Testing for Multiple Bubbles 1: Historical Episodes of Exuberance and Collapse in the S&P 500*

Peter C. B. Phillips
Yale University, University of Auckland,
University of Southampton & Singapore Management University

Shu-Ping Shi
The Australian National University

Jun Yu
Singapore Management University

July 28, 2013

Abstract

Recent work on econometric detection mechanisms has shown the effectiveness of recursive procedures in identifying and dating financial bubbles. These procedures are useful as warning alerts in surveillance strategies conducted by central banks and fiscal regulators with real time data. Use of these methods over long historical periods presents a more serious econometric challenge due to the complexity of the nonlinear structure and break mechanisms that are inherent in multiple bubble phenomena within the same sample period. To meet this challenge the present paper develops a new recursive flexible window method that is better suited for practical implementation with long historical time series. The method is a generalized version of the sup ADF test of Phillips, Wu and Yu (2011, PWY) and delivers a consistent date-stamping strategy for the origination and termination of multiple bubbles. Simulations show that the test significantly improves discriminatory power and leads to distinct power gains when multiple bubbles occur. An empirical application of the methodology is conducted on S&P 500 stock market data over a long historical period from January 1871 to December 2010. The new approach successfully identifies the well-known historical episodes of exuberance and collapse over this period, whereas the strategy of PWY and a related CUSUM dating procedure locate far fewer episodes in the same sample range.

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

JEL classification: C15, C22

*The current paper and its technical companion "Testing for Multiple Bubbles 2: Limit Theory of Real Time Detectors" build on work that was originally circulated in 2011 in a long paper entitled "Testing for Multiple Bubbles" accompanied by a 110 page supplement of technical results. We are grateful to Heather Anderson, Farshid Vahid, Tom Smith and Anthony Lynch for valuable discussions. Phillips acknowledges support from the NSF under Grant No. SES 09-56687. Shi acknowledges the Financial Integrity Research Network (FIRN) for funding support.

Peter C.B. Phillips email: peter.phillips@yale.edu. Shuping Shi, email: shuping.shi@anu.edu.au. Jun Yu, email: yujun@smu.edu.sg.

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 recursive method which can detect exuberance in asset price series

[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).

during 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 arises from the complex nonlinear structure involved in the multiple breaks that produce the bubble phenomena. Multiple breaks typically diminish the discriminatory power of existing test mechanisms such as the recursive tests 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 may be multiple bubbles in the data. The mechanisms developed here extend those of PWY by allowing for flexible window widths in the recursive regressions on which the test procedures are based. The approach adopted in PWY uses a sup ADF (SADF) test based on sequence of forward recursive right-tailed ADF unit root tests. This procedure also gives rise to a 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,

[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).

as was first shown in an unpublished working paper by Phillips and Yu (2009) whose results are subsumed as a special case within the present work. Other break testing procedures such as Chow tests, model selection, and CUSUM tests may also be applied as dating mechanisms. Extensive simulations conducted recently by Homm and Breitung (2012) 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.

When the sample period includes multiple episodes of exuberance and collapse, the SADF test may suffer from reduced power and can be inconsistent, thereby 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 and deal with multiple breaks of exuberance and collapse, the present paper proposes an alternative approach named the *generalized sup ADF* (GSADF) test. The GSADF test also relies on recursive right-tailed ADF tests but uses flexible window widths in the implementation. Instead of fixing the starting point of the recursion on the first observation, the GSADF test extends the sample coverage by changing both the starting point and the ending point of the recursion over a feasible range of flexible windows. Since the GSADF test covers more subsamples of the data and has greater window flexibility, it is designed to outperform the SADF test in detecting explosive behavior when multiple episodes occur in the data. This expected enhancement in performance by the GSADF test is demonstrated here in simulations which compare the two tests in terms of their size and power in bubble detection. The new procedure delivers a consistent dating mechanism when multiple bubbles occur, in contrast to the original version of the PWY dating strategy which can be inconsistent when multiple bubbles occur. The technique is therefore well suited to analyzing long historical time series.

A modified version of the original PWY dating algorithm is developed in which the detection procedure is repeated sequentially with re-initialization after the detection of each bubble.

Like the GSADF test, this sequential PWY algorithm works with subsamples of the data with different initializations in the recursions and therefore in theory is capable of detecting multiple bubbles. We also consider detection mechanism based on a recursive CUSUM test suggested recently in Homm and Breitung (2012).

An empirical application of these methodologies is conducted to S&P 500 stock market data over the period January 1871 to December 2010. The new GSADF approach successfully identifies all the well-known historical episodes of exuberance and collapse over this period, including the great crash, the post war boom in 1954, Black Monday in October 1987, the dot-com bubble and the subprime mortgage crisis. Several short episodes are also identified, including the famous banking panic of 1907, and the 1974 stock market crash. The strategy of PWY is much more conservative and locates only two episodes over the same historical period, catching the 1990s stock bubble but entirely missing the 2007-2008 subprime crisis. The sequential PWY algorithm is similarly conservative in detecting bubbles in this data set, as is the CUSUM procedure.

The organization of the paper is as follows. Section 2 discusses reduced form model specification issues for bubble testing, describes the new rolling window recursive test, and gives its limit theory. Section 3 proposes date-stamping strategies based on the new test and outlines their properties in single, multiple and no bubble scenarios. Section 4 reports the results of simulations investigating size, power, and performance characteristics of the various tests and dating strategies. In Section 5, the new procedures, the original PWY test, the sequential PWY test, and the CUSUM test are all applied to the S&P 500 price-dividend ratio data over 1871-2010. Section 6 concludes. Proofs are given in the Appendix. A companion paper (Phillips, Shi and Yu, 2013b) develops the limit theory and consistency properties of the dating procedures of the present paper covering both single and multiple bubble scenarios.

2 A Rolling Window Test for Identifying Bubbles

2.1 Models and Specification

A common starting point in the analysis of financial bubbles is the 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 and B_t satisfies the 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 (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.[3]

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

[3] 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.

of that view in Cooper, 2008[4]). 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 (Phillips and Magdalinos, 2007) behavior in asset prices is a primary indicator of market exuberance during the inflationary phase of a bubble and it is this time series manifestation that may be subjected to econometric testing using recursive testing procedures like the right sided unit root tests in PWY. As discussed above, recursive right sided unit root tests seem to be particularly effective as real time detection mechanisms for mildly explosive behavior and market exuberance.

The PWY test is a reduced form approach to bubble detection. In such tests (as distinct from left sided unit root tests), the focus is usually on the alternative hypothesis (rather than the martingale or unit root hypothesis) because of interest in possible departures from fundamentals and the presence of market excesses or mispricing. Right sided unit root tests, as discussed in PWY, are informative about mildly explosive or submartingale behavior in the data and are therefore useful as a form of market diagnostic or warning alert.

As with all testing procedures, model specification under the null is important for estimation purposes, not least because of the potential impact on asymptotic theory and the critical values used in testing. Unit root testing is a well known example where intercepts, deterministic trends, or trend breaks all materially impact the limit theory. Such issues also arise in right-tailed unit root tests of the type used in bubble detection, as studied recently in Phillips, Shi and Yu (2013a;

[4] "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).

PSY1). Their analysis allowed for a martingale null with an asymptotically negligible drift to capture the mild drift in price processes that are often empirically realistic over long historical periods. The prototypical model of this type has the following weak (local to zero) intercept form

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

where d is a constant, T is the sample size, and the parameter η is a localizing coefficient that controls the magnitude of the intercept and drift as $T \to \infty$. Solving (3) gives $y_t = d\frac{t}{T^\eta} + \sum_{j=1}^{t} \varepsilon_j + y_0$ revealing the deterministic drift dt/T^η. When $\eta > 0$ the drift is small relative to a linear trend, when $\eta > \frac{1}{2}$, the drift is small relative to the martingale component, and when $\eta = \frac{1}{2}$ the standardized output $T^{-1/2} y_t$ behaves asymptotically like a Brownian motion with drift which suits many macroeconomic and financial time series. The null specification (3) includes the pure random walk null of PWY as a special case when $\eta \to \infty$ and the order of magnitude of y_t is then identical to that of a pure random walk. Estimation of the localizing coefficient η is discussed in PSY1.[5]

The model specification (3) is usually complemented with transient dynamics in order to conduct tests for exuberance, just as in standard ADF unit root testing against stationarity. The recursive approach that we now suggest involves a rolling window ADF style regression implementation based on such a system. In particular, suppose the rolling window regression sample starts from the r_1^{th} fraction of the total sample (T) 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 can then be written as

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

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

[5] When $\eta > 0.5$ the drift component is dominated by the stochastic trend and estimates of η typically converge to 1/2, corresponding to the order of the stochastic trend. When $\eta \in [0, \frac{1}{2}]$, the parameter is consistently estimable, although only at a slow logarithmic rate when $\eta = \frac{1}{2}$. See PSY1 for details.

We proceed to use rolling regressions of this type to construct a new approach to bubble detection that is particularly useful in the case of multiple bubbles in the sample. The formulation includes the earlier SADF test procedure developed and used in PWY, which we now briefly review together with some other recursive and regression switching procedures.

2.2 The SADF Test of PWY

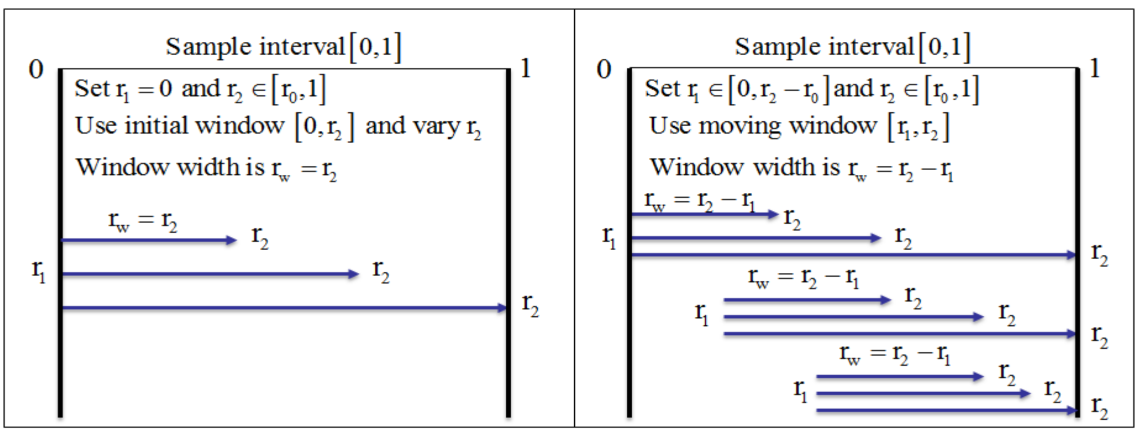
The SADF test relies on repeated estimation of the ADF model on a forward expanding sample sequence and the test is obtained as the sup value of the corresponding ADF statistic sequence. In this case, the window size r_w expands from r_0 to 1, so that r_0 is the smallest sample window width fraction (initializing computation) and 1 is the largest window fraction (the total sample size) in the recursion. The starting point r_1 of the sample sequence is fixed at 0, so the end point of each sample (r_2) equals r_w, and changes 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

$$SADF\left(r_0\right) = \sup_{r_2 \in [r_0, 1]} ADF_0^{r_2}.$$

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 (2012) 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 (2012) 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 (2012) found in their simulations that the PWY test was the most powerful in detecting multiple bubbles, we focus attention in this paper on extending the SADF test. However, our simulations and empirical implementation provide comparative results with the CUSUM procedure in view of its good overall performance recorded in the Homm and Breitung simulations.



(a) SADF Test

(b) GSADF Test

Fig. 1: The sample sequences and window widths of the SADF test and the GSADF test.

2.3 The Rolling Window GSADF Test

The GSADF test developed here continues the idea of repeatedly running the ADF test regression (4) on subsamples of the data in a recursive fashion. However, the subsamples used in the recursion are much more extensive than those of the SADF test. Besides varying the end point of the regression r_2 from r_0 (the minimum window width) to 1, the GSADF test allows the starting point r_1 in (4) to change within a feasible range, i.e. from 0 to $r_2 - r_0$. We define the GSADF statistic to be the largest ADF statistic over all feasible ranges of r_1 and r_2, and we denote this statistic by $GSADF\left(r_0\right)$. That is,

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

Fig. 1 illustrates the comparative sample sequences used in the recursive SADF and GSADF procedures.

