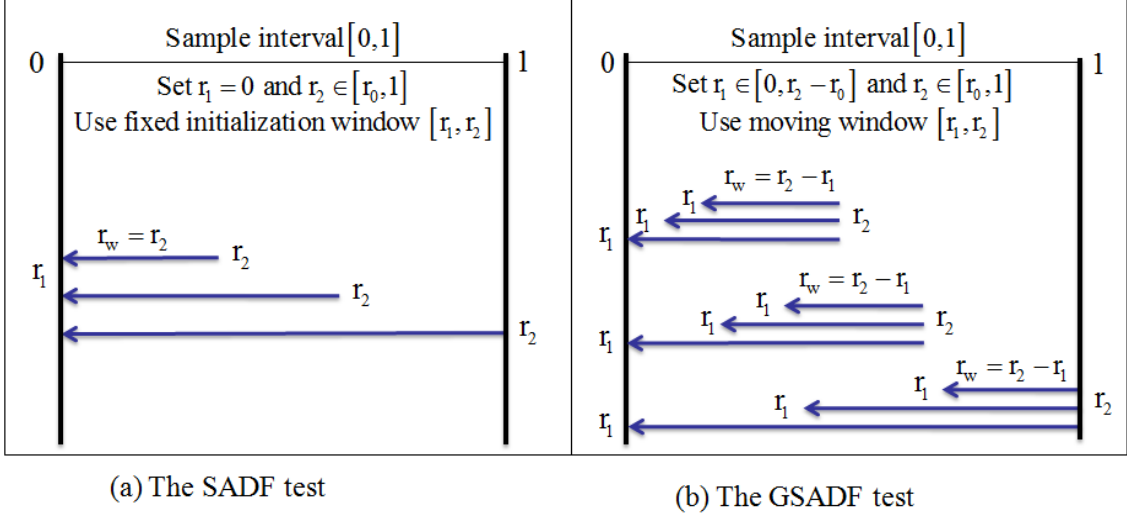(a) The SADF test (b) The GSADF test

Figure 5: An alternative illustration of the sample sequences and window widths of the SADF test and the GSADF test

4.1 The null hypothesis: no bubbles

In order to derive the consistency properties of these date-stamping strategies, we first need to obtain the asymptotic distributions of the ADF statistic and the SADF statistic with $\lfloor Tr_2 \rfloor$ observations under the null hypothesis (3). We know that the backward ADF test with observation $\lfloor Tr_2 \rfloor$ is a special case of the GSADF test with $r_1 = 0$ and a fixed r_2 and the backward sup ADF test is a special case of the GSADF test with a fixed r_2 and $r_1 = r_2 - r_w$. Therefore, based on equation (5), we can derive the asymptotic distributions of these two statistics, namely

$$F_{r_2}(W) := \frac{\frac{1}{2}r_2 \left[W(r_2)^2 - r_2 \right] - \int_0^{r_2} W(r)\, dr W(r_2)}{r_2^{1/2} \left\{ r_2 \int_0^{r_2} W(r)^2\, dr - \left[\int_0^{r_2} W(r)\, dr \right]^2 \right\}^{1/2}},$$

$$F_{r_2}^{r_0}(W) := \sup_{\substack{r_1 \in [0, r_2 - r_0] \\ r_w = r_2 - r_1}} \left\{ \frac{\frac{1}{2}r_w \left[W(r_2)^2 - W(r_1)^2 - r_w \right] - \int_{r_1}^{r_2} W(r)\, dr \left[W(r_2) - W(r_1) \right]}{r_w^{1/2} \left\{ r_w \int_{r_1}^{r_2} W(r)^2\, dr - \left[\int_{r_1}^{r_2} W(r)\, dr \right]^2 \right\}^{1/2}} \right\}.$$

We, therefore, define $cv_{r_2}^{\beta_T}$ as the $100(1 - \beta_T)\%$ quantile of $F_{r_2}(W)$ and $scv_{r_2}^{\beta_T}$ as the $100(1 - \beta_T)\%$ quantile of $F_{r_2}^{r_0}(W)$. We know that $cv_{r_2}^{\beta_T} \to \infty$ and $scv_{r_2}^{\beta_T} \to \infty$ as $\beta_T \to 0$.

21

Given $cv_{r_2}^{\beta_T} \to \infty$ and $scv_{r_2}^{\beta_T} \to \infty$, under the null hypothesis of no bubbles, the probabilities of (falsely) detecting the origination of bubble expansion and the termination of bubble collapse using the backward ADF statistic and the backward sup ADF statistic tend to zero, so that both $\Pr\{\hat{r}_e \in [r_0, 1]\} \to 0$ and $\Pr\{\hat{r}_f \in [r_0, 1]\} \to 0$.

4.2 The alternative hypothesis: a single bubble

Consider the data generating process of Phillips and Yu (2009)

$$X_t = X_{t-1}1\{t < \tau_e\} + \delta_T X_{t-1}1\{\tau_e \le t \le \tau_f\}$$
$$+ \left(\sum_{k=\tau_f+1}^{t} \varepsilon_k + X_{\tau_f}^*\right)1\{t > \tau_f\} + \varepsilon_t 1\{j \le \tau_f\}, \tag{13}$$

where $\delta_T = 1 + cT^{-\alpha}$ with $c > 0$ and $\alpha \in (0,1)$, $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$, $X_{\tau_f}^* = X_{\tau_e} + X^*$ with $X^* = O_p(1)$, $\tau_e = \lfloor Tr_e \rfloor$ is the origination of bubble expansion and $\tau_f = \lfloor Tr_f \rfloor$ is the termination of bubble collapse. The pre-bubble period $N_0 = [1, \tau_e)$ is assumed to be a pure random walk process. The bubble expansion period $B = [\tau_e, \tau_f]$ is a mildly explosive process with expansion rate δ_T. The process then collapses to $X_{\tau_f}^*$, which equals X_{τ_e} plus a small perturbation, and continues its pure random walk path in the period $N_1 = (\tau_f, \tau]$.

Notice that there is only one bubble episode in the data generating process (13). Under this mechanism we have the following consistency results, whose proofs are collected in Appendix A.

Theorem 1 *Suppose $\hat{r}_e$ and $\hat{r}_f$ are obtained from the backward DF test based on the t statistic, (10). Given an alternative hypothesis of mildly explosive behavior in model (13), if*

$$\frac{1}{cv_{r_2}^{\beta_T}} + \frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0, \tag{14}$$

we have $\hat{r}_e \overset{p}{\to} r_e$ and $\hat{r}_f \overset{p}{\to} r_f$ as $T \to \infty$.

Theorem 2 *Suppose $\hat{r}_e$ and $\hat{r}_f$ are obtained from the backward sup DF test based on the t statistic, (11) - (12). Given an alternative hypothesis of mildly explosive behavior in model (13), if*

$$\frac{1}{scv_{r_2}^{\beta_T}} + \frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0, \tag{15}$$

we have $\hat{r}_e \xrightarrow{p} r_e$ and $\hat{r}_f \xrightarrow{p} r_f$ as $T \to \infty$.

These results show that both strategies consistently estimate the origination and termination points when there is only a single bubble episode in the sample period. Suppose $cv_{r_2}^{\beta_T} = O_p(T^\gamma)$ and $scv_{r_2}^{\beta_T} = O_p(T^{\gamma_s})$. The regularity condition (14) in Theorem 1 implies that the order of magnitude (γ) of $cv_{r_2}^{\beta_T}$ needs to be greater than 0 and smaller than $1/2$, namely $\gamma \in (0, 1/2)$. Theorem 2 requires the order of magnitude (γ_s) of $scv_{r_2}^{\beta_T}$ to be greater than 0 and smaller than $1/2$ to obtain the consistency of $\hat{r}_e$ and $\hat{r}_f$.

4.3 The alternative hypothesis: two bubbles

Consider a data generating process with two bubble episodes:

$$X_t = X_{t-1}1\{t \in N_0\} + \delta_T X_{t-1}1\{t \in B_1 \cup B_2\} + \left(\sum_{k=\tau_{1f}+1}^{t} \varepsilon_k + X_{\tau_{1f}}^*\right)1\{t \in N_1\}$$

$$+ \left(\sum_{l=\tau_{2f}+1}^{t} \varepsilon_l + X_{\tau_{2f}}^*\right)1\{t \in N_2\} + \varepsilon_t 1\{j \in N_0 \cup B_1 \cup B_2\}, \tag{16}$$

where $N_0 = [1, \tau_{1e}), B_1 = [\tau_{1e}, \tau_{1f}], N_1 = (\tau_{1f}, \tau_{2e}), B_2 = [\tau_{2e}, \tau_{2f}]$ and $N_2 = (\tau_{2f}, \tau]$. $\tau_{1e} = \lfloor Tr_{1e} \rfloor, \tau_{1f} = \lfloor Tr_{1f} \rfloor$ are the origination and termination dates of the first bubble, $\tau_{2e} = \lfloor Tr_{2e} \rfloor$, $\tau_{2f} = \lfloor Tr_{2f} \rfloor$ are the origination and termination dates of the second bubble and τ is the last observation of the sample. After the collapse of the first bubble, X_t continues its pure random walk path until $\tau_{2e} - 1$ and starts another expansion process at τ_{2e}. The expansion process lasts until τ_{2f} and collapses to a value of $X_{\tau_{2f}}^*$. It then continues its pure random walk path until the end of the sample period τ. We assume that the expansion duration of the first bubble is longer than that of the second bubble, namely $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$.

The date-stamping strategy of PWY suggests calculating r_{1e}, r_{1f}, r_{2e} and r_{2f} from the following equations (based on the ADF statistic):

$$\hat{r}_{1e} = \inf_{r_2 \in [r_0, 1]} \left\{r_2 : BADF_{r_2} > cv_{r_2}^{\beta_T}\right\} \text{ and } \hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e}+\log(T)/T, 1]} \left\{r_2 : BADF_{r_2} < cv_{r_2}^{\beta_T}\right\},$$

$$\tag{17}$$

23

$$\hat{r}_{2e} = \inf_{r_2 \in [\hat{r}_{1f},1]} \left\{ r_2 : BADF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e}+\log(T)/T,1]} \left\{ r_2 : BADF_{r_2} < cv_{r_2}^{\beta_T} \right\},$$

(18)

where the duration of the bubble periods is restricted to be longer than $\log(T)$.

The new strategy recommends using the backward sup ADF test and calculating the origination and termination points according to the following equations:

$$\hat{r}_{1e} = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\},$$

(19)

$$\hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e}+\delta\log(T)/T,1]} \left\{ r_2 : BSADF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\},$$

(20)

$$\hat{r}_{2e} = \inf_{r_2 \in [\hat{r}_{1f},1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\},$$

(21)

$$\hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e}+\delta\log(T)/T,1]} \left\{ r_2 : BSADF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}.$$

(22)

An alternative implementation of the PWY procedure is to use that procedure sequentially, namely detect one bubble at a time. The dating criteria for the first bubble remains the same (i.e. equation (17)). Conditional on the first bubble having been found and terminated at $\hat{r}_{1f}$, the following dating criteria is used for a second bubble:

$$\hat{r}_{2e} = \inf_{r_2 \in (\hat{r}_{1f}+\varepsilon_T,1]} \left\{ r_2 :_{\hat{r}_{1f}} BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e}+\log(T)/T,1]} \left\{ r_2 :_{\hat{r}_{1f}} BDF_{r_2} < cv_{r_2}^{\beta_T} \right\},$$

(23)

where $_{\hat{r}_{1f}} BDF_{r_2}$ is the ADF statistic calculated over $(\hat{r}_{1f}, r_2]$. Note that we need a few observations to initialize the procedure (i.e. $r_2 \in (\hat{r}_{1f} + \varepsilon_T, 1]$ for some $\varepsilon_T > 0$).[14]

We have the following asymptotic results for these dating estimates. Proofs of the theorems are given in Appendix B.

Theorem 3 *Suppose $\hat{r}_{1e}$, $\hat{r}_{1f}$, $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are obtained from the backward DF test based on the t statistic, (17) - (18). Given an alternative hypothesis of mildly explosive behavior of model (16) with $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$, if*

$$\frac{1}{cv_{r_2}^{\beta_T}} + \frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0,$$

_____

[14] For example, $\varepsilon_T = \log T/T$ or $T^{-\delta}$ with some $\delta \in (0,1)$.

we have $\hat{r}_{1e} \xrightarrow{p} r_{1e}$ and $\hat{r}_{1f} \xrightarrow{p} r_{1f}$ as $T \to \infty$; and $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are not consistent estimators of r_{2e} and r_{2f}.

Theorem 4 *Suppose $\hat{r}_{1e}$, $\hat{r}_{1f}$, $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are obtained from the backward sup DF test based on the t statistic, (19) - (22). Given an alternative hypothesis of mildly explosive behavior of model (16) with $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$, if*

$$\frac{1}{scv_{r_2}^{\beta_T}} + \frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0,$$

we have $\hat{r}_{1e} \xrightarrow{p} r_{1e}$, $\hat{r}_{1f} \xrightarrow{p} r_{1f}$, $\hat{r}_{2e} \xrightarrow{p} r_{2e}$ and $\hat{r}_{2f} \xrightarrow{p} r_{2f}$ as $T \to \infty$.

Theorem 5 *Suppose $\hat{r}_{1e}$, $\hat{r}_{1f}$, $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are obtained from the backward DF test based on the t statistic, (17) and (23). Given an alternative hypothesis of mildly explosive behavior of model (16) with $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$, if*

$$\frac{1}{scv_{r_2}^{\beta_T}} + \frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0,$$

we have $\hat{r}_{1e} \xrightarrow{p} r_{1e}$, $\hat{r}_{1f} \xrightarrow{p} r_{1f}$, $\hat{r}_{2e} \xrightarrow{p} r_{2e}$ and $\hat{r}_{2f} \xrightarrow{p} r_{2f}$ as $T \to \infty$.

A restatement of Theorem 3 is useful. Suppose the sample period includes two bubble episodes and the duration of the first bubble is longer than the second. The strategy of PWY (corresponding to the SADF test) can consistently estimate the origination and termination of the first bubble but does not consistently estimate those of the second bubble. In contrast, Theorem 4 and Theorem 5 say that the new date-stamping strategy (corresponding to the GSADF test) and the alternative implementation of the PWY strategy can calculate the origination and termination of both bubbles consistently in this scenario.

We also analyze the consistency properties of these two date-stamping strategies when there are two bubbles and the duration of the first bubble is shorter than the second bubble. Under this circumstance, the strategy of PWY cannot estimate the origination date of the second bubble consistently[15], whereas the new strategy can consistently estimate the origination and termination dates of the two bubbles.[16]

---

[15] It can consistently estimate the origination date of the first bubble and the termination dates of both bubbles.

[16] The proof is similar to the proofs of Theorems 3, 4 and 5 (Appendix B) and is omitted for brevity.

Theorem 3, 4 and 5 can be extended to a multiple bubbles scenario. Suppose there are N bubbles ($N > 2$). If the duration of the ith bubble is longer than that of the jth bubble, where $i, j \in \{1, 2, \cdots, N\}$ and $i < j$, then, the PWY strategy can consistently estimate the origination and termination dates of the ith bubble but not those associated with the jth bubble. In contrast, the new strategy and the alternative implementation of the PWY strategy can estimate dates associated with both bubbles consistently.