Theorem 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) as in (3), the limit distribution of the GSADF test statistic is:*

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

where $r_w=r_2-r_1$ and W is a standard Wiener process. The limit theory continues to hold when the null is a unit root process with asymptotically negligible drift and innovations satisfying the error condition EC in the Appendix.

The proof of Theorem 1 is given in the Appendix. The limit distribution of the GSADF statistic is identical to the case where the regression model includes an intercept and the null hypothesis is a random walk or unit root proces 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 (5) with $r_1=0$ and $r_2=r_w\in[r_0,1]$ (see Phillips, Shi and Yu, 2012).

Similar to the limit theory of 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.[6]

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.

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 the GSADF statistic rises from 1.56 to 2.14 and the 95% finite sample critical value of the test statistic for sample size 400 increases from 1.48 to 2.21. Second, for a given r_0, the finite sample

[6] From Phillips, Shi and Yu (2012), 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.

critical values of the test statistic are almost invariant. Notice that they are very close to the corresponding asymptotic critical values, indicating that the asymptotic critical values may well be used in practical work.[7]

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. Fig. 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 $GSADF\,(0.1)$.

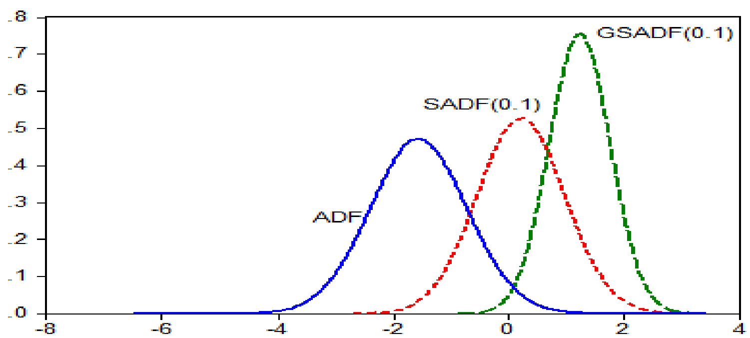


Fig. 2: Asymptotic distributions of the ADF and supADF statistics ($r_0 = 0.1$).

3 Date-stamping Strategies

As discussed in the Introduction, regulators and central banks concerned with practical policy implementation need to assess whether real time data provide evidence of financial exuberance - specifically whether any particular observation such $\tau = \lfloor Tr \rfloor$ belongs to a bubble phase in the overall trajectory. The strategy suggested in PWY is to conduct a right-tailed ADF test using information up to this observation (i.e. information embodied in $I_{\lfloor Tr \rfloor} = \{y_1, y_2, \cdots, y_{\lfloor Tr \rfloor}\}$). Since it is possible that the data $I_{\lfloor Tr \rfloor}$ may include one or more collapsing bubble episodes, the ADF test, like earlier unit root/cointegration-based tests for bubbles (e.g., Diba and Grossman, 1988), may result in finding *pseudo stationary* behavior. The strategy recommended here is to

[7]For accuracy here we use finite sample ciritical values in the simulations and the empirical applications reported below.

perform a *backward sup ADF test* on $I_{\lfloor Tr \rfloor}$ to improve identification accuracy. We use a similar flexible window recursion as that described above.

In particular, the backward SADF test performs a sup ADF test on a backward expanding sample sequence where the end point of each sample is fixed at r_2 and the start point varies from 0 to $r_2 - r_0$. The corresponding ADF statistic sequence is $\left\{ADF_{r_1}^{r_2}\right\}_{r_1 \in [0, r_2 - r_0]}$. The backward SADF statistic is defined as the sup value of the ADF statistic sequence over this interval, viz.,

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

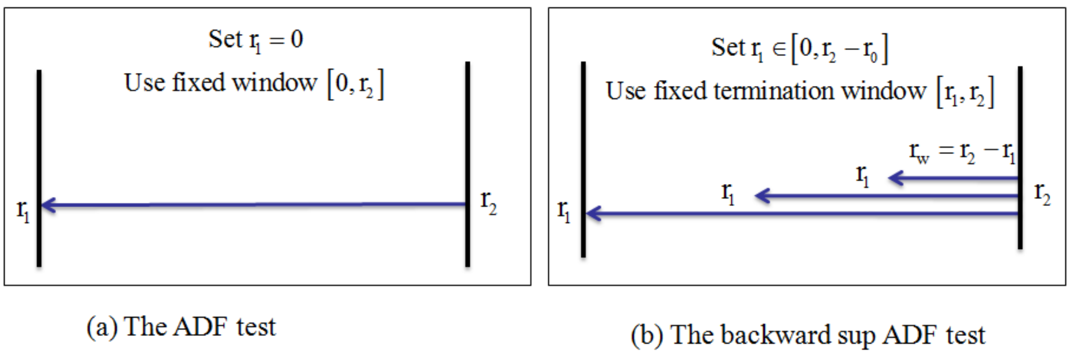


Fig. 3: The sample sequences of the ADF test and the backward SADF test.

The ADF test is a special case of the backward sup ADF test with $r_1 = 0$. We denote the corresponding ADF statistic by ADF_{r_2}. Fig. 3 illustrates the difference between the simple ADF test and the backward SADF test recursion. PWY propose comparing ADF_{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 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 ADF statistic goes below the critical value. PWY impose a condition that for a bubble to exist its duration must exceed a slowly varying (at infinity) quantity such as $L_T = \log(T)$. This requirement helps to exclude short lived blips in the fitted autoregressive coefficient and, as discussed below, can be adjusted to take into account the data frequency. The dating estimates are then delivered by the crossing time formulae

$$\hat{r}_e = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : ADF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_f = \inf_{r_2 \in [\hat{r}_e + \log(T)/T,1]} \left\{ r_2 : ADF_{r_2} < cv_{r_2}^{\beta_T} \right\}, \quad (6)$$

where $cv_{r_2}^{\beta_T}$ is the $100(1-\beta_T)\%$ critical value of the ADF statistic based on $\lfloor Tr_2 \rfloor$ observations. The significance level β_T depends on the sample size T and it is assumed that $\beta_T \to 0$ as $T \to \infty$. This control ensures that $cv_{r_2}^{\beta_T}$ diverges to infinity and thereby eliminates the type I error as $T \to \infty$. In empirical applications, however, β_T will often be fixed at some level such as 0.05 rather than using drifting significance levels.

The new strategy suggests that inference about explosiveness of the process at observation $\lfloor Tr_2 \rfloor$ be 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.[8] 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:

$$\hat{r}_e = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\}, \quad (7)$$

$$\hat{r}_f = \inf_{r_2 \in [\hat{r}_e + \delta \log(T)/T,1]} \left\{ r_2 : BSADF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}, \quad (8)$$

where $scv_{r_2}^{\beta_T}$ is the $100(1-\beta_T)\%$ 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.

[8] 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.

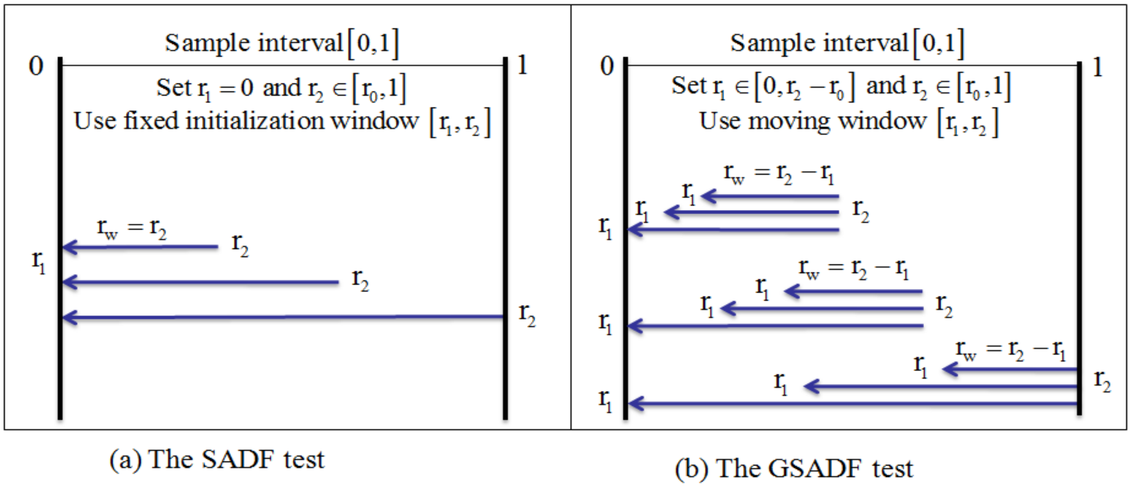


Fig. 4: An alternative illustration of the sample sequences and window widths of the SADF test and the GSADF test.

The SADF test is based on repeated implementation of the ADF test for each $r_2 \in [r_0, 1]$. The GSADF test 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 written as

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

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 Fig. 4.

3.1 Asymptotic properties of the dating algorithms

The limit theory of these date-stamping strategies requires very detailed calculations which are provided in our companion paper (Phillips, Shi andYu, 2013b; PSY2). Additional technical material needed for those derivations is contained in the online supplement to the paper. The main results and import of the theory for empirical practice are reviewed here. We look in turn at cases where there are no bubbles, a single bubble, and multiple bubbles in the data.

No bubbles Under the null hypothesis of the no bubble episodes in the data the asymptotic distributions of the ADF and SADF statistics follow from Theorem 1. 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, from the limit theory given in (5), we have the following asymptotic distributions of these two statistics

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

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

Define cv^{β_T} as the $100(1-\beta_T)\%$ quantile of $F_{r_2}(W)$ and scv^{β_T} as the $100(1-\beta_T)\%$ quantile of $F_{r_2}^{r_0}(W)$. We know that $cv^{\beta_T} \to \infty$ and $scv^{\beta_T} \to \infty$ as $\beta_T \to 0$. Given $cv^{\beta_T} \to \infty$ and $scv^{\beta_T} \to \infty$ 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 $\Pr\{\hat{r}_f \in [r_0, 1]\} \to 0$.

One bubble PSY2 study the consistency properties of the date estimates $\hat{r}_e$ and $\hat{r}_f$ under various alternatives. The simplest is a single bubble episode, like that considered in PWY. The following generating process used in PWY is an effective reduced form mechanism that switches between a martingale mechanism, a single mildly explosive episode, collapse, and subsequent renewal of martingale behavior:

$$X_t = X_{t-1}1\{t < \tau_e\} + \delta_T X_{t-1}1\{\tau_e \le t \le \tau_f\}$$

$$+\left(\sum_{k=\tau_f+1}^{t}\varepsilon_k+X^*_{\tau_f}\right)1\{t>\tau_f\}+\varepsilon_t 1\{j\leq\tau_f\}. \tag{9}$$

In (9) $\delta_T = 1+cT^{-\alpha}$ with $c>0$ and $\alpha\in(0,1)$, $\varepsilon_t \overset{iid}{\sim} (0,\sigma^2)$, $X^*_{\tau_f} = X_{\tau_e}+X^*$ with $X^* = O_p(1)$, $\tau_e = \lfloor Tr_e \rfloor$ dates the origination of bubble expansion and $\tau_f = \lfloor Tr_f \rfloor$ dates the termination of bubble collapse. The pre-bubble period $N_0 = [1,\tau_e)$ is assumed to be a pure random walk process but this is not essential to the asymptotic theory. The bubble expansion period $B = [\tau_e,\tau_f]$ is a mildly explosive process with expansion rate given by the autoregressive (AR) coefficient δ_T. As discussed in PWY, mildly explosive processes are well suited to capturing market exuberance. The process then collapses abruptly to $X^*_{\tau_f}$, which equals X_{τ_e} plus a small perturbation, and continues its random wandering martingale path over the subsequent period $N_1 = (\tau_f,\tau]$. Of course, more general models with various transitional collapse mechanisms can also be considered. The prototypical system (9) captures the main features of interest when there is a single bubble episode and is useful in analyzing test properties for a bubble alternative.