5   Empirical Application

We consider a long historical time series in which many crisis events are known to have occurred. The data comprise the real S&P 500 stock price index and the real S&P 500 stock price index dividend, both obtained from Robert Shiller's website. The data are sampled monthly over the period from January 1871 to December 2010, constituting 1,680 observations and are plotted in Figure 6 by the solid (blue) line.

We first apply the SADF test and the GSADF test to the price-dividend ratio. Table 6 presents critical values for these two tests and these were obtained from Monte Carlo simulation with $2,000$ replications (sample size $1,680$). In performing the ADF regressions and calculating critical values, the smallest window comprised 36 observations. From Table 6, the SADF and GSADF statistics for the full data series are 3.30 and 4.21. Both exceed their respective 1% right-tail critical values (i.e. $3.30 > 2.17$ and $4.21 > 3.31$), giving strong evidence that the S&P 500 price-dividend ratio had explosive subperiods. We conclude from both tests that there is evidence of bubbles in the S&P 500 stock market data.

Table 6: The SADF test and the GSADF test of the S&P 500 stock market

	SADF	GSADF
S&P500 Price-Dividend Ratio	3.30	4.21
Finite sample critical values		
90%	1.45	2.55
95%	1.70	2.80
99%	2.17	3.31

Note: Critical values of both tests are obtained from Monte Carlo simulation with $2,000$ replications (sample size 1,680). The smallest window has 36 observations.
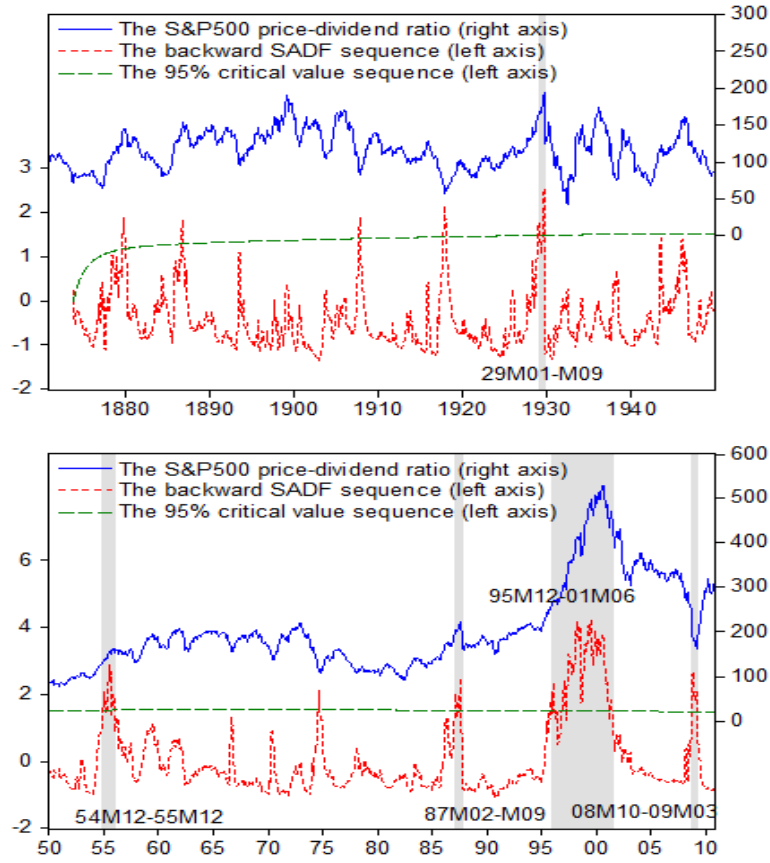
Figure 6: Date-stamping bubble periods in the S&P 500 price-dividend ratio: the GSADF test

To locate specific bubble periods, we compare the backward SADF statistic sequence with the 95% SADF critical value sequence, which were obtained from Monte Carlo simulation with 2,000 replications. The top panel of Figure 6 displays results for the date-stamping strategy over the period from January 1871 to December 1949 and the bottom panel displays results over the rest of the sample period. The identified exuberance and collapse periods include *the great crash episode* (1929M01-M09), *the postwar boom in 1954* (1954M12-1955M12), *black Monday in October 1987* (1987M02-M09), *the dot-com bubble* (1995M12-2001M06) and *the subprime mortgage crisis* (2008M10-2009M03). The durations of those episodes are greater than or equal to half a year.[17] This strategy also identifies several episodes of crisis or explosiveness whose durations are shorter than half a year, for instance, the explosive recovery phase following *the panic of 1873* (1879M10-1880M02), *the banking panic of 1907* (1907M10-M11) and *the 1974 stock market crash* (1974M09).

Notice that the new date-stamping strategy not only locates the explosive expansion periods but also identifies collapse episodes. Market collapses have occurred in the past when bubbles in other markets crashed and the collapse spread to the S&P 500 as, for instance, in the *dot-com bubble collapse* and *the subprime mortgage crisis.*

Figure 7 plots the ADF statistic sequence against the 95% ADF critical value sequence. We can see that the strategy of PWY (based on the SADF test) identifies only two explosive periods – the recovery phase of *the panic of 1873* (1879M10-1880M04) and *the dot-com bubble* (1997M07-2001M08).[18]

The PWY strategy is called FLUC monitoring in the Homm and Breitung (2011) simulation study, with some variations in the implementation procedure. Their simulation findings show that FLUC monitoring has higher power than other procedures in detecting periodically collapsing bubbles of the Evans (1991) type, the closest rival being the CUSUM procedure. For comparison, we therefore applied the CUSUM monitoring procedure to the S&P 500 price-

---

[17]If bubble duration is restricted to be longer than three months, the strategy identifies one more bubble episode, namely the explosive recovery phase following *the panic of 1873* (1879M10-1880M02).

[18]If we restrict the duration of a bubble to be longer than twelve months, the new dating strategy identifies two bubble episodes: *the postwar boom in 1954* (1954M12-1955M12) and *the dot-com bubble* (1995M12-2001M06). In that case the strategy of PWY only identifies *the dot-com bubble* (1997M07-2001M08).
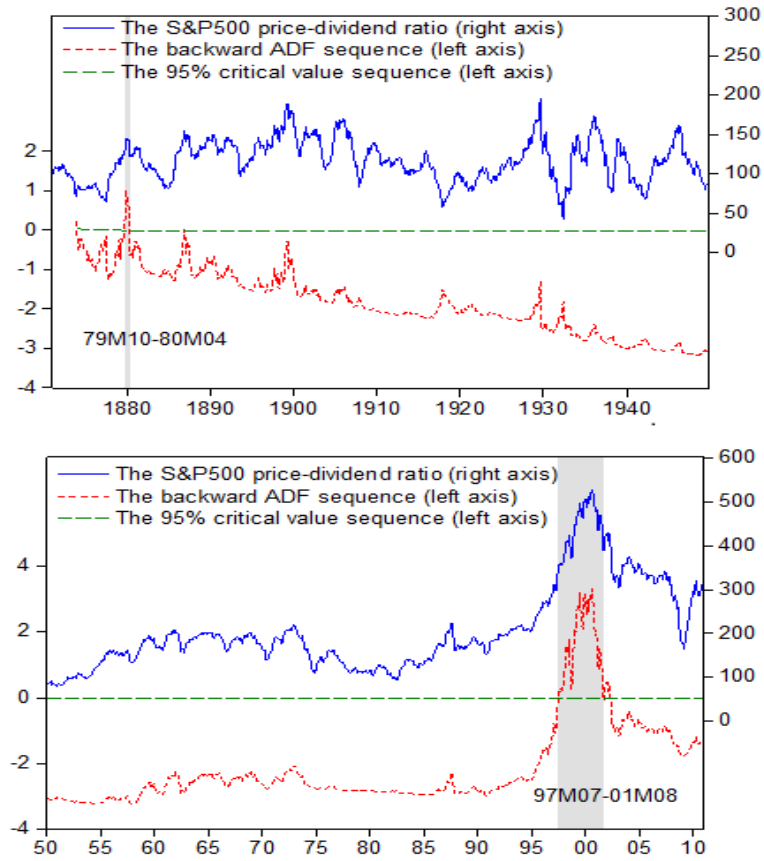
Figure 7: Date-stamping bubble periods in the S&P 500 price-dividend ratio: the SADF test.

dividend ratio (ex post). The CUSUM detector is denoted by $C_{r_0}^r$ and defined as

$$C_{r_0}^r = \frac{1}{\hat{\sigma}_r} \sum_{j=\lfloor Tr_0 \rfloor + 1}^{\lfloor Tr \rfloor} \Delta y_j \text{ with } \hat{\sigma}_r^2 = (\lfloor Tr \rfloor - 1)^{-1} \sum_{j=1}^{\lfloor Tr \rfloor} (\Delta y_j - \hat{\mu}_r)^2,$$

where $\lfloor Tr_0 \rfloor$ is the training sample,[19] $\lfloor Tr \rfloor$ is the monitoring observation, $\hat{\mu}_r$ is the mean of $\{\Delta y_1, ..., \Delta y_{\lfloor Tr \rfloor}\}$, and $r > r_0$. Under the null hypothesis of a pure random walk, it has the following asymptotic property (see Chu, Stinchcombe and White (1996))

$$\lim_{T \to \infty} P\left\{ C_{r_0}^r > c_r \sqrt{\lfloor Tr \rfloor} \text{ for some } r \in (r_0, 1] \right\} \leq \frac{1}{2} \exp\left(-\kappa_\alpha/2\right),$$

where $c_r = \sqrt{\kappa_\alpha + \log\left(r/r_0\right)}.$[20]

The CUSUM detector is applied to detrended data (i.e. to the residuals from the regression of y_t on a constant and a linear time trend). To be consistent with the SADF and GSADF dating strategies, we choose a training sample of 36 months. Figure 8 plots the CUSUM detector sequence against the 95% critical value sequence. The critical value sequence is obtained from Monte Carlo simulation (through application of the CUSUM detector to data simulated from a pure random walk) with 2,000 replications.

As it is evident in Figure 8, the CUSUM test identifies four bubble episodes for periods before 1900. For the post-1900 sample, the procedure detects only the great crash and the dot-com bubble episodes. It does not provide any warning alert or acknowledgment of black Monday in October 1987 and the subprime mortgage crisis in 2008, among other episodes identified by the GSADF dating strategy. So CUSUM monitoring may be regarded as a relatively conservative surveillance device.[21]

---

[19] It is assumed that there is no structural break in the training sample.

[20] When the significance level $\alpha = 0.05$, for instance, $\kappa_{0.05}$ equals 4.6.

[21] The conservative nature of the test arises from the fact the residual variance estimate $\hat{\sigma}_r$ (based on the data $\{y_1, ..., y_{\lfloor Tr \rfloor}\}$) can be quite large when the sample includes periodically collapsing bubble episodes, which may have less impact on the numerator due to collapses, thereby reducing the size of the CUSUM detector.
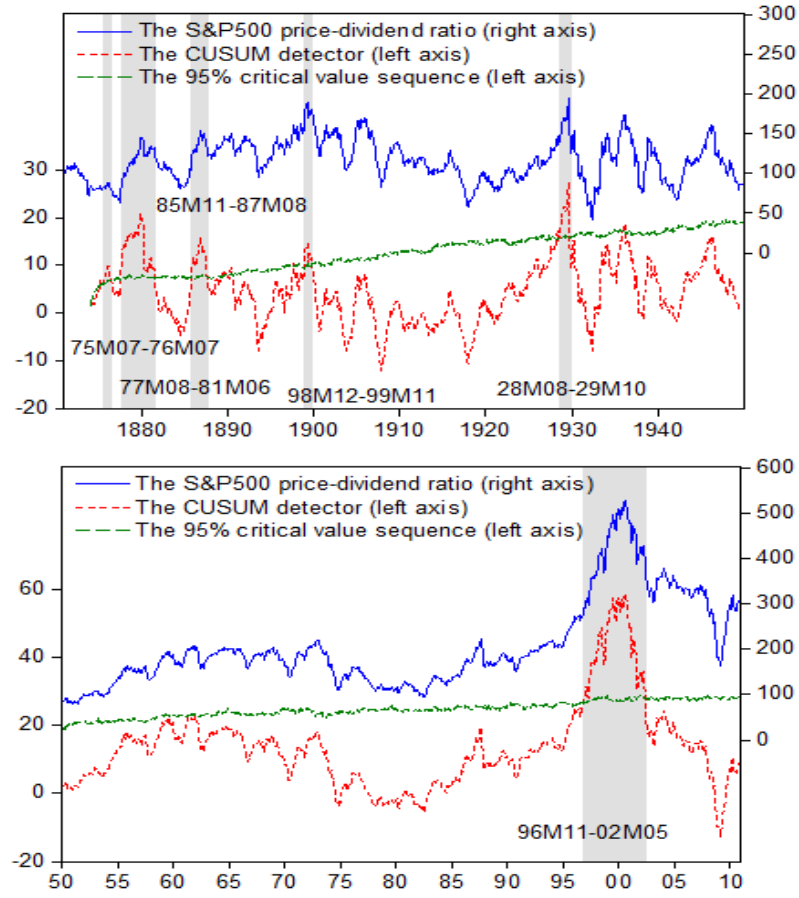
Figure 8: Date-stamping bubble periods in the S&P 500 price-dividend ratio: the CUSUM monitoring procedure.

6   Conclusion

The GSADF test is a rolling window right-sided ADF unit root test with a double-sup window selection criteria.[22] As distinct from the SADF test of PWY, we select a window size using the double-sup criteria and implement the ADF test repeatedly on a sequence of samples, which moves the window frame gradually toward the end of the sample. Experimenting on simulated asset prices reveals one of the shortcomings of the SADF test - its reduced capacity to find and locate bubbles when there are multiple collapsing episodes within the sample range. The GSADF test surmounts this problem and our simulation findings demonstrate that the GSADF test significantly improves discriminatory power in detecting multiple bubbles.

The date-stamping strategy of PWY and the new date-stamping strategy are shown to have quite different behavior under the alternative of multiple bubbles. In particular, when the sample period includes two bubbles and the duration of the first bubble is longer than the second, the strategy of PWY fails to consistently estimate the timing of the second bubble while the new strategy consistently estimates and dates both bubbles.

We apply both SADF and GSADF tests, along with their date-stamping algorithms and the alternative CUSUM monitoring procedure considered in Homm and Breitung (2011), to the S&P 500 price-dividend ratio from January 1871 to December 2010. All three tests find confirmatory evidence of multiple bubble existence. The price-dividend ratio over this historical period contains many individual peaks and troughs, a trajectory that is similar to the multiple bubble scenario for which the PWY date-stamping strategy was found to be inconsistent. The empirical test results confirm the greater discriminatory power of the GSADF strategy found in the simulations and evidenced in the asymptotic theory. The new date-stamping strategy identifies all the well known historical episodes of banking crises and financial bubbles over this long period, whereas both SADF and CUSUM procedures seem more conservative and locate fewer episodes of exuberance and collapse.

---

[22]First, we calculate the sup value of the ADF statistic over the feasible ranges of the window starting points for a fixed window size. Then, we calculate the sup value of the SADF statistic over the feasible range of window sizes.

7 References

Ahamed, L., 2009, *Lords of Finance: The Bankers Who Broke the World,* Penguin Press, New York.