Under (9) and certain rate conditions both the ADF and BSADF detectors provide consistent estimates of the origination and termination dates of the bubble.[9] When the point estimates $\hat{r}_e$ and $\hat{r}_f$ are obtained as in PWY using the ADF test and the first crossing times (6) then $(\hat{r}_e,\hat{r}_f) \xrightarrow{p} (r_e,r_f)$ as $T\to\infty$ provided the following rate condition on the critical value cv^{β_T} holds

$$\frac{1}{cv^{\beta_T}}+\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}}\to 0,\ \text{as}\ T\to\infty. \tag{10}$$

Consistency of $(\hat{r}_e,\hat{r}_f)$ was first proved in a working paper (Phillips and Yu, 2009). When the point estimates $\hat{r}_e$ and $\hat{r}_f$ are obtained from the BSADF detector using the crossing time criteria (7) - (8), we again have consistency $(\hat{r}_e,\hat{r}_f) \xrightarrow{p} (r_e,r_f)$ as $T\to\infty$ under the corresponding rate condition on the critical value scv^{β_T}, viz.,

$$\frac{1}{scv^{\beta_T}}+\frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}}\to 0,\ \text{as}\ T\to\infty. \tag{11}$$

Hence both strategies consistently estimate the origination and termination points when there is only a single bubble episode in the sample period. The rate conditions (10) and (11)

[9] Consistent estimation of the bubble origination date also requires that the minimum window size r_0 not exceed r_e otherwise the recursive regressions do not include r_e and the origination date is not identified.

require for consistency of $(\hat{r}_e, \hat{r}_f)$ that $\left(cv^{\beta_T}, scv^{\beta_T}\right)$ pass to infinity and that their orders of magnitude be smaller than $T^{1/2}\delta_T^{\tau-\tau_e}$. It is sufficient for consistency of $(\hat{r}_e, \hat{r}_f)$ that the critical values cv^{β_T} and scv^{β_T} used in the recursions expand slowly as $T \to \infty$, for example at the slowly varying rate $\log(T)$. The probability of false rejection of normal behavior then goes to zero. The upper rate condition that delimits the rate at which $\left(cv^{\beta_T}, scv^{\beta_T}\right)$ pass to infinity ensures the successful detection of mildly explosive behavior under the alternative. In effect, the critical values used in the crossing times (6) and (7) must not pass to infinity too fast relative to the strength of exuberance in the data which is governed by the value of the localizing parameter $\alpha < 1$ in the AR coefficient $\delta_T = 1 + cT^{-\alpha}$.

Multiple bubbles Multiple bubble episodes may be analyzed in a similar way using more complex alternative models and more detailed calculations, which are reported in PSY2. The key outcomes are revealed in the case of two bubble episodes, which are generated in the following system extending the prototypical model (9):

$$X_t = X_{t-1}1\{t \in N_0\} + \delta_T X_{t-1}1\{t \in B_1 \cup B_2\} + \left(\sum_{k=\tau_{1f}+1}^{t} \varepsilon_k + X^*_{\tau_{1f}}\right) 1\{t \in N_1\}$$
$$+ \left(\sum_{l=\tau_{2f}+1}^{t} \varepsilon_l + X^*_{\tau_{2f}}\right) 1\{t \in N_2\} + \varepsilon_t 1\{j \in N_0 \cup B_1 \cup B_2\}, \tag{12}$$

In (12) we use the notation $N_0 = [1, \tau_{1e}), B_1 = [\tau_{1e}, \tau_{1f}], N_1 = (\tau_{1f}, \tau_{2e}), B_2 = [\tau_{2e}, \tau_{2f}]$ and $N_2 = (\tau_{2f}, \tau]$. The observations $\tau_{1e} = \lfloor Tr_{1e} \rfloor$ and $\tau_{1f} = \lfloor Tr_{1f} \rfloor$ are the origination and termination dates of the first bubble; $\tau_{2e} = \lfloor Tr_{2e} \rfloor$ and $\tau_{2f} = \lfloor Tr_{2f} \rfloor$ 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 resumes a martingale path until time $\tau_{2e} - 1$ and a second episode of exuberance begins at τ_{2e}. The expansion process lasts until τ_{2f} and collapses to a value of $X^*_{\tau_{2f}}$. The process then continues on a martingale path until the end of the sample period τ. The expansion duration of the first bubble is assumed to be longer than that of the second bubble, namely $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$. Obvious extensions of (12) include models where the mildly explosive coefficient δ_T takes different values in regimes B_1 and B_2, and models where the transition mechanisms to martingale behavior over N_1 and N_2 take more graduated and possibly different forms, thereby distinguishing the bubble mechanisms in the two cases.

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):

$$\hat{r}_{1e} = \inf_{r_2 \in [r_0, 1]} \left\{ r_2 : ADF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e} + \log(T)/T, 1]} \left\{ r_2 : ADF_{r_2} < cv_{r_2}^{\beta_T} \right\}, \tag{13}$$

$$\hat{r}_{2e} = \inf_{r_2 \in \left[\hat{r}_{1f}, 1\right]} \left\{ r_2 : ADF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e} + \log(T)/T, 1]} \left\{ r_2 : ADF_{r_2} < cv_{r_2}^{\beta_T} \right\}, \tag{14}$$

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:

$$\hat{r}_{1e} = \inf_{r_2 \in [r_0, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\}, \tag{15}$$

$$\hat{r}_{1f} = \inf_{r_2 \in [\hat{r}_{1e} + \delta \log(T)/T, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}, \tag{16}$$

$$\hat{r}_{2e} = \inf_{r_2 \in \left[\hat{r}_{1f}, 1\right]} \left\{ r_2 : BSADF_{r_2}(r_0) > scv_{r_2}^{\beta_T} \right\}, \tag{17}$$

$$\hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e} + \delta \log(T)/T, 1]} \left\{ r_2 : BSADF_{r_2}(r_0) < scv_{r_2}^{\beta_T} \right\}. \tag{18}$$

An alternative implementation of the PWY procedure is to use that procedure sequentially, namely to detect one bubble at a time and sequentially re-apply the algorithm. The dating criteria for the first bubble remain the same (i.e. equation (13)). Conditional on the first bubble having been found and terminated at $\hat{r}_{1f}$, the following dating criteria are used to date stamp a second bubble:

$$\hat{r}_{2e} = \inf_{r_2 \in \left(\hat{r}_{1f} + \varepsilon_T, 1\right]} \left\{ r_2 :_{\hat{r}_{1f}} ADF_{r_2} > cv_{r_2}^{\beta_T} \right\} \text{ and } \hat{r}_{2f} = \inf_{r_2 \in [\hat{r}_{2e} + \log(T)/T, 1]} \left\{ r_2 :_{\hat{r}_{1f}} ADF_{r_2} < cv_{r_2}^{\beta_T} \right\}, \tag{19}$$

where ${}_{\hat{r}_{1f}}ADF_{r_2}$ is the ADF statistic calculated over $(\hat{r}_{1f}, r_2]$. This sequential application of the PWY procedure requires a few observations in order to re-initialize the test process (i.e. $r_2 \in (\hat{r}_{1f} + \varepsilon_T, 1]$ for some $\varepsilon_T > 0$) after a bubble.

The asymptotic behavior of these various dating estimates is developed in PSY2 and summarized as follows.[10]

(i) The PWY procedure: Under (12) and the rate condition (10) the ADF detector provides consistent estimates $(\hat{r}_{1e}, \hat{r}_{1f}) \xrightarrow{p} (r_{1e}, r_{1f})$ of the origination and termination of the first bubble, but does not detect the second bubble when the duration of the first bubble exceeds that of the second bubble $(\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e})$. If the duration of the first bubble is shorter than the second bubble $\tau_{1f} - \tau_{1e} \leq \tau_{2f} - \tau_{2e}$, then under the rate condition

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1-\alpha/2}} \to 0 \text{ as } T \to \infty, \tag{20}$$

PWY consistently estimates the first bubble and detects the second bubble but with a delay that misdates the bubble – specifically $(\hat{r}_{2e}, \hat{r}_{2f}) \xrightarrow{p} (r_{2e} + r_{1f} - r_{1e}, r_{2f})$.

(ii) The BSADF procedure: Under (12) and the rate condition (11) the BSADF detector provides consistent estimates $(\hat{r}_{1e}, \hat{r}_{1f}, \hat{r}_{2e}, \hat{r}_{2f}) \xrightarrow{p} (r_{1e}, r_{1f}, r_{2e}, r_{2f})$ of the origination and termination points of the first and second bubbles.

(iii) The sequential PWY procedure: Under (12) and the rate condition (10), sequential application (with re-initialization) of the ADF detectpr used in PWY provides consistent estimates $(\hat{r}_{1e}, \hat{r}_{1f}, \hat{r}_{2e}, \hat{r}_{2f}) \xrightarrow{p} (r_{1e}, r_{1f}, r_{2e}, r_{2f})$ of the origination and termination points of the first and second bubbles.

When the sample period includes successive bubble episodes the detection strategy of PWY consistently estimates the origination and termination of the first bubble but does not consistently date stamp the second bubble when the first bubble has longer duration. The new BSADF procedure and repeated implementation (with re-initialization) of the PWY strategy both provide consistent estimates of the origination and termination dates of the two bubbles. PSY2 also examine the consistency properties of the date-stamping strategies when the duration of the first bubble is shorter than the second bubble. In this case, the PWY procedure fails

[10] As mentioned earlier, we also need the condition of $r_0 \leq r_{1e}$ for consistent estimation of the first bubble origination date.

to fully consistently date-stamp the second bubble whereas the new strategy again succeeds in consistently estimating both the origination and termination dates of the two bubbles.

The reason for detection failures in the original PWY procedure lies in the asymptotic behavior of the recursive estimates of the autoregressive coefficient. Under data generating mechanisms such as (12), a recursive estimate $\hat{\delta}_{0,t}$ of $\delta_T = 1 + \frac{c}{T^\alpha}$ that is based on data up to observation $t \in B_2$ is dominated by data over the earlier domain $N_0 \cup B_1 \cup N_1$ and it turns out that $\hat{\delta}_{0,t} \sim 1 - \frac{c}{T^\alpha} < 1$. It follows that right sided unit root tests generally will not detect explosive behavior with such asymptotic behavior in the coefficient estimate. This difficulty is completely avoided by flexible rolling window methods such as the new BSADF test or by repeated use of the original PWY procedure with re-initialization that eliminates the effects of earlier bubble episodes.[11]

4 Simulations

Simulations were conducted to assess the performance of the PWY and sequential PWY procedures as well as the CUSUM approach and the new moving window detection procedure developed in the present paper. We look at size, power, and detection capability for multiple bubble episodes. The data generating process for size comparisons is the null hypothesis in (3) with $d = \eta = 1$. Discriminatory power in detecting bubbles is determined for two different generating models – the Evans (1991) collapsing bubble model (see (21) - (24) below) and an extended version of the PWY bubble model (given by (9) and (12)).

4.1 Size Comparisons

We concentrate on the SADF and GSADF tests. Size is calculated based on the asymptotic critical values displayed in Table 1 using a nominal size of 5%. The number of replications is $5,000$. From Table 2, it is clear that with $k = 0$ (no additional transient dynamic lags in the system), size performance of both tests is satisfactory. We observe that size distortion in the

[11] To consistently estimate the second bubble using PSY and sequential PWY detectors, the minimum window size needs to be small enough to distinguish the different episodes. In particular, r_0 should be less than the distance separating the two bubbles, i.e. $r_0 < r_{2e} - r_{1f}$. See PSY for more discussion.

GSADF test is smaller than that of the SADF test. For example, when $T = 200$ and $r_0 = 0.2$, size distortion of the GSADF test is 0.6% whereas that of the SADF test is 1.2%.

In addition, we explore the effect of fixed and variable transient dynamic lag length selection in the tests, using a fixed lag, BIC order selection, and sequential significance testing (e.g. Campbell and Perron, 1991) with maximum lag 12. First, as evident in Table 2, use of significance testing to determine order leads to non negligible size distortion particularly when the maximum lag length ($k_{\max}$) is large. For instance, when $T = 200, r_0 = 0.2$, and $k_{\max} = 12$, sizes of the SADF and GSADF tests are 0.116 and 0.557 (for a nominal size of 5%), indicating distortion in both tests, particularly GSADF which is vulnerable because of the short sample sizes that arise in the implementation of the flexible window width method. Second, there are downward size distortions for both tests when using a fixed lag order ($k = 3$). Third, BIC provides satisfactory sizes for SADF but less so for GSADF, where size distortion is positive and increases with sample size.