Bhargava, A. 1986, On the theory of testing for unit roots in observed time series, *Review of Economic Studies*, 53, 369–384.

Busetti, F., and A. M. R. Taylor, 2004, Tests of stationarity against a change in persistence, *Journal of Econometrics*, 123, 33–66.

Campbell, J.Y., and Perron, P., 1991, Pitfalls and opportunities: what macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, 6, 141–201.

Campbell, J.Y., and Shiller R.J., 1989, The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies*, 1, 195–228.

Charemza, W.W., and Deadman, D.F., 1995, Speculative bubbles with stochastic explosive roots: the failure of unit root testing. *Journal of Empirical Finance*, 2, 153–163.

Chu, C.J., Stinchcombe, M., and White, H., 1996, Monitoring structural change. *Econometrica*, 64:1045-1065.

Cochrane, J. H. (2005). *Asset Pricing*, Princeton: Princeton University Press.

Cooper, G., 2008, *The Origin of Financial Crises: Central Banks, Credit Bubbles and the Efficient Market Fallacy*, Vintage Books, New York.

Diba, B.T., and Grossman, H.I., 1988, Explosive rational bubbles in stock prices? *The American Economic Review*, 78, 520–530.

Evans, G.W., 1991, Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review*, 81, 922–930.

Ferguson, N., 2008, *The Ascent of Money,* Penguin Press, New York.

Funke, M., Hall, S., and Sola, M., 1994, Rational bubbles during Poland's hyperinflation: implications and empirical evidence. *European Economic Review*, 38, 1257–1276.

Gurkaynak, R. S., 2008, Econometric tests of asset price bubbles: taking stock. *Journal of Economic Surveys*, 22, 166–186.

Hall, S.G., Psaradakis, Z., and Sola M., 1999, Detecting periodically collapsing bubbles: A markov-switching unit root test. *Journal of Applied Econometrics*, 14, 143–154.

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

Homm U. and J. Breitung, 2011, Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods. *Journal of Financial Econometrics*, forthcoming.

Kim, J. Y. 2000, Detection of change in persistence of a linear time series, *Journal of Econometrics*, 95, 97–116.

Kindleberger, C. P., and Aliber, R. Z., 2005, *Manias, Panics and Crashes; A History of Financial Crises*, Hoboken, New Jersey: John Wiley and Sons, Inc.

Lee, J. and P. C. B. Phillips, 2011, Asset Pricing with Financial Bubble Risk. *Yale University, Working Paper.*

Pástor, Luboš, and Pietro Veronesi, 2006, Was there a Nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81, 61–100.

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

Phillips, P.C.B., and Perron, P., 1988, Testing for a unit root in time series regression. *Biometrika*, 75, 335–346.

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

Phillips, P. C. B. and V. Solo, 1992, Asymptotics for Linear Processes. *Annals of Statistics* 20, 971–1001.

Phillips, P.C.B., Shi, S., and Yu, J., 2011a, Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behavior. Working Paper, Sim Kee Boon Institute for Financial Economics, Singapore Management University.

Phillips, P.C.B., Shi, S., and Yu, J., 2011b, Technical Note: Testing for Multiple Bubbles. *Manuscript,* available from `https://sites.google.com/site/shupingshi/TN_GSADF.pdf?attredirects=0&d=1`.

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

Phillips, P.C.B., and Yu, J., 2009, Limit theory for dating the origination and collapse of mildly explosive periods in time series data. *Singapore Management University, Unpublished Manuscript.*

Phillips, P.C.B., and Yu, J., 2011, Dating the Timeline of Financial Bubbles During the Subprime Crisis, *Quantitative Economics*, 2, 455-491.

Psaradakis, Z., Sola, M., and Spagnolo, F., 2001, A simple procedure for detecting periodically collapsing rational bubbles. *Economics Letters*, 24, 317–323.

Shi, S. 2011, Bubbles or volatility: the Markov-switching unit root tests. *Australian National University,* Working Paper.

van Norden, S. and Vigfusson, R., 1998, Avoiding the pitfalls: can regime-switching tests reliably detect bubbles? *Studies in Nonlinear Dynamics & Econometrics*, 3,1–22.

West, K.D., 1988, Dividend Innovations and Stock Price Volatility. *Econometrica*, 56:37–61.

APPENDIX A. The date-stamping strategies (a single bubble)

Notation and useful preliminary lemmas

We define the following notation:

- The bubble period $B = [\tau_e, \tau_f]$, where $\tau_e = \lfloor Tr_e \rfloor$ and $\tau_f = \lfloor Tr_f \rfloor$.

- The normal market periods $N_0 = [1, \tau_e)$ and $N_1 = [\tau_f + 1, \tau]$, where $\tau = \lfloor Tr \rfloor$ is the last observation of the sample.

- The starting point of the regression $\tau_1 = \lfloor Tr_1 \rfloor$, the ending point of the regression $\tau_2 = \lfloor Tr_2 \rfloor$, the regression sample size $\tau_w = \lfloor Tr_w \rfloor$ with $r_w = r_2 - r_1$ and observation $t = \lfloor Tp \rfloor$.

- $B(p) \equiv \sigma W(p)$, where W is a Wiener process.

We use the data generating process

$$
X_t = \begin{cases}
X_{t-1} + \varepsilon_t & \text{for } t \in N_0 \\
\delta_T X_{t-1} + \varepsilon_t & \text{for } t \in B \\
X_{\tau_f}^* + \sum_{k=\tau_f+1}^{t} \varepsilon_k & \text{for } t \in N_1
\end{cases},
\tag{24}
$$

where $\delta_T = 1 + cT^{-\alpha}$ with $c > 0$ and $\alpha \in (0,1)$, $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$ and $X_{\tau_f}^* = X_{\tau_e} + X^*$ with $X^* = O_p(1)$. Under (24) we have the following lemmas.

Lemma 7.1 *Under the generating process (24),*
  (1) For $t \in N_0$, $X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} B(p)$.
  (2) For $t \in B$, $X_{t=\lfloor Tp \rfloor} = \delta_T^{t-\tau_e} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{1/2} \delta_T^{t-\tau_e} B(r_e)$.
  (3) For $t \in N_1$, $X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} [B(p) - B(r_f) + B(r_e)]$.

Proof. (1) For $t \in N_0$, X_t is a unit root process. We know that $T^{-1/2} X_{t=\lfloor Tp \rfloor} \overset{L}{\to} B(p)$ as $T \to \infty$. (2) For $t \in B$, the generating mechanism is

$$
X_t = \delta_T X_{t-1} + \varepsilon_t = \delta_T^{t-\tau_e} X_{\tau_e} + \sum_{j=0}^{t-\tau_e-1} \delta_T^j \varepsilon_{t-j}.
$$

Based on Phillips and Magdalinos (2007, lemma 4.2), we know that for $\alpha < 1$,

$$
T^{-\alpha/2} \sum_{j=0}^{t-\tau_e-1} \delta_T^{-(t-\tau_e)+j} \varepsilon_{t-j} \overset{L}{\to} X_c \equiv N(0, \sigma^2/2c)
$$

as $t - \tau_e \to \infty$. Furthermore, we know that $T^{-1/2} X_{\tau_e - 1} \overset{L}{\to} B(r_e)$ and hence

$$\delta_T^{-(t-\tau_e)} T^{-1/2} X_t = T^{-1/2} X_{\tau_e - 1} + T^{-(1-\alpha)/2} T^{-\alpha/2} \sum_{j=0}^{t-\tau_e-1} \delta_T^{-(t-\tau_e)+j} \varepsilon_{t-j} \overset{L}{\to} B(r_e).$$

This implies that the first term has a higher order than the second term and hence,

$$X_t = \delta_T^{t-\tau_e} X_{\tau_e} \left\{ 1 + \frac{\sum_{j=0}^{t-\tau_e-1} \delta_T^j \varepsilon_{t-j}}{\delta_T^{t-\tau_e} X_{\tau_e}} \right\} = \delta_T^{t-\tau_e} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{1/2} \delta_T^{t-\tau_e} B(r_e).$$

(3) For $t \in N_1$,

$$X_t = \sum_{k=\tau_f+1}^{t} \varepsilon_k + X_{\tau_f}^* = \sum_{k=\tau_f+1}^{t} \varepsilon_k + X_{\tau_e} + X^* \sim_a T^{1/2} [B(p) - B(r_f) + B(r_e)]$$

due to the fact that $X_{\tau_e} \sim_a T^{1/2} B(r_e)$, $\sum_{k=\tau_f+1}^{t} \varepsilon_k \sim_a T^{1/2} [B(p) - B(r_f)]$ and $X^* = O_p(1)$.

∎

Lemma 7.2 *Under the data generating process (24),*
 (1) *For $\tau_1 \in N_0$ and $\tau_2 \in B$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{\alpha - 1/2} \delta_T^{\tau_2 - \tau_e} \frac{1}{r_w c} B(r_e).$$

 (2) *For $\tau_1 \in B$ and $\tau_2 \in N_1$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_f - \tau_1}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{\alpha - 1/2} \delta_T^{\tau_f - \tau_1} \frac{1}{r_w c} B(r_e).$$

 (3) *For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = X_{\tau_e} \frac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} \{1 + o_p(1)\} \sim_a T^{\alpha - 1/2} \delta_T^{\tau_f - \tau_e} \frac{1}{r_w c} B(r_e).$$

Proof. (1) For $\tau_1 \in N_0$ and $\tau_2 \in B$, we have

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_e-1} X_j + \frac{1}{\tau_w} \sum_{j=\tau_e}^{\tau_2} X_j.$$

The first term is

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_e-1} X_j = T^{1/2} \frac{\tau_e - \tau_1}{\tau_w} \left(\frac{1}{\tau_e - \tau_1} \sum_{j=\tau_1}^{\tau_e-1} \frac{X_j}{\sqrt{T}} \right) \sim_a T^{1/2} \frac{r_e - r_1}{r_w} \int_{r_1}^{r_e} B(s)\, ds. \qquad (25)$$

37

The second term is

$$\frac{1}{\tau_w} \sum_{j=\tau_e}^{\tau_2} X_j = \frac{X_{\tau_e}}{\tau_w} \sum_{j=\tau_e}^{\tau_2} \delta_T^{j-\tau_e} \{1 + o_p(1)\} \quad \text{from Lemma 7.1}$$

$$= \frac{1}{\tau_w} \frac{\delta_T^{\tau_2-\tau_e+1} - 1}{\delta_T - 1} X_{\tau_e} \{1 + o_p(1)\}$$

$$= \frac{T^\alpha \delta_T^{\tau_2-\tau_e} + c\delta_T^{\tau_2-\tau_e} - T^\alpha}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\}$$

$$= \frac{T^\alpha \delta_T^{\tau_2-\tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_2-\tau_e} \frac{1}{r_w c} B(r_e). \tag{26}$$

Since

$$\frac{T^{\alpha-1/2}\delta_T^{\tau_2-\tau_e}}{T^{1/2}} = \frac{\delta_T^{\tau_2-\tau_e}}{T^{1-\alpha}} = \frac{e^{c(r_2-r_e)T^{1-\alpha}}}{T^{1-\alpha}} > 1,$$

$\tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} X_j$ has a higher order than $\tau_w^{-1} \sum_{j=\tau_1}^{\tau_e-1} X_j$ and hence

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_2-\tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_2-\tau_e} \frac{1}{r_w c} B(r_e).$$

(2) For $\tau_1 \in B$ and $\tau_2 \in N_1$, we have

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_f} X_j + \frac{1}{\tau_w} \sum_{j=\tau_f+1}^{\tau_2} X_j.$$

The first term is

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_f} X_j = \frac{T^\alpha \delta_T^{\tau_f-\tau_1}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_f-\tau_1} \frac{1}{r_w c} B(r_e).$$

The second term is

$$\frac{1}{\tau_w} \sum_{j=\tau_f+1}^{\tau_2} X_j$$

$$= \frac{1}{\tau_w} \sum_{j=\tau_f+1}^{\tau_2} \left[\sum_{k=\tau_f+1}^{j} \varepsilon_k + X_{\tau_e} \right]$$

$$= T^{1/2} \frac{\tau_2 - \tau_f}{\tau_w} \left[\frac{1}{\tau_2 - \tau_f} \sum_{j=\tau_f+1}^{\tau_2} \left(T^{-1/2} \sum_{k=\tau_f+1}^{j} \varepsilon_k \right) \right] + T^{1/2} \frac{\tau_2 - \tau_f}{\tau_w} \left(T^{-1/2} X_{\tau_e} \right)$$

38

$$\sim_a T^{1/2}\frac{r_2 - r_f}{r_w}\int_{r_f}^{r_2}[B(s) - B(r_f)]\,ds + T^{1/2}\frac{r_2 - r_f}{r_w}B(r_e)$$

$$= T^{1/2}\frac{r_2 - r_f}{r_w}\left\{\int_{r_f}^{r_2}[B(s) - B(r_f)]\,ds - B(r_e)\right\}. \tag{27}$$

We know that $\tau_w^{-1}\sum_{j=\tau_1}^{\tau_f}X_j$ has a higher order than $\tau_w^{-1}\sum_{j=\tau_f+1}^{\tau_2}X_j$ and hence

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_2}X_j = \frac{T^\alpha\delta_T^{\tau_f-\tau_1}}{\tau_w c}X_{\tau_e}\{1 + o_p(1)\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_f-\tau_1}\frac{1}{r_w c}B(r_e).$$

(3) For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_2}X_j = \frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_e-1}X_j + \frac{1}{\tau_w}\sum_{j=\tau_e}^{\tau_f}X_j + \frac{1}{\tau_w}\sum_{j=\tau_f+1}^{\tau_2}X_j.$$

Since

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_e-1}X_j \sim_a T^{1/2}\frac{r_e - r_1}{r_w}\int_{r_1}^{r_e}B(s)\,ds \quad \text{from (25)},$$

$$\frac{1}{\tau_w}\sum_{j=\tau_e}^{\tau_f}X_j = \frac{T^\alpha\delta_T^{\tau_f-\tau_e}}{\tau_w c}X_{\tau_e}\{1 + o_p(1)\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_f-\tau_e}\frac{1}{r_w c}B(r_e),$$

$$\frac{1}{\tau_w}\sum_{j=\tau_f+1}^{\tau_2}X_j \sim_a T^{1/2}\frac{r_2 - r_f}{r_w}\left\{\int_{r_f}^{r_2}[B(s) - B(r_f)]\,ds - B(r_e)\right\} \quad \text{from (27)},$$

it follows that $\tau_w^{-1}\sum_{j=\tau_e}^{\tau_f}X_j$ dominates $\tau_w^{-1}\sum_{j=\tau_1}^{\tau_e-1}X_j$ and $\tau_w^{-1}\sum_{j=\tau_f+1}^{\tau_2}X_j$ and hence

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_2}X_j = \frac{T^\alpha\delta_T^{\tau_f-\tau_e}}{\tau_w c}X_{\tau_e}\{1 + o_p(1)\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_f-\tau_e}\frac{1}{c r_w}B(r_e).$$

∎

Lemma 7.3 *Define the centered quantity $\tilde{X}_t = X_t - \tau_w^{-1}\sum_{j=\tau_1}^{\tau_2}X_j$.*
  (1) For $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\tilde{X}_t = \begin{cases} -\frac{T^\alpha\delta_T^{\tau_2-\tau_e}}{\tau_w c}X_{\tau_e}\{1 + o_p(1)\} & \text{if } t \in N_0 \\ \left[\delta_T^{t-\tau_e} - \frac{T^\alpha\delta_T^{\tau_2-\tau_e}}{\tau_w c}\right]X_{\tau_e}\{1 + o_p(1)\} & \text{if } t \in B \end{cases}.$$

39

(2) *For $\tau_1 \in B$ and $\tau_2 \in N_1$,*

$$\tilde{X}_t = \begin{cases} \left[\delta_T^{t-\tau_e} - \dfrac{T^\alpha \delta_T^{\tau_f - \tau_1}}{\tau_w c} \right] X_{\tau_e} \{1 + o_p(1)\} & \text{if } t \in B \\[3mm] -\dfrac{T^\alpha \delta_T^{\tau_f - \tau_1}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} & \text{if } t \in N_1 \end{cases}.$$

(3) *For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,*

$$\tilde{X}_t = \begin{cases} -\dfrac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\} & \text{if } t \in N_0 \cup N_1 \\[3mm] \left[\delta_T^{t-\tau_e} - \dfrac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} \right] X_{\tau_e} \{1 + o_p(1)\} & \text{if } t \in B \end{cases}.$$

Proof. (1) Suppose $\tau_1 \in N_0$ and $\tau_2 \in B$. If $t \in N_0$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = -\frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\}, \tag{28}$$

where the second term dominates the first term due to the fact that

$$X_t \sim_a T^{1/2} B(p) \quad \text{from Lemma 7.1}$$

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j \sim_a T^{\alpha-1/2} \delta_T^{\tau_2 - \tau_e} \frac{1}{r_w c} B(r_e) \quad \text{from Lemma 7.2.}$$

If $t \in B$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = \left[\delta_T^{t-\tau_e} - \frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{\tau_w c} \right] X_{\tau_e} \{1 + o_p(1)\}.$$

(2) Suppose $\tau_1 \in B$ and $\tau_2 \in N_1$. If $t \in B$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = \left[\delta_T^{t-\tau_e} - \frac{T^\alpha \delta_T^{\tau_f - \tau_1}}{\tau_w c} \right] X_{\tau_e} \{1 + o_p(1)\}.$$

If $t \in N_1$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = -\frac{T^\alpha \delta_T^{\tau_f - \tau_1}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\},$$

where the second term dominates the first term due to the fact that

$$X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} \left[B(p) - B(r_f) + B(r_e) \right] \quad \text{from Lemma 7.1}$$

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j \sim_a T^{\alpha-1/2} \delta_T^{\tau_f - \tau_1} \frac{1}{r_w c} B(r_e) \quad \text{from Lemma 7.2.}$$

40

(3) Suppose $\tau_1 \in N_0$ and $\tau_2 \in N_1$. If $t \in N_0$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = -\frac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\},$$

where the second term dominates the first term due to the fact that

$$X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} B(p) \text{ from Lemma 7.1}$$

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j \sim_a T^{\alpha - 1/2} \delta_T^{\tau_f - \tau_e} \frac{1}{r_w c} B(r_e) \text{ from Lemma 7.2.}$$

If $t \in B$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = \left[\delta_T^{t - \tau_e} - \frac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} \right] X_{\tau_e} \{1 + o_p(1)\}.$$

If $t \in N_1$,

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = -\frac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\},$$

due to the fact that $X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} [B(p) - B(r_f) + B(r_e)]$. ∎

Lemma 7.4 *The sample variance terms involving $\tilde{X}_t$ behave as follows.*
   (1) *For $\tau_1 \in N_0$ and $\tau_2 \in B$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_2 - \tau_e)}}{2c} X_{\tau_e}^2 \{1 + o_p(1)\} \sim_a \frac{T^{1+\alpha} \delta_T^{2(\tau_2 - \tau_e)}}{2c} B(r_e)^2.$$

   (2) *For $\tau_1 \in B$ and $\tau_2 \in N_1$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_f - \tau_e)}}{2c} X_{\tau_e}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_f - \tau_e)}}{2c} B(r_e)^2.$$

   (3) *For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_f - \tau_e)}}{2c} X_{\tau_e}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_f - \tau_e)}}{2c} B(r_e)^2.$$

Proof. (1) For $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \sum_{j=\tau_1}^{\tau_e} \tilde{X}_{j-1}^2 + \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}^2.$$

41

The first term is

$$\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 = \sum_{j=\tau_1}^{\tau_e-1} \frac{T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)}}{\tau_w^2 c^2} X_{\tau_e}^2 \{1+o_p(1)\} \ \text{ from Lemma 7.3}$$

$$= \frac{\tau_e-\tau_1}{\tau_w^2 c^2} T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)} X_{\tau_e}^2 \{1+o_p(1)\} \sim_a \frac{r_e-r_1}{r_w^2 c} T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)} B(r_e).$$

Given that

$$\sum_{j=\tau_e}^{\tau_2} \delta_T^{2(j-1-\tau_e)} = \frac{\delta_T^{2(\tau_2-\tau_e)}-\delta_T^{-2}}{\delta_T^2-1} = \frac{T^{\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c}\{1+o_p(1)\}$$

$$\sum_{j=\tau_e}^{\tau_2} \delta_T^{j-1-\tau_e} = \frac{\delta_T^{\tau_2-\tau_e}-\delta_T^{-1}}{\delta_T-1} = \frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{c}\{1+o_p(1)\},$$

the second term is

$$\sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}^2$$

$$= \sum_{j=\tau_e}^{\tau_2} \left[\delta_T^{j-1-\tau_e} - \frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c}\right]^2 X_{\tau_e}^2 \{1+o_p(1)\}$$

$$= \sum_{j=\tau_e}^{\tau_2} \left[\delta_T^{2(j-1-\tau_e)} - 2\delta_T^{j-1-\tau_e}\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c} + \frac{T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)}}{\tau_w^2 c^2}\right] X_{\tau_e}^2 \{1+o_p(1)\}$$

$$= \left[\frac{T^{\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c} - 2\frac{T^{2\alpha-1}\delta_T^{2(\tau_2-\tau_e)}}{r_w c^2} + \frac{r_2-r_e+\frac{1}{T}}{r_w^2 c^2}T^{2\alpha-1}\delta_T^{2(\tau_2-\tau_e)}\right] X_{\tau_e}^2 \{1+o_p(1)\}$$

$$= \frac{T^{\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c} X_{\tau_e}^2 \{1+o_p(1)\} \ \text{(since } \alpha > 2\alpha - 1\text{)}$$

$$\sim_a \frac{T^{1+\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c} B(r_e)^2.$$

Since $1+\alpha > 2\alpha$, $\sum_{j=\tau_e}^{\tau_2}\tilde{X}_{j-1}^2$ dominates $\sum_{j=\tau_1}^{\tau_e}\tilde{X}_{j-1}^2$ and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^{\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c} X_{\tau_e}^2 \{1+o_p(1)\} \sim_a \frac{T^{1+\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c} B(r_e)^2.$$

(2) For $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2 + \sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}^2.$$

Since

$$\sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2 = \sum_{j=\tau_1}^{\tau_f} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_f-\tau_1}}{\tau_w c} \right]^2 X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_f-\tau_e)}}{2c} B(r_e)^2,$$

$$\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}^2 = \sum_{j=\tau_f+1}^{\tau_2} \frac{T^{2\alpha} \delta_T^{2(\tau_f-\tau_1)}}{\tau_w^2 c^2} X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a \frac{r_2 - r_f}{r_w^2 c^2} T^{2\alpha} \delta_T^{2(\tau_f-\tau_1)} B(r_e)^2,$$

the quantity $\sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2$ dominates $\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}^2$ and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_f-\tau_e)}}{2c} X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_f-\tau_e)}}{2c} B(r_e)^2.$$

(3) For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 + \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2 + \sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}^2.$$

Since

$$\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 = \sum_{j=\tau_1}^{\tau_e-1} \frac{T^{2\alpha} \delta_T^{2(\tau_f-\tau_e)}}{\tau_w^2 c^2} X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a T^{2\alpha} \delta_T^{2(\tau_f-\tau_e)} \frac{r_e - r_1}{r_w^2 c^2} B(r_e)^2,$$

$$\sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2 = \sum_{j=\tau_e}^{\tau_f} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{\tau_w c} \right]^2 X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a T^{\alpha+1} \delta_T^{2(\tau_f-\tau_e)} \frac{B(r_e)^2}{2c},$$

$$\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}^2 = \sum_{j=\tau_f+1}^{\tau_2} \frac{T^{2\alpha} \delta_T^{2(\tau_f-\tau_e)}}{\tau_w^2 c^2} X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a \frac{r_2 - r_f}{r_w^2 c^2} T^{2\alpha} \delta_T^{2(\tau_f-\tau_e)} B(r_e)^2,$$

the component $\sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2$ dominates the other two terms and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_f-\tau_e)}}{2c} X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_f-\tau_e)}}{2c} B(r_e)^2.$$

∎

Lemma 7.5 *The sample covariance of $\tilde{X}_t$ and ε_t behaves as follows.*
   (1) *For $\tau_1 \in N_0$ and $\tau_2 \in B$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j = \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j \left\{ 1 + o_p(1) \right\} \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_2-\tau_e} X_c B(r_e).$$

43

(2) For $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f - \tau_e} X_c B(r_e).$$

(3) For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f - \tau_e} X_c B(r_e).$$

Proof. (1) For $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j.$$

The first term is

$$\sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_e - 1} -\frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{\tau_w c} X_{\tau_e}\varepsilon_j \{1 + o_p(1)\}$$

$$= -\frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{r_w c} \left(T^{-1/2}X_{\tau_e}\right) \left(T^{-1/2}\sum_{j=\tau_1}^{\tau_e - 1} \varepsilon_j\right) \{1 + o_p(1)\}$$

$$\sim_a -\frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{r_w c} B(r_e)[B(r_e) - B(r_1)].$$

The second term is

$$\sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j$$

$$= \sum_{j=\tau_e}^{\tau_2} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_2 - \tau_e}}{\tau_w c}\right] X_{\tau_e}\varepsilon_j \{1 + o_p(1)\}$$

$$= \left[T^{\alpha/2}\delta_T^{\tau_2 - \tau_e}\left(\frac{1}{T^{\alpha/2}}\sum_{j=\tau_e}^{\tau_2}\delta_T^{-(\tau_2 - j + 1)}\varepsilon_j\right) - \frac{\delta_T^{\tau_2 - \tau_e}}{T^{1/2-\alpha}r_w c}\left(\frac{1}{\sqrt{T}}\sum_{j=\tau_e}^{\tau_2}\varepsilon_j\right)\right] X_{\tau_e} \{1 + o_p(1)\}$$

$$= T^{\alpha/2}\delta_T^{\tau_2 - \tau_e}\left(T^{-\alpha/2}\sum_{j=\tau_e}^{\tau_2}\delta_T^{-(\tau_2 - j + 1)}\varepsilon_j\right) X_{\tau_e} \{1 + o_p(1)\} \ \ (\text{since } \alpha/2 > \alpha - 1/2)$$

$$\sim_a T^{(\alpha+1)/2}\delta_T^{\tau_2 - \tau_e} X_c B(r_e).$$

44

Since $(\alpha + 1)/2 > \alpha$, $\sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j$ dominates $\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\varepsilon_j$ and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_2-\tau_e}X_cB(r_e).$$

(2) For $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j.$$

Since

$$\sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_f} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_f-\tau_1}}{\tau_w c}\right] X_{\tau_e}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e}X_cB(r_e),$$

$$\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_f+1}^{\tau_2} -\frac{T^\alpha \delta_T^{\tau_f-\tau_1}}{\tau_w c}X_{\tau_e}\varepsilon_j \{1 + o_p(1)\} \sim_a -\frac{T^\alpha \delta_T^{\tau_f-\tau_1}}{r_w c}B(r_e)[B(r_2) - B(r_f)],$$

the quantity $\sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j$ dominates $\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j$ and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e}X_cB(r_e).$$

(3) For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j.$$

Since

$$\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_1}^{\tau_e-1} -\frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{\tau_w c}X_{\tau_e}\varepsilon_j \{1 + o_p(1)\} \sim_a -\frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{r_w c}B(r_e)[B(r_e) - B(r_1)],$$

$$\sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_e}^{\tau_f} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{\tau_w c}\right] X_{\tau_e}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e}X_cB(r_e),$$

$$\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_f+1}^{\tau_2} -\frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{\tau_w c}X_{\tau_e}\varepsilon_j \{1 + o_p(1)\} \sim_a -\frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{r_w c}B(r_e)[B(r_2) - B(r_f)],$$

and $(\alpha + 1)/2 > \alpha$, the component $\sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j$ dominates the other two terms and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j \{1 + o_p(1)\} \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e}X_cB(r_e).$$

∎

45

Lemma 7.6

The sample covariance of $\tilde{X}_{j-1}$ and $X_j - \delta_T X_{j-1}$ behaves as follows.
  (1) For $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\left(X_j - \delta_T X_{j-1}\right) \sim_a \frac{r_e - r_1}{r_w} T\delta_T^{\tau_2 - \tau_e} B\left(r_e\right) \int_{r_1}^{r_e} B\left(s\right) ds.$$

  (2) For $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\left(X_j - \delta_T X_{j-1}\right) \sim_a -T\delta_T^{2\left(\tau_f - \tau_e\right)+1} B\left(r_e\right)^2.$$

  (3) For $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\left(X_j - \delta_T X_{j-1}\right) \sim_a -T\delta_T^{2\left(\tau_f - \tau_e\right)+1} B\left(r_e\right)^2.$$

Proof. (1) When $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\left(X_j - \delta_T X_{j-1}\right) = \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1}\left(\varepsilon_j - \frac{c}{T^\alpha}X_{j-1}\right)$$

$$= \sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j - \frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1}X_{j-1}. \qquad (29)$$

The first term is

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_2 - \tau_e} X_c B\left(r_e\right) \text{ from Lemma 7.5.}$$

The second term is

$$\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1}X_{j-1}$$

$$= \frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e - 1} -\frac{T^\alpha\delta_T^{\tau_2 - \tau_e}}{\tau_w c}X_{\tau_e}X_{j-1}\left\{1 + o_p\left(1\right)\right\}$$

$$= -\frac{\tau_e - \tau_1}{\tau_w}T\delta_T^{\tau_2 - \tau_e}\left(T^{-1/2}X_{\tau_e}\right)\left[\frac{1}{\tau_e - \tau_1}\sum_{j=\tau_1}^{\tau_e - 1}\left(T^{-1/2}X_{j-1}\right)\right]\left\{1 + o_p\left(1\right)\right\}$$

$$\sim_a -\frac{r_e - r_1}{r_w}T\delta_T^{\tau_2 - \tau_e}B\left(r_e\right)\int_{r_1}^{r_e} B\left(s\right) ds.$$

46

Since $(\alpha+1)/2 < 1$, the quantity $\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}X_{j-1}$ dominates $\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j$ and hence

$$\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\left(X_j-\delta_T X_{j-1}\right)=-\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}X_{j-1}\left\{1+o_p(1)\right\}\sim_a\frac{r_e-r_1}{r_w}T\delta_T^{\tau_2-\tau_e}B\left(r_e\right)\int_{r_1}^{r_e}B\left(s\right)ds.$$