Overall, the magnitude of the size distortion seems smallest when a fixed lag length is used in the recursive tests. The tests are conservative in this case and GSADF is noticeably less affected than SADF. There are advantages to conservative testing because size must go to zero for consistent date stamping of bubbles. So conservative testing helps to reduce the false detection probability. We therefore recommend using fixed lag length methods in the GSADF testing and dating algorithms. This approach is used later in the paper in the empirical application.

4.2 Power Comparisons

4.2.1 Collapsing Bubble Alternatives

We first simulate asset price series based on the Lucas asset pricing model and Evans (1991) collapsing 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 $B_t = 0$

Table 2: 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%.

	$k = 0$	$k = 3$	BIC	Significance Test
$T = 100$ and $r_0 = 0.4$				
SADF	0.043	0.008	0.040	0.115
GSADF	0.048	0.021	0.064	0.378
$T = 200$ and $r_0 = 0.2$				
SADF	0.038	0.007	0.050	0.116
GSADF	0.044	0.024	0.105	0.557
$T = 400$ and $r_0 = 0.1$				
SADF	0.034	0.007	0.056	0.137
GSADF	0.059	0.037	0.131	0.790

Note: size calculations are based on 5000 replications.

for all t to obtain[12]

$$D_t = \mu + D_{t-1} + \varepsilon_{Dt}, \quad \varepsilon_{Dt} \sim N\left(0, \sigma_D^2\right) \tag{21}$$

$$P_t^f = \frac{\mu\rho}{(1-\rho)^2} + \frac{\rho}{1-\rho} D_t, \tag{22}$$

and the Evans bubble component

$$B_{t+1} = \rho^{-1} B_t \varepsilon_{B,t+1}, \qquad \text{if} \;\; B_t < b \tag{23}$$

$$B_{t+1} = \left[\zeta + (\pi\rho)^{-1} \theta_{t+1} \left(B_t - \rho\zeta\right)\right] \varepsilon_{B,t+1}, \qquad \text{if} \;\; B_t \geq b \;. \tag{24}$$

This series has the submartingale property $\mathbb{E}_t\left(B_{t+1}\right) = \left(1 + r_f\right) B_t$. 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 $y_t \sim N\left(0, \tau^2\right)$. 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 (23) - (24) state that a bubble

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

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

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

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 3 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 and dividends from 1871 to 1980. The value for the discount factor ρ is equivalent to a 5% yearly interest rate.

Table 3: 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

In our empirical application of the SADF and GSADF tests to S&P 500 data we use monthly data. Correspondingly in our simulations, the parameters μ and σ_D^2 are set to match the sample mean and sample variance of the first differenced monthly real S&P 500 stock price index and dividend series described in the application section below. The discount value ρ equals 0.985 (we allow ρ to vary from 0.975 to 0.999 in the power comparisons). The new setting is labeled *monthly* in Table 3.

Fig. 5 depicts one realization of the data generating process with the monthly parameter settings. As is apparent in the figure, there are several collapsing episodes of different magnitudes within this particular sample trajectory. Implementation of the SADF and GSADF tests on this particular realization reveals some of the advantages and disadvantages of the two approaches.

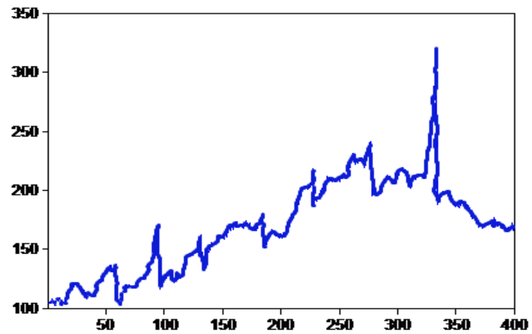


Fig. 5: Simulated time series of $P_t = P_t^f + \kappa B_t$ using the Evans collapsing bubble model (21) - (24) with sample size 400 and monthly parameter settings.

We first implement the SADF test on the whole sample range of this trajectory. We then repeat the test on a sub-sample which contains fewer collapsing episodes. The smallest window size considered in the SADF test for the whole sample contains 40 observations (setting $r_0 = 0.1, T = 400$). The SADF statistic for the full trajectory is 0.71,[13] which is smaller than the 10% finite sample critical value 1.19 (from Table 1). According to this test, therefore, we would conclude that there are no bubbles in the sample.

Next suppose that the SADF test starts from the 201^{st} observation, and the smallest regression window again contains 40 observations (setting $r_0 = 0.2, T = 200$). The SADF statistic obtained from this sample is 1.39[14], which is greater than the 5% finite sample critical value 1.30 (from Table 1). In this case, we reject the null hypothesis of no bubble in the data at the 5% level. These conflicting results point to some instability in the SADF test: evidently the SADF test fails to find bubbles when the full sample is utilized whereas when the sample is truncated to exclude some of the collapse episodes the test succeeds in finding supportive evidence of bubbles.

These two experiments can be viewed as particular (fixed) component runs within the flexible window GSADF test. In the first experiment, the sample starting point of the GSADF test r_1

[13] The SADF statistic is obtained from the subsample regression running from the first observation to the peak of the most significant explosive episode within the sample period (i.e. the $333^{\mathbf{th}}$ observation).

[14] This value comes from the subsample regression starting with the 201^{st} observation up to the 333^{th} observation.

is set to 0, while in the second experiment the sample starting point r_1 is fixed at 0.5. The conflicting results obtained from these two experiments demonstrate the importance of allowing for variable starting points in the implementation of the test, as is done in the GSADF test. When the GSADF procedure is applied to the data, the test statistic value is 8.59,[15] which substantially exceeds the 1% finite sample critical value 2.71 (from Table 1). Thus, the GSADF test finds strong evidence of bubbles in the simulated data. Compared to the SADF test, the GSADF identifies bubbles without having to arbitrarily re-select sample starting points, giving an obvious improvement that is useful in empirical applications.[16]

We now proceed to discuss the results for the full simulation of $R = 5000$ replications. The powers shown in Tables 4 and 5 are calculated using 95% quantiles of the finite sample distributions (Table 1). The smallest window size for both the SADF test and the GSADF test has 40 observations. The data generating process is the periodically collapsing explosive process given in (21) - (24). 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 powers of the GSADF test are respectively 7% and 15.2% higher than those of the SADF test when $T = 100$ and 200.[17]

Table 4: Powers of the SADF and GSADF tests. The data generating process is equation (21)-(24).

	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.

[15] This value is obtained from the subsample regression which covers the most significant expansion period, spanning from the 289^{th} observation to the 333^{th} observation.

[16] Similar phenomena (not reported in detail here) were observed with an alternative data generating process where the logarithmic dividend is a random walk with drift. Parameters in the alternative data generating process (monthly) were 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 $P_t = P_t^f + 500B_t$.

[17] We also considered test results when 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.

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 setting. First, due to the fact that the rate of bubble expansion in this model is inversely related to the discount factor, powers of both the 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% (99.3% to 91.0%, respectively) as the discount factor rises from 0.975 to 0.990. Second, as apparent in Table 5, 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 (21)-(24) 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.

4.2.2 Mildly Explosive Alternatives

We next consider mildly explosive bubble alternatives of the form generated by (9) and (12). These models allow for both single and double bubble scenarios and enable us to compare the finite sample performance of the PWY strategy, the sequential PWY approach, the new dating method and the CUSUM procedure.[18] The CUSUM detector is denoted by $C_{r_0}^r$ and defined as

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

where $\lfloor Tr_0 \rfloor$ is the training sample,[19] $\lfloor Tr \rfloor$ is the monitoring observation, $\hat{\mu}_r$ is the mean of $\left\{\Delta y_1, ..., \Delta y_{\lfloor Tr \rfloor}\right\}$, and $r > r_0$. Under the null hypothesis of a pure random walk, the recursive statistic $C_{r_0}^r$ 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] For the sequential PWY method, we use an automated procedure to re-initialize the process following bubble detection. Specifically, if the PWY strategy identifies a collapse in the market at time t (i.e. $ADF_{t-1} > cv_{t-1}^{0.95}$ and $ADF_t < cv_t^{0.95}$)[21] we re-initialize the test from observation t.

We set the parameters $y_0 = 100$ and $\sigma = 6.79$ so that they match the initial value and the sample standard deviation of the differenced series of the normalized S&P 500 price-dividend ratio described later in our empirical application. The remaining parameters are set to $c = 1, \alpha = 0.6$ and $T = 100$. For the one bubble experiment, we set the duration of the bubble to be 15% of the total sample and let the bubble originate 40% into the sample (i.e. $\tau_f - \tau_e = \lfloor 0.15T \rfloor$ and $\tau_e = \lfloor 0.4T \rfloor$). For the two-bubble experiment, the bubbles originate 20% and 60% into the sample, and the durations are $\lfloor 0.20T \rfloor$ and $\lfloor 0.10T \rfloor$, respectively. Fig. 6 displays typical realizations of these two data generating processes.

---

[18] Simulations in Homm and Breitung (2012) show that the PWY strategy has higher power than other procedures in detecting periodically collapsing bubbles of the Evans (1991) type, the closest rival being the CUSUM procedure.

[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] We impose the additional restriction of successive realizations $ADF_{t+1} < cv_{t+1}^{0.95}$ and $ADF_{t+2} < cv_{t+2}^{0.95}$ to confirm a bubble collapse.

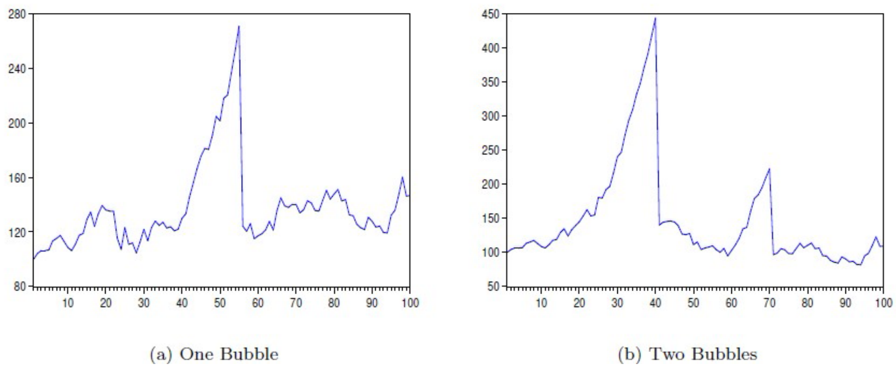


Fig. 6: Typical sample paths generated according to (9) for panel (a) and (12) for panel (b).

We report here a summary of the simulation findings for the main experimental designs. In simulations, we allow τ_e, the bubble location parameter in the single bubble process (13), to be $\lfloor 0.2T \rfloor$, $\lfloor 0.4T \rfloor$ and $\lfloor 0.6T \rfloor$ and the duration of bubbles to vary from $\lfloor 0.10T \rfloor$ to $\lfloor 0.20T \rfloor$. Other parameter configurations were considered and led to broadly similar findings, so are not reported here. For each parameter constellation, 5,000 replications were used. Bubbles were identified using respective finite sample 95% quantiles, obtained from Monte Carlo simulations with 5,000 replications. The minimum window size has 12 observations. We report the empirical mean and standard deviation (in parentheses) of the number of bubbles identified within the sample range, along with the proportions of sample paths identified with bubbles (in squared brackets). Sample paths with no evidence of bubbles found are omitted from the calculation of the mean and the standard deviation. Tables 6 - 7 provide a selection of the results. The main findings are as follows.

1. For the single bubble case, powers of the PWY strategy, the sequential PWY approach and the CUSUM procedure are similar and are slightly lower than that of the new strategy (Table 6).

2. The power of the tests increase with the duration of bubble expansion, shown in square brackets in Table 6, and with the value of the autoregressive coefficient δ_T, although this is not reported here. Hence, bubble detection is more successful when duration of the bubble is longer (and its expansion rate faster).

Table 6: Number of bubbles identified for the single bubble DGP with different bubble durations and locations. Parameters are set as: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, T = 100$. Figures in parentheses and square brackets are standard deviations and powers of the tests, respectively.