(2) When $\tau_1\in B$ and $\tau_2\in N$,

$$\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\left(X_j-\delta_T X_{j-1}\right)=\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j-\delta_T\tilde{X}_{\tau_f}X_{\tau_f}-\frac{c}{T^\alpha}\sum_{j=\tau_f+2}^{\tau_2}\tilde{X}_{j-1}X_{j-1}.$$

Since

$$\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j\sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e}X_cB\left(r_e\right)\text{ from Lemma 7.5,}$$

$$\delta_T\tilde{X}_{\tau_f}X_{\tau_f}=\delta_T^{\tau_f-\tau_e+1}X_{\tau_e}X_{\tau_f}\left\{1+o_p(1)\right\}\sim_a T\delta_T^{2\left(\tau_f-\tau_e\right)}B\left(r_e\right)^2,$$

$$\frac{c}{T^\alpha}\sum_{j=\tau_f+2}^{\tau_2}\tilde{X}_{j-1}X_{j-1}=\frac{c}{T^\alpha}\sum_{j=\tau_f+2}^{\tau_2}-\frac{T^\alpha\delta_T^{\tau_f-\tau_1}}{\tau_w c}X_{\tau_e}X_{j-1}\left\{1+o_p(1)\right\}\sim_a-\frac{r_2-r_f}{r_w}T\delta_T^{\tau_f-\tau_1}B\left(r_e\right)\int_{r_f}^{r_2}B\left(s\right)ds,$$

the component $\delta_T\tilde{X}_{\tau_f}X_{\tau_f}$ dominates the other two terms and hence

$$\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\left(X_j-\delta_T X_{j-1}\right)=-\delta_T\tilde{X}_{\tau_f}X_{\tau_f}\left\{1+o_p(1)\right\}\sim_a-T\delta_T^{2\left(\tau_f-\tau_e\right)}B\left(r_e\right)^2.$$

(3) When $\tau_1\in N_0$ and $\tau_2\in N_1$,

$$\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\left(X_j-\delta_T X_{j-1}\right)=\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}\left(\varepsilon_j-\frac{c}{T^\alpha}X_{j-1}\right)+\sum_{j=\tau_e}^{\tau_f}\tilde{X}_{j-1}\varepsilon_j$$

$$+\tilde{X}_{\tau_f}\left(X_{\tau_f+1}-\delta_T X_{\tau_f}\right)+\sum_{j=\tau_f+2}^{\tau_2}\tilde{X}_{j-1}\left(\varepsilon_j-\frac{c}{T^\alpha}X_{j-1}\right)$$

$$=\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j-\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}X_{j-1}-\delta_T\tilde{X}_{\tau_f}X_{\tau_f}-\frac{c}{T^\alpha}\sum_{j=\tau_f+2}^{\tau_2}\tilde{X}_{j-1}X_{j-1}.$$

Since

$$\sum_{j=\tau_1}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j\sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e}X_cB\left(r_e\right)\text{ from Lemma 7.5,}$$

$$\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}X_{j-1}=\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1}-\frac{T^\alpha\delta_T^{\tau_f-\tau_e}}{\tau_w c}X_{\tau_e}X_{j-1}\left\{1+o_p(1)\right\}\sim_a-\frac{r_e-r_1}{r_w}T\delta_T^{\tau_f-\tau_e}B\left(r_e\right)\int_{r_1}^{r_e}B\left(s\right)ds,$$

47

$$\delta_T \tilde{X}_{\tau_f} X_{\tau_f} = \delta_T^{\tau_f - \tau_e + 1} X_{\tau_e} X_{\tau_f} \{1 + o_p(1)\} \sim_a T \delta_T^{2(\tau_f - \tau_e)} B(r_e)^2,$$

$$\frac{c}{T^\alpha} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1} X_{j-1} = \frac{c}{T^\alpha} \sum_{j=\tau_f+2}^{\tau_2} -\frac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} X_{\tau_e} X_{j-1} \{1 + o_p(1)\} \sim_a -\frac{r_2 - r_f}{r_w} T \delta_T^{\tau_f - \tau_e} B(r_e) \int_{r_f}^{r_2} B(s)\, ds,$$

the quantity $\delta_T \tilde{X}_{\tau_f} X_{\tau_f}$ dominates the other three terms and hence

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1}) = -\delta_T \tilde{X}_{\tau_f} X_{\tau_f} \{1 + o_p(1)\} \sim_a -T \delta_T^{2(\tau_f - \tau_e)} B(r_e)^2.$$

∎

Test asymptotics

The regression model used for the Dickey-Fuller test is

$$X_t = \alpha_{r_1, r_2} + \delta_{r_1, r_2} X_{t-1} + \varepsilon_t, \varepsilon_t \overset{iid}{\sim} N\left(0, \sigma_{r_2 . r_w}^2\right).$$

First, we calculate the asymptotic distribution of the Dickey-Fuller statistic under the alternative hypothesis. Based on Lemma 7.4 and Lemma 7.6, we can obtain the limit distribution of $\hat{\delta}_{r_1, r_2} - \delta_T$. When $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$T^\alpha \delta_T^{\tau_2 - \tau_e} \left(\hat{\delta}_{r_1, r_2} - \delta_T\right) = \frac{\frac{1}{T\delta_T^{\tau_2 - \tau_e}} \sum_{j=1}^{\tau} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1})}{\frac{1}{T^{1+\alpha} \delta_T^{2(\tau_2 - \tau_e)}} \sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2} \overset{L}{\to} \frac{\frac{r_e - r_1}{r_w} \int_{r_1}^{r_e} B(s)\, ds}{B(r_e)};$$

when $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$T^\alpha \left(\hat{\delta}_{r_1, r_2} - \delta_T\right) = \frac{\frac{1}{T\delta_T^{2(\tau_f - \tau_e)+1}} \sum_{j=1}^{\tau} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1})}{\frac{1}{T^{\alpha+1} \delta_T^{2(\tau_f - \tau_e)}} \sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2} \overset{L}{\to} -2c;$$

when $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$T^\alpha \left(\hat{\delta}_{r_1, r_2} - \delta_T\right) = \frac{\frac{1}{T\delta_T^{2(\tau_f - \tau_e)+1}} \sum_{j=1}^{\tau} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1})}{\frac{1}{T^{\alpha+1} \delta_T^{2(\tau_f - \tau_e)}} \sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2} \overset{L}{\to} -2c.$$

The asymptotic distribution of Dickey-Fuller coefficient statistic (denoted DF^z) is as follows. When $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$DF_{r_1, r_2}^z = \tau_w \left(\hat{\delta}_{r_1, r_2} - 1\right) = \tau_w (\delta_T - 1) + \tau_w \left(\hat{\delta}_{r_1, r_2} - \delta_T\right)$$

48

$$= \tau_w \left(\delta_T - 1 \right) + o_p \left(r_w \frac{T^{1-\alpha}}{\delta_T^{\tau_2 - \tau_e}} \right)$$

$$= \frac{\tau_w c}{T^\alpha} + o_p \left(r_w \frac{T^{1-\alpha}}{e^{c(r_2 - r_e)T^{1-\alpha}}} \right)$$

$$= r_w c T^{1-\alpha} + o_p \left(1 \right) \to \infty.$$

When $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$DF_{r_1, r_2}^z = \tau_w \left(\hat{\delta}_{r_1, r_2} - 1 \right) = \tau_w \left(\delta_T - 1 \right) + \tau_w \left(\hat{\delta}_{r_1, r_2} - \delta_T \right)$$

$$= \frac{\tau_w c}{T^\alpha} + o_p \left(\frac{\tau_w}{T^\alpha} \right)$$

$$= -c r_w T^{1-\alpha} \to -\infty.$$

When $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$DF_{r_1, r_2}^z = \tau_w \left(\hat{\delta}_{r_1, r_2} - 1 \right) = \tau_w \left(\delta_T - 1 \right) + \tau_w \left(\hat{\delta}_{r_1, r_2} - \delta_T \right)$$

$$= \frac{\tau_w c}{T^\alpha} + o_p \left(\frac{\tau_w}{T^\alpha} \right)$$

$$= -c r_w T^{1-\alpha} \to -\infty;$$

This implies that when $\tau_1 \in N_0$ and $\tau_2 \in B$, we have $\hat{\delta}_{r_1, r_2} - 1 \sim_a T^{-\alpha} c$; and for the other two cases, $\hat{\delta}_{r_1, r_2} - 1 \sim_a -T^{-\alpha} c$.

To obtain the asymptotic distribution of the Dickey-Fuller t-statistic, we need to estimate the standard error of $\hat{\delta}_{r_1, r_2}$. (1) When $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$Var \left(\hat{\delta}_{r_1, r_2} \right) = \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \left(\tilde{X}_j - \hat{\delta}_{r_1, r_2} \tilde{X}_{j-1} \right)^2$$

$$= \tau_w^{-1} \left[\sum_{j=\tau_1}^{\tau_e - 1} \left[\varepsilon_j - \left(\hat{\delta}_{r_1, r_2} - 1 \right) \tilde{X}_{j-1} \right]^2 + \sum_{j=\tau_e}^{\tau_2} \left[\varepsilon_j - \left(\hat{\delta}_{r_1, r_2} - \delta_T \right) \tilde{X}_{j-1} \right]^2 \right]$$

$$= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \varepsilon_j^2 + \left(\hat{\delta}_{r_1, r_2} - 1 \right)^2 \tau_w^{-1} \sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1}^2 + \left(\hat{\delta}_{r_1, r_2} - \delta_T \right)^2 \tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}^2$$

$$- 2 \left(\hat{\delta}_{r_1, r_2} - 1 \right) \tau_w^{-1} \sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1} \varepsilon_j - 2 \left(\hat{\delta}_{r_1, r_2} - \delta_T \right) \tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j$$

49

$$= \left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2 \tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}^2 \sim_a \frac{2c}{T^\alpha} \frac{(r_e - r_1)^2}{r_w^3} \left[\int_{r_1}^{r_e} B(s)\,ds\right]^2.$$

The term $\left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2 \tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}^2$ dominates the other terms due to the fact that

$$\left(\hat{\delta}_{r_1,r_2} - 1\right)^2 \tau_w^{-1} \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 = O_p\left(T^{-2\alpha}\right) O_p\left(T^{2\alpha-1}\delta_T^{2(\tau_2-\tau_e)}\right) = O_p\left(T^{-1}\delta_T^{2(\tau_2-\tau_e)}\right),$$

$$\left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2 \tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}^2 = O_p\left(\frac{1}{T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)}}\right) O_p\left(T^\alpha \delta_T^{2(\tau_2-\tau_e)}\right) = O_p\left(T^{-\alpha}\right),$$

$$2\left(\hat{\delta}_{r_1,r_2} - 1\right) \tau_w^{-1} \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\varepsilon_j = O_p\left(T^{-\alpha}\right) O_p\left(\frac{\delta_T^{\tau_2-\tau_e}}{T^{1-\alpha}}\right) = O_p\left(T^{-1}\delta_T^{\tau_2-\tau_e}\right),$$

$$2\left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \tau_w^{-1} \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j = O_p\left(\frac{1}{T^\alpha \delta_T^{\tau_2-\tau_e}}\right) O_p\left(\frac{\delta_T^{\tau_2-\tau_e}}{T^{(1-\alpha)/2}}\right) = O_p\left(T^{-(1+3\alpha)/2}\right).$$

(2) When $\tau_1 \in B$ and $\tau_2 \in N_1$, we know that

$$\tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2}\tilde{X}_{\tau_f} - \varepsilon_{\tau_f+1}$$

$$= \left[X_{\tau_f+1} - \bar{X}\right] - \hat{\delta}_{r_1,r_2}\left[X_{\tau_f} - \bar{X}\right] - \varepsilon_{\tau_f+1}$$

$$= \left[\varepsilon_{\tau_f+1} + X_{\tau_e} + X^* - \bar{X}\right] - \hat{\delta}_{r_1,r_2}\left[X_{\tau_f} - \bar{X}\right] - \varepsilon_{\tau_f+1}$$

$$= X_{\tau_e} + X^* - \tilde{X}_{\tau_f} - \left[\hat{\delta}_{r_1,r_2} - 1\right]\tilde{X}_{\tau_f}$$

$$= O_p\left(T^{1/2}\right) + O_p(1) - O_p\left(T^{1/2}\delta_T^{\tau_f-\tau_e}\right) - O_p\left(T^{1/2-\alpha}\delta_T^{\tau_f-\tau_e}\right)$$

$$= -\tilde{X}_{\tau_f} = -\delta_T^{\tau_f-\tau_e}X_{\tau_e}\{1 + o_p(1)\} \text{ from Lemma 7.3.}$$

The variance of $\hat{\delta}_{r_1,r_2}$ is

$$Var\left(\hat{\delta}_{r_1,r_2}\right) = \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \left(\tilde{X}_j - \hat{\delta}_{r_1,r_2}\tilde{X}_{j-1}\right)^2$$

$$= \tau_w^{-1} \left\{\sum_{j=\tau_f+2}^{\tau_2} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - 1\right)\tilde{X}_{j-1}\right]^2 + \sum_{j=\tau_1}^{\tau_f} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - \delta_T\right)\tilde{X}_{j-1}\right]^2 \right.$$

$$\left. + \left[\tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2}\tilde{X}_{\tau_f}^2 - \varepsilon_{\tau_f+1} + \varepsilon_{\tau_f+1}\right]^2\right\}$$

$$= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \varepsilon_j^2 + \left(\hat{\delta}_{r_1,r_2} - 1\right)^2 \tau_w^{-1} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}^2 + \left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2 \tau_w^{-1} \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2$$

$$- 2 \left(\hat{\delta}_{r_1, r_2} - 1 \right) \tau_w^{-1} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j - 2 \left(\hat{\delta}_{r_1, r_2} - \delta_T \right) \tau_w^{-1} \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1} \varepsilon_j + \tau_w^{-1} \tilde{X}_{\tau_f}^2$$

$$= \tau_w^{-1} \tilde{X}_{\tau_f}^2 = \tau_w^{-1} \delta_T^{2(\tau_f - \tau_e)} X_{\tau_e}^2 \left\{ 1 + o_p(1) \right\} \sim_a \frac{1}{r_w} \delta_T^{2(\tau_f - \tau_e)} B(r_e)^2.$$

The term $\tau_w^{-1} \tilde{X}_{\tau_f}^2$ dominates the other terms due to the fact that

$$\left(\hat{\delta}_{r_1, r_2} - 1 \right)^2 \tau_w^{-1} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}^2 = O_p \left(\frac{\delta_T^{2(\tau_f - \tau_1)}}{T} \right),$$

$$\left(\hat{\delta}_{r_1, r_2} - \delta_T \right)^2 \tau_w^{-1} \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2 = O_p \left(\frac{\delta_T^{2(\tau_f - \tau_e)}}{T^\alpha} \right),$$

$$2 \left(\hat{\delta}_{r_1, r_2} - 1 \right) \tau_w^{-1} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j = O_p \left(\frac{\delta_T^{\tau_f - \tau_1}}{T} \right),$$

$$2 \left(\hat{\delta}_{r_1, r_2} - \delta_T \right) \tau_w^{-1} \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1} \varepsilon_j = O_p \left(\frac{\delta_T^{\tau_f - \tau_e}}{T^{(1+\alpha)/2}} \right),$$

$$\tau_w^{-1} \tilde{X}_{\tau_f}^2 = O_p \left(\delta_T^{2(\tau_f - \tau_e)} \right).$$

(3) When $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$\tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1, r_2} \tilde{X}_{\tau_f} - \varepsilon_{\tau_f+1}$$

$$= X_{\tau_e} + X^* - \tilde{X}_{\tau_f} - \left[\hat{\delta}_{r_1, r_2} - 1 \right] \tilde{X}_{\tau_f}$$

$$= O_p \left(T^{1/2} \right) + o_p(1) - O_p \left(T^{1/2} \delta_T^{\tau_f - \tau_e} \right) - O_p \left(T^{1/2 - \alpha} \delta_T^{\tau_f - \tau_e} \right)$$

$$= -\tilde{X}_{\tau_f} = -\delta_T^{\tau_f - \tau_e} X_{\tau_e} \left\{ 1 + o_p(1) \right\} \text{ from Lemma 7.3.}$$

The variance of $\hat{\delta}_{r_1, r_2}$ is

$$Var \left(\hat{\delta}_{r_1, r_2} \right)$$

$$= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \left(\tilde{X}_j - \hat{\delta}_{r_1, r_2} \tilde{X}_{j-1} \right)^2$$

$$= \tau_w^{-1} \left\{ \sum_{j=\tau_f+2}^{\tau_2} \left[\varepsilon_j - \left(\hat{\delta}_{r_1, r_2} - 1 \right) \tilde{X}_{j-1} \right]^2 + \sum_{j=\tau_1}^{\tau_e-1} \left[\varepsilon_j - \left(\hat{\delta}_{r_1, r_2} - 1 \right) \tilde{X}_{j-1} \right]^2 \right.$$

$$+ \sum_{j=\tau_e}^{\tau_f} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - \delta_T \right) \tilde{X}_{j-1} \right]^2 + \tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2} \tilde{X}_{\tau_f}^2 \Bigg\}^2$$

$$= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \varepsilon_j^2 + \left(\hat{\delta}_{r_1,r_2} - 1 \right)^2 \tau_w^{-1} \left[\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}^2 + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 \right] + \left(\hat{\delta}_{r_1,r_2} - \delta_T \right)^2 \tau_w^{-1} \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2$$

$$- 2 \left(\hat{\delta}_{r_1,r_2} - 1 \right) \tau_w^{-1} \left[\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1} \varepsilon_j \right] - 2 \left(\hat{\delta}_{r_1,r_2} - \delta_T \right) \tau_w^{-1} \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1} \varepsilon_j + \tau_w^{-1} \tau_f^2$$

$$= \tau_w^{-1} \tilde{X}_{\tau_f}^2 = \frac{\delta_T^{2(\tau_f-\tau_e)}}{\tau_w} X_{\tau_e}^2 \left\{ 1 + o_p (1) \right\} \sim_a \frac{\delta_T^{2(\tau_f-\tau_e)}}{r_w} B \left(r_e \right)^2 .$$

The term $\tau_w^{-1} \tilde{X}_{\tau_f}^2$ dominates the other terms due to the fact that

$$\left(\hat{\delta}_{r_1,r_2} - 1 \right)^2 \frac{1}{\tau_w} \left[\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}^2 + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 \right] = O_p \left(\frac{\delta_T^{2(\tau_f-\tau_e)}}{T} \right) ,$$

$$\left(\hat{\delta}_{r_1,r_2} - \delta_T \right)^2 \frac{1}{\tau_w} \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2 = O_p \left(\frac{\delta_T^{2(\tau_f-\tau_e)}}{T^\alpha} \right) ,$$

$$2 \left(\hat{\delta}_{r_1,r_2} - 1 \right) \frac{1}{\tau_w} \left[\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1} \varepsilon_j \right] = O_p \left(\frac{\delta_T^{\tau_f-\tau_e}}{T} \right) ,$$

$$2 \left(\hat{\delta}_{r_1,r_2} - \delta_T \right) \frac{1}{\tau_w} \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1} \varepsilon_j = O_p \left(\frac{\delta_T^{\tau_f-\tau_e}}{T^{(1+\alpha)/2}} \right) ,$$

$$\tau_w^{-1} \tilde{X}_{\tau_f}^2 = O_p \left(\delta_T^{2(\tau_f-\tau_e)} \right) .$$

The asymptotic distribution of the DF t-statistic can be calculated as follows. When $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$DF_{r_1,r_2}^t = \left(\frac{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2}{\hat{\sigma}^2} \right)^{1/2} \left(\hat{\delta}_{r_1,r_2} - 1 \right) \sim_a \frac{T^{1/2} \delta_T^{\tau_2-\tau_e} r_w^{3/2} B \left(r_e \right)}{2 \left(r_e - r_1 \right) \int_{r_1}^{r_e} B \left(s \right) ds} \to \infty.$$

When $\tau_1 \in B$ and $\tau_2 \in N_1$,

$$DF_{r_1,r_2}^t = \left(\frac{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2}{\hat{\sigma}^2} \right)^{1/2} \left(\hat{\delta}_{r_1,r_2} - 1 \right) \sim_a - \left(\frac{1}{2} c r_w \right)^{1/2} T^{(1-\alpha)/2} \to -\infty.$$

When $\tau_1 \in N_0$ and $\tau_2 \in N_1$,

$$DF_{r_1,r_2}^t = \left(\frac{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2}{\hat{\sigma}^2}\right)^{1/2} \left(\hat{\delta}_{r_1,r_2} - 1\right) \sim_a - \left(\frac{1}{2}cr_w\right)^{1/2} T^{(1-\alpha)/2} \to -\infty.$$

The date-stamping strategy of PWY

The origination of the bubble expansion and the termination of the bubble collapse based on the backward DF test are identified as

$$\hat{r}^e = \inf_{r_2 \in [r_0,1]} \left\{r_2 : BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} \text{ and } \hat{r}^f = \inf_{r_2 \in [\hat{r}_e + \log(T)/T, 1]} \left\{r_2 : BDF_{r_2} < cv_{r_2}^{\beta_T}\right\}.$$

We know that when $\beta_T \to 0$, $cv_{r_2}^{\beta_T} \to \infty$.

The asymptotic distributions of the backward DF statistic under the alternative hypothesis are

$$BDF_{r_2} \sim_a \begin{cases} F_{r_2}(W) & \text{if } r_2 \in N_0 \\ T^{1/2}\delta_T^{\tau_2-\tau_e} \frac{r_w^{3/2} B(r_e)}{2(r_e-r_1)\int_{r_1}^{r_e} B(s)ds} \to \infty & \text{if } r_2 \in B \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} \to -\infty & \text{if } r_2 \in N_1 \end{cases}.$$

It is obvious that if $r_2 \in N_0$,

$$\lim_{T \to \infty} \Pr\left\{BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} = \Pr\left\{F_{r_2}(W) = \infty\right\} = 0.$$

If $r_2 \in B$, $\lim_{T\to\infty} \Pr\left\{BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} = 1$ provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_e}} \to 0$. This implies that provided $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T\to\infty} \Pr\left\{BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} = 1$ for any $r_2 \in B$. If $r_2 \in N_1$, $\lim_{T\to\infty} \Pr\left\{BDF_{r_2} < cv_{r_2}^{\beta_T}\right\} = \lim_{T\to\infty} \Pr\left\{-T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} < cv_{r_2}^{\beta_T}\right\} = 1$.

It follows that for any $\eta, \gamma > 0$,

$$\Pr\left\{\hat{r}_e > r_e + \eta\right\} \to 0 \text{ and } \Pr\left\{\hat{r}_f < r_f - \gamma\right\} \to 0$$

due to the fact that $\Pr\left\{BDF_{r_e+a_\eta} > cv_{r_2}^{\beta_T}\right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{BDF_{r_f-a_\gamma} > cv_{r_2}^{\beta_T}\right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary, $\Pr\left\{\hat{r}_e < r_e\right\} \to 0$ and $\Pr\left\{\hat{r}_f > r_f\right\} \to 0$, we deduce that $\Pr\left\{|\hat{r}_e - r_e| > \eta\right\} \to 0$ and $\Pr\left\{|\hat{r}_f - r_f| > \gamma\right\} \to 0$ as $T \to \infty$, provided that

$$\frac{1}{cv_{r_2}^{\beta_T}} + \frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0.$$

Therefore, $\hat{r}_e$ and $\hat{r}_f$ are consistent estimators of r_e and r_f.

The new date-stamping strategy

The origination of the bubble expansion and the termination of the bubble collapse based on the backward sup DF test are identified as

$$\hat{r}^e = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\},$$

$$\hat{r}^f = \inf_{r_2 \in [\hat{r}_e + \delta \log(T)/T, 1]} \left\{ r_2 : BSDF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}.$$

We know that when $\beta_T \to 0$, $scv_{r_2}^{\beta_T} \to \infty$.

The asymptotic distributions of the backward sup DF statistic under the alternative hypothesis are

$$BSDF_{r_2}(r_0) \sim_a \begin{cases} F_{r_2}^{r_0}(W) & \text{if } r_2 \in N_0 \\ T^{1/2} \delta_T^{\tau_2 - \tau_e} \sup_{r_1 \in [0, r_2 - r_0]} \left\{ \frac{r_w^{3/2} B(r_e)}{2(r_e - r_1) \int_{r_1}^{r_e} B(s) ds} \right\} & \text{if } r_2 \in B \\ -T^{(1-\alpha)/2} \sup_{r_1 \in [0, r_2 - r_0]} \left\{ \left(\frac{1}{2} cr_w \right)^{1/2} \right\} \to \infty & \text{if } r_2 \in N_1 \end{cases}.$$

It is obvious that if $r_2 \in N_0$,

$$\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} = \Pr \left\{ F_{r_2}^{r_0}(W) = \infty \right\} = 0.$$

If $r_2 \in B$, $\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} = 1$ provided that $\frac{scv_{r_2}^{\beta_T}}{T^{1/2} \delta_T^{\tau_2 - \tau_e}} \to 0$. This implies that provided $\frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} = 1$ for any $r_2 \in B$. If $r_2 \in N_1$, $\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\} = 1$.

It follows that for any $\eta, \gamma > 0$,

$$\Pr \left\{ \hat{r}_e > r_e + \eta \right\} \to 0 \text{ and } \Pr \left\{ \hat{r}_f < r_f - \gamma \right\} \to 0,$$

since $\Pr \left\{ BSDF_{r_e + a_\eta}(r_0) > scv_{r_2}^{\beta_T} \right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr \left\{ BSDF_{r_f - a_\gamma}(r_0) > scv_{r_2}^{\beta_T} \right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary, $\Pr \left\{ \hat{r}_e < r_e \right\} \to 0$ and $\Pr \left\{ \hat{r}_f > r_f \right\} \to 0$, we deduce that $\Pr \left\{ |\hat{r}_e - r_e| > \eta \right\} \to 0$ and $\Pr \left\{ |\hat{r}_f - r_f| > \gamma \right\} \to 0$ as $T \to \infty$, provided that

$$\frac{1}{scv_{r_2}^{\beta_T}} + \frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0.$$

Therefore, $\hat{r}_e$ and $\hat{r}_f$ are consistent estimators of r_e and r_f.

APPENDIX B. Date-stamping strategies (two bubbles)

Notation and lemmas

- The two bubble periods are $B_1 = [\tau_{1e}, \tau_{1f}]$ and $B_2 = [\tau_{2e}, \tau_{2f}]$, where $\tau_{1e} = \lfloor Tr_{1e} \rfloor$, $\tau_{1f} = \lfloor Tr_{1f} \rfloor$, $\tau_{2e} = \lfloor Tr_{2e} \rfloor$ and $\tau_{2f} = \lfloor Tr_{2f} \rfloor$.

- The normal periods are $N_0 = [1, \tau_{1e})$, $N_1 = (\tau_{1f}, \tau_{2e})$, $N_2 = (\tau_{2f}, \tau]$, where $\tau = \lfloor Tr \rfloor$ is the last observation of the sample.

- We assume that $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$.

We use the data generating process

$$
X_t = \begin{cases}
X_{t-1} + \varepsilon_t & \text{for } t \in N_0 \\
\delta_T X_{t-1} + \varepsilon_t & \text{for } t \in B_i \text{ with } i = 1, 2 \\
X^*_{\tau_{if}} + \sum_{k=\tau_{if}+1}^{t} \varepsilon_k & \text{for } t \in N_i \text{ with } i = 1, 2
\end{cases}
\tag{30}
$$

where $\delta_T = 1 + cT^{-\alpha}$ with $c > 0$ and $\alpha \in (0, 1)$, $\varepsilon_t \overset{iid}{\sim} N(0, \sigma^2)$ and $X^*_{\tau_{if}} = X_{\tau_{ie}} + X^*$ with $X^* = O_p(1)$ for $i = 1, 2$. We have the following lemmas.