	PWY	PSY	Seq	CUSUM
$\boldsymbol{\tau_e = \lfloor 0.2T \rfloor}$				
$\tau_f - \tau_e = \lfloor 0.10T \rfloor$	1.05 (0.27) [0.81]	1.25 (0.51) [0.84]	1.36 (0.62) [0.80]	1.09 (0.34) [0.71]
$\tau_f - \tau_e = \lfloor 0.15T \rfloor$	1.04 (0.20) [0.92]	1.23 (0.50) [0.95]	1.36 (0.61) [0.91]	1.04 (0.24) [0.90]
$\tau_f - \tau_e = \lfloor 0.20T \rfloor$	1.04 (0.20) [0.96]	1.22 (0.48) [0.98]	1.35 (0.61) [0.95]	1.03 (0.20) [0.96]
$\boldsymbol{\tau_e = \lfloor 0.4T \rfloor}$				
$\tau_f - \tau_e = \lfloor 0.10T \rfloor$	1.18 (0.45) [0.73]	1.24 (0.48) [0.81]	1.35 (0.60) [0.73]	1.17 (0.45) [0.76]
$\tau_f - \tau_e = \lfloor 0.15T \rfloor$	1.17 (0.43) [0.88]	1.22 (0.47) [0.91]	1.33 (0.58) [0.88]	1.13 (0.40) [0.91]
$\tau_f - \tau_e = \lfloor 0.20T \rfloor$	1.17 (0.44) [0.94]	1.22 (0.47) [0.96]	1.31 (0.56) [0.93]	1.13 (0.38) [0.96]
$\boldsymbol{\tau_e = \lfloor 0.6T \rfloor}$				
$\tau_f - \tau_e = \lfloor 0.10T \rfloor$	1.32 (0.64) [0.72]	1.26 (0.51) [0.79]	1.37 (0.62) [0.73]	1.26 (0.58) [0.78]
$\tau_f - \tau_e = \lfloor 0.15T \rfloor$	1.30 (0.61) [0.86]	1.25 (0.50) [0.90]	1.32 (0.58) [0.86]	1.23 (0.55) [0.90]
$\tau_f - \tau_e = \lfloor 0.20T \rfloor$	1.30 (0.61) [0.92]	1.26 (0.50) [0.95]	1.28 (0.55) [0.92]	1.23 (0.55) [0.94]

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

3. Bubble location has little impact on the accuracy of the PSY and sequential PWY estimators. But the PWY and CUSUM estimators both become less accurate if the bubble originates at a later stage of the sample period, as shown in parentheses in Table 6. Overall, in the one bubble scenario, the sequential PWY procedure tends to over-estimate the bubble number, the PSY estimator to slightly overestimate bubble number, and the PWY and CUSUM estimators to be more accurate.

4. In the two-bubble scenario, bubble duration can have an especially large impact on the PWY strategy, as is clear in Table 7. When the duration of the first bubble is longer than the second bubble, the mean values of the PWY bubble number estimates are far from the true value and close to one, indicating substantial underestimation. This is consistent with the asymptotic theory which shows that when the duration of the first bubble is longer than the second bubble, the PWY strategy consistently identifies the first bubble but not the second bubble. When the duration of the second bubble is longer than the first

bubble, the bias of the estimate is much smaller. For instance, when $\tau_{1f} - \tau_{1e} = \lfloor 0.1T \rfloor$ and $\tau_{2f} - \tau_{2e} = \lfloor 0.2T \rfloor$, the mean value of the PWY estimates is 1.71, which is much closer to the true value. This simulation finding corroborates the asymptotic theory, which shows that the PWY strategy can detect both bubbles under these conditions.

5. Similar to the weakness of the PWY strategy, when the duration of first bubble is longer than that of the second bubble, the performance of the CUSUM procedure is also biased downwards to selecting a single bubble. Also, like the PWY procedure, there is obvious improvement in the performance of the CUSUM procedure when the second bubble lasts longer (Table 7 final column).

6. As expected, the sequential PWY procedure performs nearly as well as the PSY strategy in the two bubble case but tends to have higher variation and less power than PSY. Estimation accuracy of both estimators improves with the durations of bubbles. (Table 7).

7. Overall, substantially better performance in the two bubble case is delivered by the PSY and sequential PWY estimators, with higher power and much greater accuracy in determining the presence of more than one bubble (Table 7 column 2 and 3).

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 Fig. 6 by the solid (blue) line, which shows the price-dividend ratio over this period to reflect asset prices in relation to fundamentals, according to the pricing equation (1). One might allow also for a time-varying discount factor in that equation. If there were no unobservable component in fundamentals, it follows from the pricing equation that in the absence of bubbles the price-dividend ratio is a function of the discount factor and the dividend growth rate (e.g., Cochrane,

Table 7: Number of bubbles identified for the two-bubble DGP with varying bubble duration. Parameters are set as: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, \tau_{1e} = \lfloor 0.20T \rfloor, \tau_{2e} = \lfloor 0.60T \rfloor, T = 100$. Figures in parentheses and square brackets are standard deviations and powers of the tests, respectively.

	PWY	PSY	Seq	CUSUM
$\boldsymbol{\tau_{1f} - \tau_{1e} = \lfloor 0.10T \rfloor}$				
$\tau_{2f} - \tau_{2e} = \lfloor 0.10T \rfloor$	1.25 (0.46) [0.87]	1.77 (0.57) [0.92]	1.85 (0.63) [0.86]	1.41 (0.53) [0.84]
$\tau_{2f} - \tau_{2e} = \lfloor 0.15T \rfloor$	1.55 (0.53) [0.91]	1.84 (0.54) [0.96]	1.89 (0.57) [0.91]	1.62 (0.52) [0.93]
$\tau_{2f} - \tau_{2e} = \lfloor 0.20T \rfloor$	1.71 (0.49) [0.95]	1.86 (0.52) [0.97]	1.85 (0.53) [0.95]	1.68 (0.50) [0.97]
$\boldsymbol{\tau_{1f} - \tau_{1e} = \lfloor 0.15T \rfloor}$				
$\tau_{2f} - \tau_{2e} = \lfloor 0.10T \rfloor$	1.08 (0.30) [0.94]	1.83 (0.54) [0.97]	1.87 (0.61) [0.94]	1.15 (0.39) [0.93]
$\tau_{2f} - \tau_{2e} = \lfloor 0.15T \rfloor$	1.28 (0.47) [0.95]	1.94 (0.47) [0.98]	1.95 (0.53) [0.96]	1.48 (0.53) [0.96]
$\tau_{2f} - \tau_{2e} = \lfloor 0.20T \rfloor$	1.60 (0.53) [0.97]	1.97 (0.43) [0.99]	1.93 (0.47) [0.97]	1.75 (0.48) [0.98]
$\boldsymbol{\tau_{1f} - \tau_{1e} = \lfloor 0.20T \rfloor}$				
$\tau_{2f} - \tau_{2e} = \lfloor 0.10T \rfloor$	1.05 (0.23) [0.97]	1.83 (0.53) [0.99]	1.89 (0.60) [0.97]	1.06 (0.26) [0.97]
$\tau_{2f} - \tau_{2e} = \lfloor 0.15T \rfloor$	1.10 (0.31) [0.97]	1.97 (0.43) [0.99]	1.96 (0.52) [0.98]	1.16 (0.39) [0.97]
$\tau_{2f} - \tau_{2e} = \lfloor 0.20T \rfloor$	1.29 (0.48) [0.98]	2.01 (0.38) [0.99]	1.95 (0.45) [0.98]	1.46 (0.54) [0.98]

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

1992; Ang and Bekaert, 2006). In such cases, tests for a unit root in the price-dividend ratio do not preclude the presence of a (stationary or nonstationary) time-varying discount factor influencing the ratio.

Table 8: The SADF test and the GSADF test of the S&P500 price-dividend ratio

	Test Stat.	Finite Sample Critical Values		
		90%	95%	99%
SADF	3.30	1.45	1.70	2.17
GSADF	4.21	2.55	2.80	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.

We first apply the summary SADF and GSADF tests to the price-dividend ratio. Table 8 presents critical values for these two tests obtained by simulation with 2,000 replications (sample size 1,680). In performing the ADF regressions and calculating the critical values, the smallest window comprised 36 observations. From Table 8, the SADF and GSADF statistics for the full data series are 3.30 and 4.21, obtained from subsamples 1987M01-2000M07 and 1976M04-1999M06, respectively. 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 summary tests that there is evidence of bubbles in the S&P 500 stock market data. These calculations used a transient dynamic lag order $k = 0$. The findings are robust to other choices. For example, the same conclusion applies when $k = 3$, where the SADF and GSADF tests for the full data series are 2.16 and 3.88 with corresponding 5% critical values of 1.70 and 3.40.

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 simulations with $2,000$ replications. The top panel of Fig. 7 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.

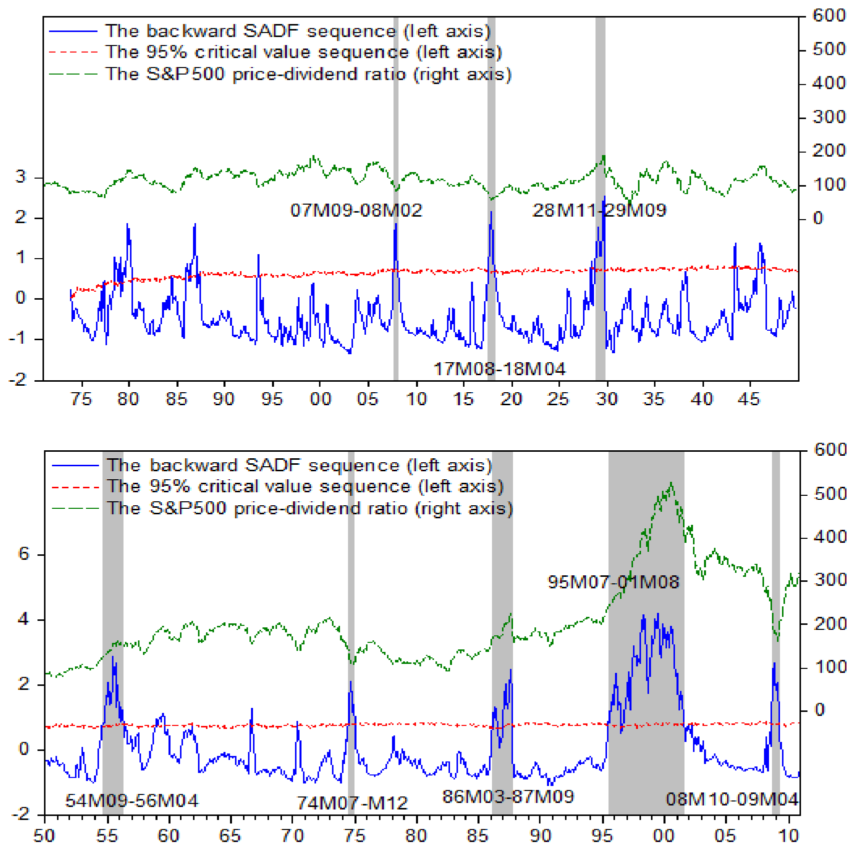


Fig. 7: Date-stamping bubble periods in the S&P 500 price-dividend ratio: the GSADF test.