Lemma 7.7 *Under the data generating process (30),*
(1) *For $t \in N_0$, $X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} B(p)$.*
(2) *For $t \in B_i$ with $i = 1, 2$, $X_{t=\lfloor Tp \rfloor} = \delta_T^{t-\tau_{ie}} X_{\tau_{ie}} \{1 + o_p(1)\} \sim_a T^{1/2} \delta_T^{t-\tau_{ie}} B(r_{ie})$.*
(3) *For $t \in N_i$ with $i = 1, 2$, $X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} [B(p) - B(r_{if}) + B(r_{ie})]$.*

Lemma 7.8 *Under the data generating process (30),*
(1) *For $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,*

$$
\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_2 - \tau_{ie}}}{\tau_w c} X_{\tau_{ie}} \{1 + o_p(1)\} \sim_a T^{\alpha - 1/2} \delta_T^{\tau_2 - \tau_{ie}} \frac{1}{r_w c} B(r_{ie}).
$$

(2) *For $\tau_1 \in B_i$ and $\tau_2 \in N_i$ with $i = 1, 2$,*

$$
\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_{if} - \tau_1}}{\tau_w c} X_{\tau_{ie}} \{1 + o_p(1)\} \sim_a T^{\alpha - 1/2} \delta_T^{\tau_{if} - \tau_1} \frac{1}{r_w c} B(r_{ie}).
$$

(3) *For $\tau_1 \in N_{i-1}$ and $\tau_2 \in N_i$ with $i = 1, 2$,*

$$
\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = X_{\tau_{ie}} \frac{T^\alpha \delta_T^{\tau_{if} - \tau_{ie}}}{\tau_w c} \{1 + o_p(1)\} \sim_a T^{\alpha - 1/2} \delta_T^{\tau_{if} - \tau_{ie}} \frac{1}{r_w c} B(r_{ie}).
$$

(4) *For $\tau_1 \in N_0$ and $\tau_2 \in N_2$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_{1f}-\tau_{1e}}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2} \delta_T^{\tau_{1f}-\tau_{1e}} \frac{1}{r_w c} B(r_{1e}).$$

(5) *For $\tau_1 \in B_1$ and $\tau_2 \in B_2$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \begin{cases} \frac{T^\alpha \delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2} \delta_T^{\tau_{1f}-\tau_1} \frac{1}{r_w c} B(r_{1e}) & if \ \tau_{1f} - \tau_1 > \tau_2 - \tau_{2e} \\ \frac{T^\alpha \delta_T^{\tau_2-\tau_{2e}}}{\tau_w c} X_{\tau_{2e}} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2} \delta_T^{\tau_2-\tau_{2e}} \frac{1}{r_w c} B(r_{2e}) & if \ \tau_{1f} - \tau_1 \le \tau_2 - \tau_{2e} \end{cases}.$$

(6) *For $\tau_1 \in B_1$ and $\tau_2 \in N_2$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \begin{cases} \frac{T^\alpha \delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2} \delta_T^{\tau_{1f}-\tau_1} \frac{1}{r_w c} B(r_1) & if \ \tau_{1f} - \tau_1 > \tau_{2f} - \tau_{2e} \\ \frac{T^\alpha \delta_T^{\tau_{2f}-\tau_{2e}}}{\tau_w c} X_{\tau_{2e}} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2} \delta_T^{\tau_{2f}-\tau_{2e}} \frac{1}{r_w c} B(r_{2e}) & if \ \tau_{1f} - \tau_1 \le \tau_{2f} - \tau_{2e} \end{cases}.$$

(7) *For $\tau_1 \in N_0$ and $\tau_2 \in B_2$,*

$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^\alpha \delta_T^{\tau_{1f}-\tau_{1e}}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} \sim_a T^{\alpha-1/2} \delta_T^{\tau_{1f}-\tau_{1e}} \frac{1}{r_w c} B(r_{1e}).$$

Lemma 7.9 *Define the centered quantity $\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j$.*
(1) *For $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,*

$$\tilde{X}_t = \begin{cases} -\frac{T^\alpha \delta_T^{\tau_2-\tau_{ie}}}{\tau_w c} X_{\tau_{ie}} \{1 + o_p(1)\} & if \ t \in N_{i-1} \\ \left[\delta_T^{t-\tau_{ie}} - \frac{T^\alpha \delta_T^{\tau_2-\tau_{ie}}}{\tau_w c} \right] X_{\tau_{ie}} \{1 + o_p(1)\} & if \ t \in B_i \end{cases}.$$

(2) *For $\tau_1 \in B_i$ and $\tau_2 \in N_i$ with $i = 1, 2$,*

$$\tilde{X}_t = \begin{cases} \left[\delta_T^{t-\tau_{ie}} - \frac{T^\alpha \delta_T^{\tau_{if}-\tau_1}}{\tau_w c} \right] X_{\tau_{ie}} \{1 + o_p(1)\} & if \ t \in B_i \\ -\frac{T^\alpha \delta_T^{\tau_{if}-\tau_1}}{\tau_w c} X_{\tau_{ie}} \{1 + o_p(1)\} & if \ t \in N_i \end{cases}.$$

(3) *For $\tau_1 \in N_{i-1}$ and $\tau_2 \in N_i$ with $i = 1, 2$,*

$$\tilde{X}_t = \begin{cases} -\frac{T^\alpha \delta_T^{\tau_{if}-\tau_{ie}}}{\tau_w c} X_{\tau_{ie}} \{1 + o_p(1)\} & if \ t \in N_{i-1} \cup N_i \\ \left[\delta_T^{t-\tau_{ie}} - \frac{T^\alpha \delta_T^{\tau_{if}-\tau_{ie}}}{\tau_w c} \right] X_{\tau_{ie}} \{1 + o_p(1)\} & if \ t \in B_i \end{cases}.$$

(4) *For* $\tau_1 \in N_0$ *and* $\tau_2 \in N_2$,

$$\tilde{X}_t = \begin{cases} -\dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_{1e}}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} & \text{if } t \in N_i \\ \left[\delta_T^{t-\tau_{ie}} X_{\tau_{ie}} - \dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_{1e}}}{\tau_w c} X_{\tau_{1e}}\right] \{1 + o_p(1)\} & \text{if } t \in B_i, \ i = 1,2, \end{cases}.$$

(5) *For* $\tau_1 \in B_1$ *and* $\tau_2 \in B_2$, *if* $\tau_{1f} - \tau_1 > \tau_2 - \tau_{2e}$,

$$\tilde{X}_t = \begin{cases} \left[\delta_T^{t-\tau_{ie}} X_{\tau_{ie}} - \dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}}\right] \{1 + o_p(1)\} & \text{if } t \in B_i, \ i = 1,2, \\ -\dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} & \text{if } t \in N_1 \end{cases}.$$

and if $\tau_{1f} - \tau_1 \leq \tau_2 - \tau_{2e}$

$$\tilde{X}_t = \begin{cases} \left[\delta_T^{t-\tau_{ie}} X_{\tau_{ie}} - \dfrac{T^\alpha \delta_T^{\tau_2-\tau_{2e}}}{\tau_w c} X_{\tau_{2e}}\right] \{1 + o_p(1)\} & \text{if } t \in B_i, \ i = 1,2, \\ -\dfrac{T^\alpha \delta_T^{\tau_2-\tau_{2e}}}{\tau_w c} X_{\tau_{2e}} \{1 + o_p(1)\} & \text{if } t \in N_1 \end{cases}.$$

(6) *For* $\tau_1 \in B_1$ *and* $\tau_2 \in N_2$, *if* $\tau_{1f} - \tau_1 > \tau_{2f} - \tau_{2e}$,

$$\tilde{X}_t = \begin{cases} \left[\delta_T^{t-\tau_{ie}} X_{\tau_{ie}} - \dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}}\right] \{1 + o_p(1)\} & \text{if } t \in B_i, \ i = 1,2, \\ -\dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} & \text{if } t \in N_i, \ i = 1,2, \end{cases}.$$

and if $\tau_{1f} - \tau_1 \leq \tau_{2f} - \tau_{2e}$,

$$\tilde{X}_t = \begin{cases} \left[\delta_T^{t-\tau_{ie}} X_{\tau_{ie}} - \dfrac{T^\alpha \delta_T^{\tau_{2f}-\tau_{2e}}}{\tau_w c} X_{\tau_{2e}}\right] \{1 + o_p(1)\} & \text{if } t \in B_i, \ i = 1,2, \\ -\dfrac{T^\alpha \delta_T^{\tau_{2f}-\tau_{2e}}}{\tau_w c} X_{\tau_{2e}} \{1 + o_p(1)\} & \text{if } t \in N_i, \ i = 1,2, \end{cases}.$$

(7) *For* $\tau_1 \in N_0$ *and* $\tau_2 \in B_2$,

$$\tilde{X}_t = \begin{cases} -\dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_{1e}}}{\tau_w c} X_{\tau_{1e}} \{1 + o_p(1)\} & \text{if } t \in N_i, \ i = 1,2, \\ \left[\delta_T^{t-\tau_{ie}} X_{\tau_{ie}} - \dfrac{T^\alpha \delta_T^{\tau_{1f}-\tau_{1e}}}{\tau_w c} X_{\tau_{1e}}\right] \{1 + o_p(1)\} & \text{if } t \in B_i, \ i = 1,2, \end{cases}.$$

Lemma 7.10 *The sample variance of* $\tilde{X}_t$ *has the following limit form:*
(1) *For* $\tau_1 \in N_{i-1}$ *and* $\tau_2 \in B_i$ *with* $i = 1,2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_2-\tau_{ie})}}{2c} X_{\tau_{ie}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{1+\alpha}\delta_T^{2(\tau_2-\tau_{ie})}}{2c} B(r_{ie})^2.$$

57

(2) For $\tau_1 \in B_i$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_{if} - \tau_{ie})}}{2c} X_{\tau_{ie}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_{if} - \tau_{ie})}}{2c} B(r_{ie})^2.$$

(3) For $\tau_1 \in N_{i-1}$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_{if} - \tau_{ie})}}{2c} X_{\tau_{ie}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_{if} - \tau_{ie})}}{2c} B(r_{ie})^2.$$

(4) For $\tau_1 \in N_0$ and $\tau_2 \in N_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} X_{\tau_{1e}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} B(r_{1e})^2.$$

(5) For $\tau_1 \in B_1$ and $\tau_2 \in B_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} X_{\tau_{1e}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} B(r_{1e})^2.$$

(6) For $\tau_1 \in B_1$ and $\tau_2 \in N_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} X_{\tau_{1e}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} B(r_{1e})^2.$$

(7) For $\tau_1 \in N_0$ and $\tau_2 \in B_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \frac{T^\alpha \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} X_{\tau_{1e}}^2 \{1 + o_p(1)\} \sim_a \frac{T^{\alpha+1} \delta_T^{2(\tau_{1f} - \tau_{1e})}}{2c} B(r_{1e})^2.$$

Lemma 7.11 *The sample covariance of $\tilde{X}_t$ and ε_t has the following limit form:*
(1) *For $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_2 - \tau_{ie}} X_c B(r_{ie}).$$

(2) *For $\tau_1 \in B_i$ and $\tau_2 \in N_i$ with $i = 1, 2$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_{if} - \tau_{ie}} X_c B(r_{ie}).$$

(3) For $\tau_1 \in N_{i-1}$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_{if} - \tau_{ie}} X_c B(r_{ie}).$$

(4) For $\tau_1 \in N_0$ and $\tau_2 \in N_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j \sim_a T^{(1+\alpha)/2} \delta_T^{\tau_{1f} - \tau_{1e}} X_c B(r_{1e}).$$

(5) For $\tau_1 \in B_1$ and $\tau_2 \in B_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_{1f} - \tau_{1e}} X_c B(r_{1e}).$$

(6) For $\tau_1 \in B_1$ and $\tau_2 \in N_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j \sim_a T^{(1+\alpha)/2} \delta_T^{\tau_{1f} - \tau_{1e}} X_c B(r_{1e}).$$

(7) For $\tau_1 \in N_0$ and $\tau_2 \in B_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_{1f} - \tau_{1e}} X_c B(r_{1e}).$$

Lemma 7.12 *The sample covariance of $\tilde{X}_{j-1}$ and $X_j - \delta_T X_{j-1}$ has the following limit form:*
(1) For $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1}) \sim_a \frac{r_{ie} - r_1}{r_w} T \delta_T^{\tau_2 - \tau_{ie}} B(r_{ie}) \int_{r_1}^{r_{ie}} B(s) \, ds.$$

(2) For $\tau_1 \in B_i$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1}) \sim_a -T \delta_T^{2(\tau_{if} - \tau_{ie})} B(r_{ie})^2.$$

(3) For $\tau_1 \in N_{i-1}$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} (X_j - \delta_T X_{j-1}) \sim_a \sim_a -T \delta_T^{2(\tau_{if} - \tau_{ie})} B(r_{ie})^2.$$

(4) *For $\tau_1 \in N_0$ and $\tau_2 \in N_2$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \left(X_j - \delta_T X_{j-1} \right) \sim_a -T\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)} B\left(r_{1e}\right)^2.$$

(5) *For $\tau_1 \in B_1$ and $\tau_2 \in B_2$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \left(X_j - \delta_T X_{j-1} \right) \sim_a -T\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)} B\left(r_{1e}\right)^2.$$

(6) *For $\tau_1 \in B_1$ and $\tau_2 \in N_2$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \left(X_j - \delta_T X_{j-1} \right) \sim_a -T\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)} B\left(r_{1e}\right)^2.$$

(7) *For $\tau_1 \in N_0$ and $\tau_2 \in B_2$,*

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \left(X_j - \delta_T X_{j-1} \right) \sim_a -T\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)} B\left(r_{1e}\right)^2.$$

We refer to Appendix A for the proof of Lemma 7.7, Lemma 7.8, Lemma 7.9, Lemma 7.10, Lemma 7.11 and Lemma 7.12. A more detailed proof of Appendix B is available online at `https://sites.google.com/site/shupingshi/TN_GSADF.pdf?attredirects=0&d=1`.

Test asymptotics

The regression model for the Dickey-Fuller test is

$$X_t = \alpha_{r_1,r_2} + \delta_{r_1,r_2} X_{t-1} + \varepsilon_t, \varepsilon_t \overset{iid}{\sim} N\left(0, \sigma_{r_2.r_w}^2\right).$$

First, we calculate the asymptotic distribution of the Dickey-Fuller statistic under the alternative hypothesis. Based on Lemma 7.10 and Lemma 7.12, we can obtain the limit distribution of $\hat{\delta}_{r_1,r_2} - \delta_T$. When $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$T^\alpha \delta_T^{\tau_2 - \tau_{ie}} \left(\hat{\delta}_T - \delta_T \right) \overset{L}{\to} \frac{\left(r_{ie} - r_1\right) \int_{r_1}^{r_{ie}} B\left(s\right) ds}{r_w B\left(r_{ie}\right)};$$

for all other cases,

$$T^\alpha \left(\hat{\delta}_T - \delta_T \right) \overset{L}{\to} -2c.$$

The asymptotic distribution of Dickey-Fuller coefficient statistic is

$$DF^z_{r_1,r_2} = \tau_w \left(\hat{\delta}_{r_1,r_2} - 1 \right) = \tau_w \left(\delta_T - 1 \right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T \right).$$

When $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$DF^z_{r_1,r_2} = r_w c T^{1-\alpha} + o_p(1) \rightarrow \infty;$$

for all other cases,

$$DF^z_{r_1,r_2} = -r_w c T^{1-\alpha} + o_p(1) \rightarrow -\infty.$$

This implies that when $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$\hat{\delta}_{r_1,r_2} - 1 \sim_a cT^{-\alpha};$$

for all other cases,

$$\hat{\delta}_{r_1,r_2} - 1 \sim_a -cT^{-\alpha}.$$

To obtain the asymptotic distribution of the Dickey-Fuller t-statistic, we need to estimate the standard error of $\hat{\delta}_{r_1,r_2}$. When $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$Var\left(\hat{\delta}_{r_1,r_2} \right) \sim_a \frac{2c}{T^\alpha} \frac{(r_{ie} - r_1)^2}{r_w^3} \left[\int_{r_1}^{r_{ie}} B(s) \, ds \right]^2;$$

when $\tau_1 \in B_i$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$Var\left(\hat{\delta}_{r_1,r_2} \right) \sim_a \delta_T^{2\left(\tau_{if} - \tau_{ie} \right)} r_w^{-1} B(r_{ie})^2;$$

when $\tau_1 \in N_{i-1}$ and $\tau_2 \in N_i$ with $i = 1, 2$,

$$Var\left(\hat{\delta}_{r_1,r_2} \right) \sim_a \delta_T^{2\left(\tau_{if} - \tau_{ie} \right)} r_w^{-1} B(r_{ie})^2;$$

when $\tau_1 \in N_0$ and $\tau_2 \in N_2$,

$$Var\left(\hat{\delta}_{r_1,r_2} \right) \sim_a \delta_T^{2\left(\tau_{1f} - \tau_{1e} \right)} r_w^{-1} B(r_{1e})^2;$$

when $\tau_1 \in B_1$ and $\tau_2 \in B_2$,

$$Var\left(\hat{\delta}_{r_1,r_2} \right) \sim_a \delta_T^{2\left(\tau_{1f} - \tau_{1e} \right)} r_w^{-1} B(r_{1e})^2;$$

61

when $\tau_1 \in B_1$ and $\tau_2 \in N_2$,

$$Var\left(\hat{\delta}_{r_1,r_2}\right) \sim_a \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)} r_w^{-1} B\left(r_{1e}\right)^2;$$

when $\tau_1 \in N_0$ and $\tau_2 \in B_2$,

$$Var\left(\hat{\delta}_{r_1,r_2}\right) \sim_a \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)} \frac{1}{r_w} B\left(r_{1e}\right)^2.$$

The asymptotic distributions of the DF t-statistic can be calculated as

$$DF_{r_1,r_2}^t = \left(\frac{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2}{\hat{\sigma}^2}\right)^{1/2} \left(\hat{\delta}_{r_1,r_2} - 1\right).$$

When $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$DF_{r_1,r_2}^t \sim_a T^{1/2} \delta_T^{\tau_2 - \tau_{ie}} \frac{r_w^{3/2} B\left(r_{ie}\right)}{2\left(r_{ie} - r_1\right) \int_{r_1}^{r_{ie}} B\left(s\right) ds} \to \infty;$$

for all other cases,

$$DF_{r_1,r_2}^t \sim_a -\left(\frac{1}{2} cr_w\right)^{1/2} T^{(1-\alpha)/2} \to -\infty.$$

The date-stamping strategy of PWY

The origination of the bubble expansion r_{1e}, r_{2e} and the termination of the bubble collapse r_{1f}, r_{2f} based on the backward DF test are identified as

$$\hat{r}_{1e} = \inf_{r_2 \in [r_0,1]} \left\{r_2 : BDF_{r_2} > cv_{r_2}^{\beta_T}\right\},$$

$$\hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e}+\log(T)/T,1]} \left\{r_2 : BDF_{r_2} < cv_{r_2}^{\beta_T}\right\},$$

$$\hat{r}_{2e} = \inf_{r_2 \in (\hat{r}_{1f},1]} \left\{r_2 : BDF_{r_2} > cv_{r_2}^{\beta_T}\right\},$$

$$\hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e}+\log(T)/T,1]} \left\{r_2 : BDF_{r_2} < cv_{r_2}^{\beta_T}\right\}.$$

We know that when $\beta_T \to 0$, $cv_{r_2}^{\beta_T} \to \infty$.