The identified exuberance and collapse periods after 1900 include *the banking panic of 1907* (1907M09-1908M02), *the 1917 stock market crash* (1917M08-1918M04), *the great crash episode* (1928M11-1929M09), *the postwar boom in 1954* (1954M09-1956M04), *black Monday in October 1987* (1986M03-1987M09), *the dot-com bubble* (1995M07-2001M08) and *the subprime mortgage crisis* (2008M10-2009M04). The durations of those episodes are greater than or equal to half a year. This strategy also identifies several episodes of explosiveness and collapse whose durations are shorter than six months – including *the 1974 stock market crash* (1974M07-M12). Importantly, the new date-stamping strategy not only locates explosive expansion periods but also identifies collapse episodes. Market collapses have occurred in the past when bubbles in other markets crashed and contagion spread to the S&P 500 as occured, for instance, during the *dot-com bubble collapse* and *the subprime mortgage crisis.*

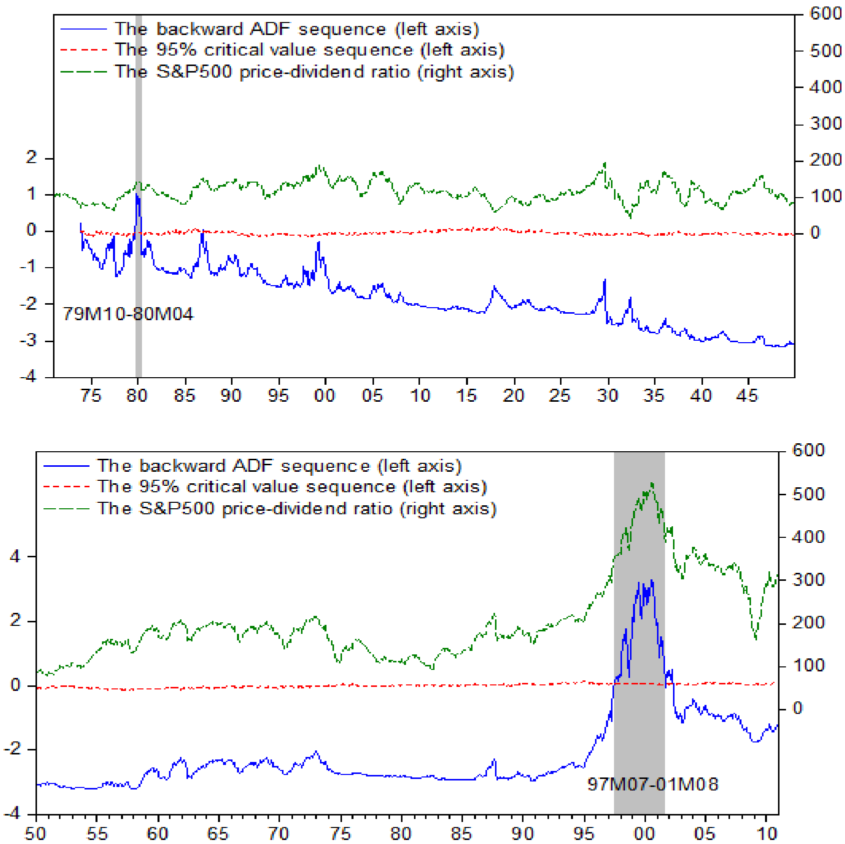


Fig. 8: Date-stamping bubble periods in the S&P 500 price-dividend ratio: the SADF test.

Fig. 8 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). If we restrict the duration of bubbles to be longer than twelve months. The new dating strategy identifies three bubble episodes: *the postwar boom in 1954*, *black Monday in October 1987* and *the dot-com bubble* whereas the strategy of PWY identifies only *the dot-com bubble* in that case.

Empirical results from the sequential PWY procedure are shown in Fig. 9 which plots the ADF statistic sequence against the 95% ADF critical value sequence (as for the PWY dating strategy). As in the simulation exercise (see Section 4.2.2) we use automated re-initialization in the implementation of sequential PWY. A minimum window size $\lfloor r_0 T \rfloor$ is needed to initiate the recursive regression test, so the sequential PWY procedure is unable to perform detection (and hence will fail to identify any bubbles that may occur) over the intervening period $(t, t + \lfloor r_0 T \rfloor)$ following a re-intialization at time t. Futhermore, if the PWY strategy fails to detect a bubble, no re-initialization occurs and the recursive test continues through the sample until a bubble is detected and a subsequent re-initialization is triggered. Hence, the sequential PWY strategy, just like PWY, has some inherent disadvantages in detecting multiple bubbles. In practice, one could potentially pre-divide the sample period into sub-samples for testing but, as shown in the example of Section 4.2.1, the subsample approach may well be sensitive to the pre-selection of the sample periods.

The sequential ADF plot has several breaks in the Figure, each corresponding to the re-initialization of the test procedure following a collapse. The findings from the sequential PWY test indicate two bubbles after 1900 – the *dot-com bubble* (1997M12 - 2002M04) and *the subprime mortgage crisis* (2008M10 -2009M03). Interestingly, after excluding data from the dot-com bubble collapse and earlier data, the sequential PWY strategy successfully identifies an additional episode – *the subprime mortgage crisis* – which the PWY strategy fails to catch (Fig. 8).[22]

[22] If the transient dynamic lag order is $k = 3$, the backward SADF strategy identifies two additional bubble episodes (namely, 1945M12-1946M07 and 1969M11-1970M12). The PWY and sequential PWY strategies identify the same bubble episodes with slight changes in dates.

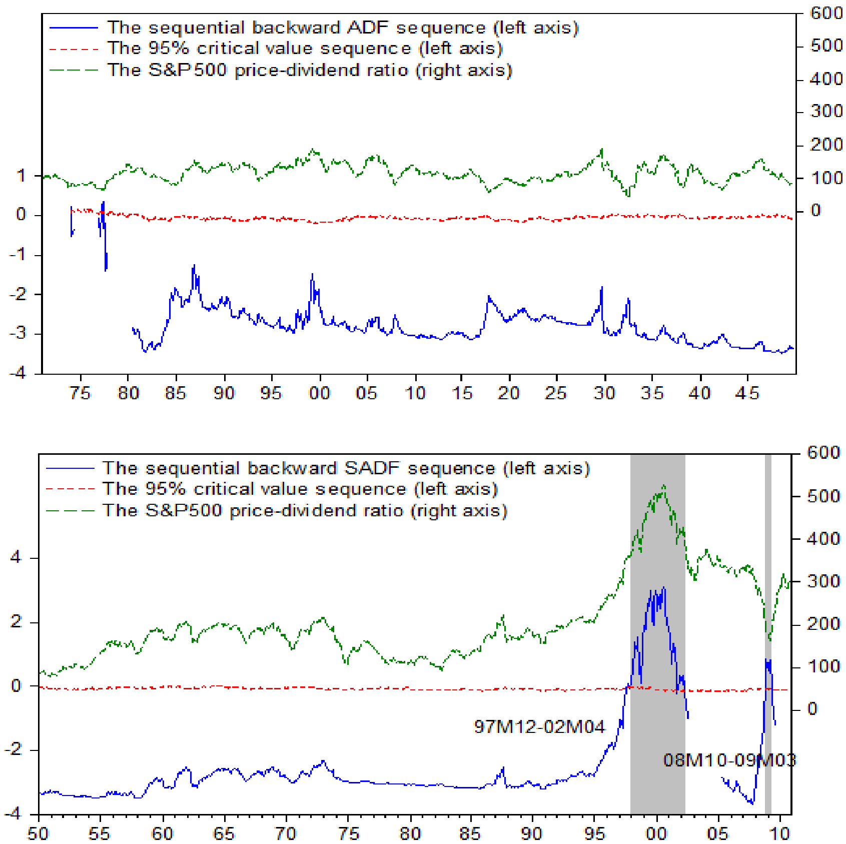


Fig. 9: Date-stamping bubble periods in the S&P 500 price-dividend ratio: the sequential PWY strategy.

For comparison, we applied the CUSUM monitoring procedure to the detrended S&P 500 price-dividend ratio (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. Fig. 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.

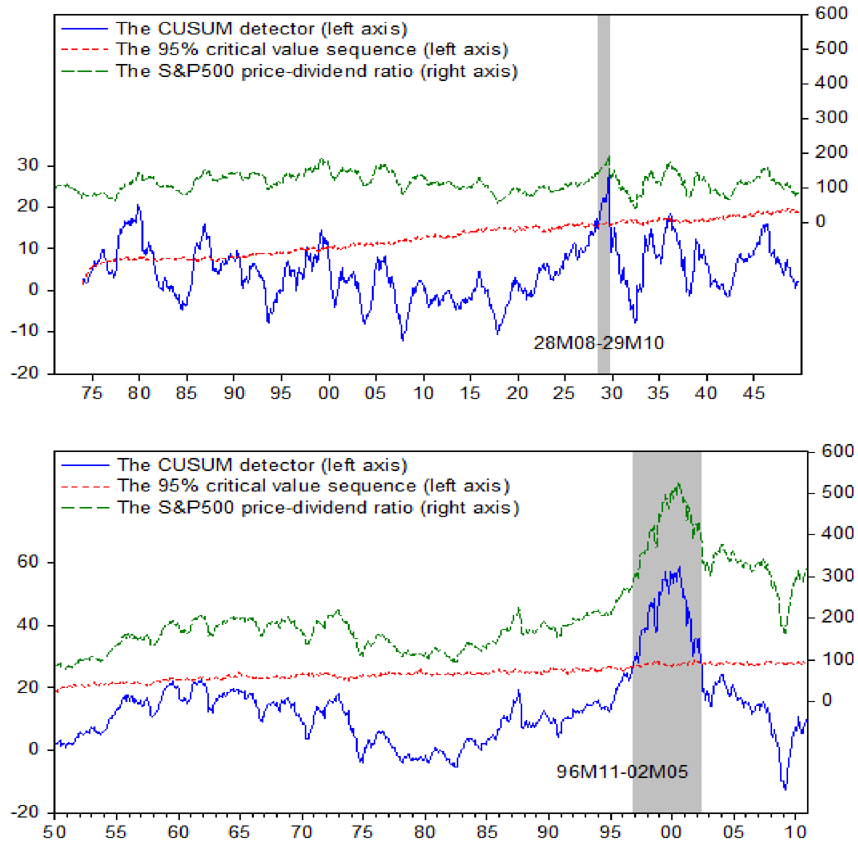


Fig. 10. Date-stamping bubble periods in the S&P 500 price-dividend ratio: the CUSUM monitoring procedure.

As is evident in Fig. 10, 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 for 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.[23]

[23] 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.

6 Conclusion

This paper introduces a new recursive testing procedure and dating algorithm that is useful in detecting multiple bubble events. The GSADF test is a rolling window right-sided ADF unit root test with a double-sup window selection criteria. The reason for the double sup is that the ADF statistic is computed over feasible ranges of the window start points and over a feasible range of window sizes. As distinct from the SADF test of PWY, the window size is selected using the double-sup criteria and the ADF test is implemented 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 limited ability to find and locate bubbles when there are multiple collapsing episodes within the sample range. The GSADF test surmounts this limitation and our simulation findings demonstrate that the GSADF test significantly improves discriminatory power in detecting multiple bubbles. This advantage is particularly important in the empirical study of long historical data series.

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 the strategy of PWY often fails to identify or consistently date stamp the second bubble, whereas the new strategy consistently estimates and dates both bubbles. The PWY dating algorithm may be applied sequentially by re-initializing the detection process after a bubble is found. This sequential application of the PWY dating algorithm has improved asymptotic properties over PWY in the detection of multiple bubbles but both simulations and empirical applications show its performance to be more limited in this capacity.

We apply both SADF and GSADF tests, the sequential PWY dating algorithm, and the CUSUM monitoring procedure, along with their date-stamping algorithms, to the S&P 500 price-dividend ratio from January 1871 to December 2010. All four 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 turns out 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 all other procedures seem more conservative and locate fewer episodes of exuberance and collapse.

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., 1992. Explaining the Variance of Price-Dividend Ratios. *The Review of Financial Studies*, 5(2), 243-280.

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.

Homm, U., and Breitung, J., 2012. Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1), 198-231.

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., 2013a, Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behavior. *Oxford Bulletin of Economics and Statistics*, forthcoming.

Phillips, P.C.B., Shi, S., and Yu, J., 2013b. Testing for Multiple Bubbles 2: Limit Theory of Dating Algorithms. Working Paper.

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., 2012. Specification sensitivities in the Markov-switching unit root test for bubbles. *Empirical Economics*, forthcoming.

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. Asymptotic Distribution of the GSADF test

Before proving Theorem 1 we give conditions on the innovations and state two preliminary lemmas whose proofs follow directly by standard methods (Phillips, 1987; Phillips and Perron, 1988; Phillips and Solo, 1992).

Assumption (EC) *Let $u_t = \psi(L)\varepsilon_t = \Sigma_{j=0}^{\infty}\psi_j\varepsilon_{t-j}$, where $\Sigma_{j=0}^{\infty}j\left|\psi_j\right| < \infty$ and $\{\varepsilon_t\}$ is an i.i.d sequence with mean zero, variance σ^2 and finite fourth moment.*

Lemma 7.1 *Suppose u_t satisfies error condition EC. Define $M_T(r) = 1/T\sum_{s=1}^{[Tr]}u_s$ with $r \in [r_0, 1]$ and $\xi_t = \sum_{s=1}^{t}u_s$. Let $r_2, r_w \in [r_0, 1]$ and $r_1 = r_2 - r_w$. The following hold:*

(1) $\sum_{s=1}^{t}u_s = \psi(1)\sum_{s=1}^{t}\varepsilon_s + \eta_t - \eta_0$, *where* $\eta_t = \sum_{j=0}^{\infty}\alpha_j\varepsilon_{t-j}$, $\eta_0 = \sum_{j=0}^{\infty}\alpha_j\varepsilon_{-j}$ *and* $\alpha_j = -\sum_{i=1}^{\infty}\psi_{j+i}$, *which is absolutely summable.*

(2) $\frac{1}{T}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\varepsilon_t^2 \xrightarrow{p} \sigma^2 r_w$.