The asymptotic distributions of the backward DF statistic under the alternative hypothesis are (given $r_1 \in N_0$ and $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$)

$$
BDF_{r_2} \sim_a \begin{cases} F_{r_2}(W) & \text{if } r_2 \in N_0 \\ T^{1/2}\delta_T^{\tau_2-\tau_{1e}} \frac{r_w^{3/2} B(r_{1e})}{2(r_{1e}-r_1)\int_{r_1}^{r_{1e}} B(s)ds} & \text{if } r_2 \in B_1 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r_2 \in N_1 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r_2 \in B_2 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r_2 \in N_2 \end{cases}.
$$

It is obvious that if $r_2 \in N_0$,

$$
\lim_{T \to \infty} \Pr\left\{ BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} = \Pr\{F_{r_2}(W) = \infty\} = 0.
$$

If $r_2 \in B_1$, $\lim_{T \to \infty} \Pr\left\{ BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} = 1$ provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_{1e}}} \to 0$. This implies that provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T \to \infty} \Pr\left\{ BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} = 1$ for any $r_2 \in B_1$. If $r_2 \in N_1$, $\lim_{T \to \infty} \Pr\left\{ BDF_{r_2} < cv_{r_2}^{\beta_T} \right\} = 1$.

It follows that for any $\eta, \gamma > 0$,

$$
\Pr\{\hat{r}_{1e} > r_{1e} + \eta\} \to 0 \text{ and } \Pr\{\hat{r}_{1f} < r_{1f} - \gamma\} \to 0,
$$

due to the fact that $\Pr\left\{ BDF_{r_{1e}+a_\eta} > cv_{r_2}^{\beta_T} \right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{ BDF_{r_{1f}-a_\gamma} > cv_{r_2}^{\beta_T} \right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary, $\Pr\{\hat{r}_{1e} < r_{1e}\} \to 0$ and $\Pr\{\hat{r}_{1f} > r_{1f}\} \to 0$, we deduce that $\Pr\{|\hat{r}_{1e} - r_{1e}| > \eta\} \to 0$ and $\Pr\{|\hat{r}_{1f} - r_{1f}| > \gamma\} \to 0$ as $T \to \infty$, provided that

$$
\frac{1}{cv_{r_2}^{\beta_T}} + \frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0.
$$

The strategy can consistently estimate r_{1e} and r_{1f}.

Since $\lim_{T \to \infty} \Pr\left\{ BDF_{r_2} < cv_{r_2}^{\beta_T} \right\} = 1$ when $r_2 \in N_1 \cup B_2 \cup N_2$, the strategy cannot estimate r_{2e} and r_{2f} consistently when $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$.

The new date-stamping strategy

The origination of the bubble expansion r_{1e}, r_{2e} and the termination of the bubble collapse r_{1f}, r_{2f} based on the backward sup DF test are identified as

$$
\hat{r}_{1e} = \inf_{r_2 \in [r_0, 1]} \left\{ r_2 : BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\},
$$

63

$$\hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e} + \delta \log(T)/T, 1]} \left\{ r_2 : BSDF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\},$$

$$\hat{r}_{2e} = \inf_{r_2 \in (\hat{r}_{1f}, 1]} \left\{ r_2 : BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\},$$

$$\hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e} + \delta \log(T)/T, 1]} \left\{ r_2 : BSDF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}.$$

We know that when $\beta_T \to 0$, $scv_{r_2}^{\beta_T} \to \infty$.

The asymptotic distributions of the backward sup DF statistic under the alternative hypothesis are (given $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$)

$$BSDF_{r_2}(r_0) \sim_a \begin{cases} F_{r_2}^{r_0}(W) & \text{if } r_2 \in N_0 \\ T^{1/2} \delta_T^{\tau_2 - \tau_{1e}} \sup_{r_1 \in [0, r_2 - r_0]} \left\{ \frac{r_w^{3/2} B(r_{1e})}{2(r_{1e} - r_1) \int_{r_1}^{r_{1e}} B(s) ds} \right\} & \text{if } r_2 \in B_1 \\ -T^{(1-\alpha)/2} \sup_{r_1 \in [0, r_2 - r_0]} \left(\frac{1}{2} cr_w \right)^{1/2} & \text{if } r_2 \in N_1 \\ T^{1/2} \delta_T^{\tau_2 - \tau_{2e}} \sup_{r_1 \in [0, r_2 - r_0]} \left\{ \frac{r_w^{3/2} B(r_{2e})}{2(r_{2e} - r_1) \int_{r_1}^{r_{2e}} B(s) ds} \right\} & \text{if } r_2 \in B_2 \\ -T^{(1-\alpha)/2} \sup_{r_1 \in [0, r_2 - r_0]} \left(\frac{1}{2} cr_w \right)^{1/2} & \text{if } r_2 \in N_2 \end{cases}.$$

It is obvious that if $r_2 \in N_0$,

$$\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} = \Pr \left\{ F_{r_2}^{r_0}(W) = \infty \right\} = 0.$$

If $r_2 \in B_i$ with $i = 1, 2$, $\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} = 1$ provided that $\frac{scv_{r_2}^{\beta_T}}{T^{1/2} \delta_T^{\tau_2 - \tau_{ie}}} \to 0$. This implies that provided that $\frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\} = 1$ for any $r_2 \in B_i$. If $r_2 \in N_i$ with $i = 1, 2$, $\lim_{T \to \infty} \Pr \left\{ BSDF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\} = 1$.

It follows that for any $\eta, \gamma > 0$,

$$\Pr \left\{ \hat{r}_{ie} > r_{ie} + \eta \right\} \to 0 \text{ and } \Pr \left\{ \hat{r}_{if} < r_{if} - \gamma \right\} \to 0,$$

since $\Pr \left\{ BSDF_{r_{ie} + a_\eta}(r_0) > scv_{r_2}^{\beta_T} \right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr \left\{ BSDF_{r_{if} - a_\gamma}(r_0) > scv_{r_2}^{\beta_T} \right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary, $\Pr \left\{ \hat{r}_{ie} < r_{ie} \right\} \to 0$ and $\Pr \left\{ \hat{r}_{if} > r_{if} \right\} \to 0$, we deduce that $\Pr \left\{ |\hat{r}_{ie} - r_{ie}| > \eta \right\} \to 0$ and $\Pr \left\{ |\hat{r}_{if} - r_{if}| > \gamma \right\} \to 0$ as $T \to \infty$, provided that

$$\frac{1}{scv_{r_2}^{\beta_T}} + \frac{scv_{r_2}^{\beta_T}}{T^{1/2}} \to 0.$$

Therefore, the date-stamping strategy based on the generalized sup ADF test can consistently estimate r_{1e}, r_{1f}, r_{2e} and r_{2f}.

Sequential implementation of the date-stamping strategy of PWY

The origination of the bubble expansion r_{1e}, r_{2e} and the termination of the bubble collapse r_{1f}, r_{2f} based on the backward DF test are identified as

$$\hat{r}_{1e} = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BDF_{r_2} > cv_{r_2}^{\beta_T} \right\},$$

$$\hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e} + \log(T)/T, 1]} \left\{ r_2 : BDF_{r_2} < cv_{r_2}^{\beta_T} \right\},$$

$$\hat{r}_{2e} = \inf_{r_2 \in (\hat{r}_{1f} + \varepsilon_T, 1]} \left\{ r_2 :_{\hat{r}_{1f}} BDF_{r_2} > cv_{r_2}^{\beta_T} \right\},$$

$$\hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e} + \log(T)/T, 1]} \left\{ r_2 :_{\hat{r}_{1f}} BDF_{r_2} < cv_{r_2}^{\beta_T} \right\}.$$

where $_{\hat{r}_{1f}} BDF_{r_2}$ is the backward DF statistic calculate over $(\hat{r}_{1f}, r_2]$. We know that when $\beta_T \to 0$, $cv_{r_2}^{\beta_T} \to \infty$.

The asymptotic distributions of the backward DF statistic under the alternative hypothesis are (given $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$)

$$BDF_{r_2} \sim_a \begin{cases} F_{r_2}(W) & \text{if } r_1 \in N_0 \text{ and } r_2 \in N_0 \\ T^{1/2} \delta_T^{\tau_2 - \tau_{1e}} \frac{r_w^{3/2} B(r_{1e})}{2(r_{1e} - r_1) \int_{r_1}^{r_{1e}} B(s)ds} & \text{if } r_1 \in N_0 \text{ and } r_2 \in B_1 \\ -T^{(1-\alpha)/2} \left(\frac{1}{2} cr_w \right)^{1/2} & \text{if } r_1 \in N_0 \text{ and } r_2 \in N_1 \end{cases}$$

and

$$_{\hat{r}_{1f}} BDF_{r_2} \sim_a \begin{cases} F_{r_2}(W) & \text{if } r_2 \in N_1 \\ T^{1/2} \delta_T^{\tau_2 - \tau_{2e}} \frac{r_w^{3/2} B(r_{ie})}{2(r_{ie} - r_1) \int_{r_1}^{r_{ie}} B(s)ds} & \text{if } r_2 \in B_2 \\ -T^{(1-\alpha)/2} \left(\frac{1}{2} cr_w \right)^{1/2} & \text{if } r_2 \in N_2 \end{cases} .$$

It is obvious that if $r_2 \in N_0$,

$$\lim_{T \to \infty} \Pr \left\{ BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} = \Pr \left\{ F_{r_2}(W) = \infty \right\} = 0.$$

If $r_2 \in B_1$, $\lim_{T \to \infty} \Pr \left\{ BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} = 1$ provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2} \delta_T^{\tau_2 - \tau_{1e}}} \to 0$. So, provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T \to \infty} \Pr \left\{ BDF_{r_2} > cv_{r_2}^{\beta_T} \right\} = 1$ for any $r_2 \in B_1$. If $r_2 \in N_1$, $\lim_{T \to \infty} \Pr \left\{ BDF_{r_2} < cv_{r_2}^{\beta_T} \right\} =$

1 and $\lim_{T\to\infty} \Pr\left\{\hat{r}_{1f} BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} = \Pr\left\{F_{r_2}(W) = \infty\right\} = 0$. If $r_2 \in B_2$, $\lim_{T\to\infty} \Pr\left\{\hat{r}_{1f} BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} = 1$ provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_{2e}}} \to 0$. This implies that provided that $\frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T\to\infty} \Pr\left\{\hat{r}_{1f} BDF_{r_2} > cv_{r_2}^{\beta_T}\right\} = 1$ for any $r_2 \in B_2$. If $r_2 \in N_2$, $\lim_{T\to\infty} \Pr\left\{\hat{r}_{1f} BDF_{r_2} < cv_{r_2}^{\beta_T}\right\} = 1$.

It follows that for any $\eta, \gamma > 0$,

$$\Pr\left\{\hat{r}_{1e} > r_{1e} + \eta\right\} \to 0 \text{ and } \Pr\left\{\hat{r}_{1f} < r_{1f} - \gamma\right\} \to 0,$$

since $\Pr\left\{BDF_{r_{1e}+a_\eta} > cv_{r_2}^{\beta_T}\right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{BDF_{r_{1f}-a_\gamma} > cv_{r_2}^{\beta_T}\right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary and $\Pr\left\{\hat{r}_{1e} < r_{1e}\right\} \to 0$ and $\Pr\left\{\hat{r}_{1f} > r_{1f}\right\} \to 0$, we deduce that $\Pr\left\{|\hat{r}_{1e} - r_{1e}| > \eta\right\} \to 0$ and $\Pr\left\{|\hat{r}_{1f} - r_{1f}| > \gamma\right\} \to 0$ as $T \to \infty$, provided that

$$\frac{1}{cv_{r_2}^{\beta_T}} + \frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0.$$

Thus, this date-stamping strategy consistently estimates r_{1e} and r_{1f}.

For any $\phi, \kappa > 0$,

$$\Pr\left\{\hat{r}_{2e} > r_{2e} + \phi\right\} \to 0 \text{ and } \Pr\left\{\hat{r}_{2f} < r_{2f} - \kappa\right\} \to 0,$$

since $\Pr\left\{\hat{r}_{1f} BDF_{r_{2e}+a_\phi} > cv_{r_2}^{\beta_T}\right\} \to 1$ for all $0 < a_\phi < \phi$ and $\Pr\left\{\hat{r}_{1f} BDF_{r_{2f}-a_\kappa} > cv_{r_2}^{\beta_T}\right\} \to 1$ for all $0 < a_\kappa < \kappa$. Since $\phi, \kappa > 0$ is arbitrary, $\Pr\left\{r_{1f} < \hat{r}_{2e} < r_{2e}\right\} \to 0$ and $\Pr\left\{\hat{r}_{2f} > r_{2f}\right\} \to 0$, we deduce that $\Pr\left\{|\hat{r}_{2e} - r_{2e}| > \eta\right\} \to 0$ and $\Pr\left\{|\hat{r}_{2f} - r_{2f}| > \gamma\right\} \to 0$ as $T \to \infty$, provided that

$$\frac{1}{cv_{r_2}^{\beta_T}} + \frac{cv_{r_2}^{\beta_T}}{T^{1/2}} \to 0.$$

Therefore, the alternative sequential implementation of the PWY procedure consistently estimates r_{2e} and r_{2f}.