(3) $T^{-1/2}\sum_{t=1}^{[Tr]}\varepsilon_t \xrightarrow{L} \sigma W(r)$.

(4) $T^{-1}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\sum_{s=1}^{t-1}\varepsilon_s\varepsilon_t \xrightarrow{L} \frac{1}{2}\sigma^2\left[W(r_2)^2 - W(r_1)^2 - r_w\right]$.

(5) $T^{-3/2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\varepsilon_t t \xrightarrow{L} \sigma\left[r_2 W(r_2) - r_1 W(r_1) - \int_{r_1}^{r_2}W(s)\,ds\right]$.

(6) $T^{-1}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\left(\eta_{t-1} - \eta_0\right)\varepsilon_t \xrightarrow{p} 0$.

(7) $T^{-1/2}\left(\eta_{[Tr]} - \eta_0\right) \xrightarrow{p} 0$.

(8) $\sqrt{T}M_T(r) \xrightarrow{L} \psi(1)\sigma W(r)$.

(9) $T^{-3/2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\xi_{t-1} \xrightarrow{L} \psi(1)\sigma\int_{r_1}^{r_2}W(s)\,ds$.

(10) $T^{-5/2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\xi_{t-1}t \xrightarrow{L} \psi(1)\sigma\int_{r_1}^{r_2}W(s)\,s ds$.

(11) $T^{-2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\xi_{t-1}^2 \xrightarrow{L} \sigma^2\psi(1)^2\int_{r_1}^{r_2}W(s)^2\,ds$.

(12) $T^{-3/2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}\xi_t u_{t-j} \xrightarrow{p} 0, \forall j \geq 0$.

Lemma 7.2 *Define $y_t = \tilde{\alpha}_T t + \sum_{s=1}^{t}u_s$, $\tilde{\alpha}_T = \tilde{\alpha}\psi(1)T^{-\eta}$ with $\eta > 1/2$ and let u_t satisfy condition EC. Then*

$$(a)\;\; T^{-1}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}y_{t-1}\varepsilon_t \xrightarrow{L} \frac{1}{2}\sigma^2\psi(1)\left[W(r_2)^2 - W(r_1)^2 - r_w\right].$$

$$(b)\;\; T^{-3/2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor}y_{t-1} \xrightarrow{L} \psi(1)\sigma\int_{r_1}^{r_2}W(s)\,ds.$$

$$(c)\ \ T^{-2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} y_{t-1}^2 \xrightarrow{L} \sigma^2\psi(1)^2\int_{r_1}^{r_2} W(s)^2\, ds.$$

$$(d)\ \ T^{-3/2}\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} y_{t-1}u_{t-j} \xrightarrow{p} 0, , j = 0, 1, \cdots .$$

Proof of Theorem 1. The regression model is

$$\Delta y_t = \alpha_{r_1,r_2} + \beta_{r_1,r_2} y_{t-1} + \sum_{k=1}^{p-1} \phi_{r_1,r_2}^k \Delta y_{t-k} + \varepsilon_t.$$

Under the null hypothesis that $\alpha_{r_1,r_2} = \tilde{\alpha}T^{-\eta}$ and $\beta_{r_1,r_2} = 0$, we have $y_t = \tilde{\alpha}_T t + \sum_{s=1}^t u_s$ and $\triangle y_t = \tilde{\alpha}_T + u_t$, where $\tilde{\alpha}_T = \psi_{r_1,r_2}(1)\,\alpha_{r_1,r_2}$ and $u_t = \psi_{r_1,r_2}(1)\,\varepsilon_t$ with $\psi_{r_1,r_2}(1) = \left(1 - \phi_{r_1,r_2}^1 L - \phi_{r_1,r_2}^2 L^2 - \cdots - \phi_{r_1,r_2}^{p-1} L^{p-1}\right)^{-1}$.

The deviation of the OLS estimate $\hat{\theta}_{r_1,r_2}$ from the true value θ_{r_1,r_2} is given by

$$\hat{\theta}_{r_1,r_2} - \theta_{r_1,r_2} = \left[\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t X_t'\right]^{-1}\left[\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t\varepsilon_t\right], \qquad (25)$$

where $X_t = [\tilde{\alpha}_T + u_{t-1}\ \tilde{\alpha}_T + u_{t-2} \ldots \tilde{\alpha}_T + u_{t-p+1}\ 1\ y_{t-1}]'$, $\theta = [\phi_{r_1,r_2}^1\ \phi_{r_1,r_2}^2 \ldots \phi_{r_1,r_2}^{p-1}\ \alpha_{r_1,r_2}\ \beta_{r_1,r_2}]'$. The probability limit of $\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t X_t'$ is block diagonal from (d) of Lemma 7.2. Therefore, we only need to obtain the last 2×2 components of $\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t X_t'$ and the last 2×1 component of $\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t\varepsilon_t$ to calculate the ADF statistics, which are

$$\begin{bmatrix} \Sigma' 1 & \Sigma y_{t-1} \\ \Sigma' y_{t-1} & \Sigma y_{t-1}^2 \end{bmatrix} \text{ and } \begin{bmatrix} \Sigma'\varepsilon_t \\ \Sigma' y_{t-1}\varepsilon_t \end{bmatrix},$$

respectively, where Σ' denotes summation over $t = \lfloor Tr_1\rfloor, \lfloor Tr_1\rfloor + 1, \cdots, \lfloor Tr_2\rfloor$. Based on (3) in Lemma 7.1 and (a) in Lemma 7.2, the scaling matrix should be $\Upsilon_T = diag\left(\sqrt{T}, T\right)$. Premultiplying equation (25) by Υ_T, results in

$$\Upsilon_T\begin{bmatrix} \widehat{\alpha}_{r_1,r_2} - \alpha_{r_1,r_2} \\ \widehat{\beta}_{r_1,r_2} - \beta_{r_1,r_2} \end{bmatrix} = \left\{\Upsilon_T^{-1}\left[\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t X_t'\right]_{(-2)\times(-2)} \Upsilon_T^{-1}\right\}^{-1}\left\{\Upsilon_T^{-1}\left[\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t\varepsilon_t\right]_{(-2)\times 1}\right\}.$$

Consider the matrix $\Upsilon_T^{-1}\left[\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t X_t'\right]_{(-2)\times(-2)} \Upsilon_T^{-1}$, whose partitioned form is

$$\begin{bmatrix} \sqrt{T} & 0 \\ 0 & T \end{bmatrix}^{-1}\begin{bmatrix} \Sigma' 1 & \Sigma' y_{t-1} \\ \Sigma' y_{t-1} & \Sigma' y_{t-1}^2 \end{bmatrix}\begin{bmatrix} \sqrt{T} & 0 \\ 0 & T \end{bmatrix}^{-1} = \begin{bmatrix} T^{-1}\Sigma' 1 & T^{-3/2}\Sigma' y_{t-1} \\ T^{-3/2}\Sigma' y_{t-1} & T^{-2}\Sigma' y_{t-1}^2 \end{bmatrix}$$

$$\xrightarrow{L} \begin{bmatrix} r_w & \psi_{r_1,r_2}(1)\,\sigma_{r_1,r_2}\int_{r_1}^{r_2} W(s)\,ds \\ \psi_{r_1,r_2}(1)\,\sigma_{r_1,r_2}\int_{r_1}^{r_2} W(s)\,ds & \sigma^2_{r_1,r_2}\psi_{r_1,r_2}(1)^2\int_{r_1}^{r_2} W(s)^2\,ds \end{bmatrix}$$

and the matrix $\Upsilon_T^{-1}\left[\sum_{t=\lfloor Tr_1\rfloor}^{\lfloor Tr_2\rfloor} X_t\varepsilon_t\right]_{(-2)\times 1}$, for which

$$\begin{bmatrix} T^{-1/2}\Sigma\varepsilon_t \\ T^{-1}\Sigma y_{t-1}\varepsilon_t \end{bmatrix} \xrightarrow{L} \begin{bmatrix} \sigma_{r_1,r_2}\left[W(r_2)-W(r_1)\right] \\ \frac{1}{2}\sigma^2_{r_1,r_2}\psi_{r_1,r_2}(1)\left[W(r_2)^2-W(r_1)^2-r_w\right] \end{bmatrix}.$$

Under the null hypothesis that $\alpha_{r_1,r_2}=T^{-\eta}$ and $\beta_{r_1,r_2}=0$,

$$\begin{bmatrix} \sqrt{T}\left(\hat{\alpha}_{r_1,r_2}-\alpha_{r_1,r_2}\right) \\ T\hat{\beta}_{r_1,r_2} \end{bmatrix} \xrightarrow{L} \begin{bmatrix} r_w & A_{r_1,r_2} \\ A_{r_1,r_2} & B_{r_1,r_2} \end{bmatrix}^{-1} \begin{bmatrix} C_{r_1,r_2} \\ D_{r_1,r_2} \end{bmatrix},$$

where

$$\begin{aligned} A_{r_1,r_2} &= \psi_{r_1,r_2}(1)\,\sigma_{r_1,r_2}\int_{r_1}^{r_2} W(s)\,ds, \\ B_{r_1,r_2} &= \sigma^2_{r_1,r_2}\psi_{r_1,r_2}(1)^2\int_{r_1}^{r_2} W(s)^2\,ds, \\ C_{r_1,r_2} &= \sigma_{r_1,r_2}\left[W(r_2)-W(r_1)\right], \\ D_{r_1,r_2} &= \frac{1}{2}\sigma^2_{r_1,r_2}\psi_{r_1,r_2}(1)\left[W(r_2)^2-W(r_1)^2-r_w\right]. \end{aligned}$$

Therefore, $\hat{\beta}_{r_1,r_2}$ converges at rate T to the following limit variate

$$T\hat{\beta}_{r_1,r_2} \xrightarrow{L} \frac{A_{r_1,r_2}C_{r_1,r_2}-r_wD_{r_1,r_2}}{A^2_{r_1,r_2}-r_wB_{r_1,r_2}}.$$

To calculate the t-statistic $t_{r_1,r_2}=\frac{\hat{\beta}_{r_1,r_2}}{se\left(\hat{\beta}_{r_1,r_2}\right)}$ of $\hat{\beta}_{r_1,r_2}$, we first find the standard error $se\left(\hat{\beta}_{r_1,r_2}\right)$. We have

$$var\left(\begin{bmatrix} \hat{\alpha}_{r_1,r_2} \\ \hat{\beta}_{r_1,r_2} \end{bmatrix}\right) = \sigma^2_{r_1,r_2}\begin{bmatrix} \Sigma'1 & \Sigma'y_{t-1} \\ \Sigma'y_{t-1} & \Sigma'y^2_{t-1} \end{bmatrix}^{-1},$$

so the variance of $T\hat{\beta}_{r_1,r_2}$ can be calculated from

$$\begin{aligned} &var\left(\begin{bmatrix} \sqrt{T}\left(\hat{\alpha}_{r_1,r_2}-\alpha_{r_1,r_2}\right) \\ T\hat{\beta}_{r_1,r_2} \end{bmatrix}\right) \\ &= \sigma^2_{r_1,r_2}\left\{\begin{bmatrix} \sqrt{T} & 0 \\ 0 & T \end{bmatrix}^{-1}\begin{bmatrix} \Sigma'1 & \Sigma'y_{t-1} \\ \Sigma'y_{t-1} & \Sigma'y^2_{t-1} \end{bmatrix}\begin{bmatrix} \sqrt{T} & 0 \\ 0 & T \end{bmatrix}^{-1}\right\}^{-1} \end{aligned}$$

$$= \sigma^2_{r_1,r_2} \left[\begin{array}{cc} T^{-1}\Sigma'1 & T^{-3/2}\Sigma' y_{t-1} \\ T^{-3/2}\Sigma' y_{t-1} & T^{-2}\Sigma' y^2_{t-1} \end{array} \right]^{-1} \xrightarrow{L} \sigma^2_{r_1,r_2} \left[\begin{array}{cc} r_w & A_{r_1,r_2} \\ A_{r_1,r_2} & B_{r_1,r_2} \end{array} \right]^{-1}.$$

It follows that the t-statistic t_{r_1,r_2} of $\hat{\beta}_{r_1,r_2}$ satisfies

$$t_{r_1,r_2} \xrightarrow{L} \frac{\frac{1}{2} r_w \left[W(r_2)^2 - W(r_1)^2 - r_w \right] - \int_{r_1}^{r_2} W(s)\, ds \left[W(r_2) - W(r_1) \right]}{r_w^{1/2} \left\{ r_w \int_{r_1}^{r_2} W(s)^2\, ds - \left[\int_{r_1}^{r_2} W(s)\, ds \right]^2 \right\}^{1/2}}.$$

By continuous mapping the asymptotic distribution of the GSADF statistic is

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

giving the stated result. ■

Testing for Multiple Bubbles 2: Limit Theory of Real Time Detectors*

Peter C. B. Phillips
Yale University, University of Auckland,
University of Southampton & Singapore Management University

Shu-Ping Shi
The Australian National University

Jun Yu
Singapore Management University

July 28, 2013

Abstract

This paper provides the limit theory of real time dating algorithms for bubble detection that were suggested in Phillips, Wu and Yu (2011, PWY) and Phillips, Shi and Yu (2013b, PSY). Bubbles are modeled using mildly explosive bubble episodes that are embedded within longer periods where the data evolves as a stochastic trend, thereby capturing normal market behavior as well as exuberance and collapse. Both the PWY and PSY estimates rely on recursive right tailed unit root tests (each with a different recursive algorithm) that may be used in real time to locate the origination and collapse dates of bubbles. Under certain explicit conditions, the moving window detector of PSY is shown to be a consistent dating algorithm even in the presence of multiple bubbles. The other algorithms are consistent detectors for bubbles early in the sample and, under stronger conditions, for subsequent bubbles in some cases. These asymptotic results and accompanying simulations guide the practical implementation of the procedures. They indicate that the PSY moving window detector is more reliable than the PWY strategy, sequential application of the PWY procedure and the CUSUM procedure.

Keywords: Bubble duration, Consistency, Dating algorithm, Limit theory, Multiple bubbles, Real time detector.

JEL classification: C15, C22

*The current paper and its empirical companion "Testing for Multiple Bubbles 1: Historical Episodes of Exuberance and Collapse in the S&P 500" build on work that was originally circulated in 2011 in a long paper entitled "Testing for Multiple Bubbles" and a supplement of related technical results. We are grateful to Heather Anderson, Farshid Vahid, Tom Smith and Anthony Lynch for valuable discussions. Phillips acknowledges support from the NSF under Grant Nos. SES 09-56687 and SES 12-58258. Shi acknowledges the Financial Integrity Research Network (FIRN) for funding support. Yu acknowledges the financial support from Singapore Ministry of Education Academic Research Fund Tier 2 under the grant number MOE2011-T2-2-096. Peter C.B. Phillips email: peter.phillips@yale.edu. Shuping Shi, email: shuping.shi@anu.edu.au. Jun Yu, email: yujun@smu.edu.sg.

1 Introduction

A recent article by Phillips, Wu and Yu (2011, PWY) developed new econometric methodology for real time bubble detection. When it was applied to Nasdaq data in the 1990s, the algorithm revealed that evidence in the data supported Greenspan's declaration of 'irrational exuberance' in December 1996 and that this evidence of market exuberance had existed for some 16 months prior to that declaration. Greenspan's remark therefore amounted to an assertion that could have been evidence-based if the test had been conducted at the time.

Greenspan formulated his comment as a question: "How do we know when irrational exuberance has unduly escalated asset values?" It was this very question that the recursive test procedure in PWY was designed to address. Correspondingly, an element of the methodology that is critical for empirical applications and policy assessment is the consistency of the test. Ideally we want a test whose size goes to zero and whose power goes to unity as the sample size passes to infinity. Then in very large samples there will be no false positive declarations of exuberance and no false negative assessments where asset price bubbles are missed.

PWY gave heuristic arguments showing that their recursive methodology produced a consistent test for exuberance and they provided a real time dating algorithm for finding the bubble origination and termination dates that was used in analyzing the Nasdaq data. The present paper provides a rigorous limit theory showing formal test consistency of the PWY bubble detection procedure and the consistency of its associated dating algorithm under certain conditions, notably the existence of a single bubble period in the data.[1] This limit theory is part of a much larger formal investigation undertaken here which examines the asymptotic properties of bubble detection algorithms when there may be multiple episodes of exuberance in the data, under which the PWY procedure does not perform nearly as well. As argued in our companion paper Phillips, Shi and Yu (2013b, PSY), data over long historical periods often include several crises involving financial exuberance and collapse. Bubble detection in this context of multiple episodes of exuberance and collapse is much more complex and is the main subject of the

[1] The present paper therefore subsumes the results contained in the unpublished working paper of Phillips and Yu (2009) which is referenced in PWY and which first analyzed the asymptotic properties of the PWY procedure.

PSY paper, which develops a new moving window bubble detector that has some substantial advantages for long data series characterised by multiple financial crises.

The dating algorithms of PWY and PSY are now being applied to a wide range of markets that include energy, real estate, and commodities, as well as financial assets[2]. This methodology and its various applications have also attracted the attention of central bank economists, fiscal regulators, and the financial press.[3] It is therefore important that the limit properties and performance characteristics of these dating algorithms be well understood to assist in guiding practitioners about the suitable choice of procedures for implementation in empirical work and policy assessment exercises.

The PWY and PSY strategies for bubble detection and the estimation of any bubble origination and termination dates involve the comparison of a sequence of recursive test statistics with a corresponding critical value sequence, the crossing times of the lines being used to provide the date estimates. The PWY procedure uses recursively calculated right sided unit root test statistics based on a full sample of observations up to the current data point, whereas PSY use a moving window recursion of sup statistics based on a sequence of right sided unit root tests calculated over flexible windows of varying length taken up to the current data point. Inferences from the PWY and PSY strategies about the presence of exuberance in the data, including the dating of any exuberance or collapse, are drawn from these test sequences and the corresponding critical value sequences. The goals of the present paper are to explore the asymptotic and finite sample properties of these two procedures and to build a methodology for analyzing real time detector asymptotics in this context.

Our findings can be summarized as follows. First, under some general conditions both the PWY and PSY detectors are consistent when there is a single bubble in the sample period.

[2]See Phillips and Yu (2011b), Das et al. (2011), Homm and Breitung (2012), Gutierrez (2013), Bohl et al. (2013), Etienne et al. (2013), among others.

[3]For example, a *Financial Times* article (Meyer, 2013) reports the work of Etienne et al. (2013) which employs the PSY dating algorithm to identify agricultural commodity bubbles. Recent working papers from the Hong Kong Monetary Authority (Yiu et al, 2012) and the Central Bank of Colombia (Gómez-González, et al, 2013) use PSY in studying real estate bubbles in Hong Kong and Columbia. Work for UNCTAD by Gilbert (2010) applies PWY to date bubbles in commodity prices and test congressional testimony reasoning by Masters (2008), and recent financial press articles (Phillips and Yu, 2011a, 2013) use PWY to assess current real estate and world stock market data for evidence of bubbles.

Second, when there are two bubbles in the sample period, the PWY detector for the first bubble is consistent, whereas the PWY estimates associated with the second bubble are duration-dependent. Specifically, the PWY strategy fails to detect the existence of the second bubble (and hence cannot provide consistent date estimates for the timing of that bubble) when the first bubble has longer duration than the second. But when the duration of the second bubble exceeds the first, the PWY strategy can detect the second bubble but only with some delay. Third, the PSY strategy and (under additional conditions) a sequential implementation of the PWY strategy (to each individual bubble in turn) do provide consistent detectors for both bubbles and these results hold irrespective of bubble duration. Thus, the PSY dating algorithm and sequential application of the PWY procedure have desirable asymptotic properties in a multiple bubbles scenario. One disadvantage of sequentially applying the PWY procedure is that sufficient data is needed between bubbles to implement the procedure and therefore some origination dates may not be consistently estimated if the origination date is excluded from the PWY sample recursion.

The paper also reports simulations to evaluate the finite sample performance of these detectors and date estimators, along with an alternative procedure based on CUSUM tests, as proposed in recent work by Homm and Breitung (2012). The simulation results strongly corroborate the asymptotic theory, indicating that the PSY detector is much more reliable than PWY. On the other hand and with some exceptions that will be discussed in detail below, sequential application of the PWY procedure may perform nearly as well as the PSY algorithm. The performance characteristics of the CUSUM procedure are found to be similar to those of PWY. Overall, the results suggest that the PSY detector is a preferred procedure for practical implementation, especially with long data series involving more than one crisis episode.

The rest of the paper is organized as follows. Section 2 introduces the date stamping procedures that use recursive regressions and right tailed unit root tests of the type considered in PWY and PSY. This section also describes the models used to capture mildly explosive bubble behaviour when there are single and multiple bubble episodes in the data. Section 3 derives the limit theory for the dating procedures under both single bubble and multiple bubble alternatives. Finite sample performance is studied in Section 4 and Section 5 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, 2013c) provides a complete set of additional mathematical derivations that are needed for the limit theory presented here.

2 Bubble Dating Algorithms

This Section introduces three different dating algorithms – the original PWY detector, the PSY detector, and a sequential version of the PWY detector. The approach in all of these algorithms is to use recursive right tailed unit root tests to assess evidence for mildly explosive bubble behaviour. In what follows we use the same models, tests, and notation as the companion paper PSY to assist joint reading of the two papers.

The null hypothesis is specified as suggested in Phillips, Shi and Yu (2013a): a random walk (or more generally a martingale) process with an asymptotically negligible drift which we write in the form

$$X_t = kT^{-\eta} + X_{t-1} + \varepsilon_t, \text{ with constant } k \text{ and } \eta > 1/2, \tag{1}$$

where T is the sample size, $\varepsilon_t \overset{i.i.d}{\sim} \left(0, \sigma^2\right)$ and $X_0 = O_p(1)$. Under these simple conditions, partial sums of ε_t satisfy the functional law

$$T^{-1/2} \sum_{t=1}^{\lfloor T\cdot \rfloor} \varepsilon_t \Rightarrow B(\cdot) := \sigma W(\cdot), \tag{2}$$

where W is standard Brownian motion. The framework can be extended to allow for martingale difference sequence and more general weakly dependent innovations under conditions that allow the limit theory to continue to hold under the null (1), based on the functional law (2), and under mildly explosive alternatives as in (4) below, the latter based on results in Phillips and Magdalinos (2007a, 2007b). We maintain the *iid* error assumption here to keep the exposition as simple as possible and the paper to manageable length.

The fitted regression model is

$$\Delta X_t = \alpha + \beta X_{t-1} + \varepsilon_t, \;\; \varepsilon_t \overset{i.i.d}{\sim} \left(0, \sigma^2\right), \tag{3}$$

which includes an intercept but no time trend. As in PSY, the fitted model may also be formulated in ADF regression format to allow for any short memory dependence in the innovations. The results given below continue to hold in that event but full extension to this case will substantially complicate derivations that are already extremely lengthy.

The test alternative is a mildly explosive bubble process with either a single bubble or sequence of multiple bubble episodes. The data generating processes that are used to capture bubble effects are extended versions of the PWY bubble model. That model has a single explosive episode and collapse within the sample period $[1, T]$ and has the following form

$$\begin{aligned} X_t &= X_{t-1} 1\{t < \tau_e\} + \delta_T X_{t-1} 1\{\tau_e \leq t \leq \tau_f\} \\ &+ \left(\sum_{k=\tau_f+1}^{t} \varepsilon_k + X_{\tau_f}^* \right) 1\{t > \tau_f\} + \varepsilon_t 1\{j \leq \tau_f\}. \end{aligned} \tag{4}$$

As usual, it is convenient to work with fractions of the sample T and we use the notation $t = \lfloor Tr \rfloor$ to denote the integer part of Tr for $r \in [0, 1]$. In the process (4) a mildly explosive bubble runs from $\tau_e = \lfloor Tr_e \rfloor$ to $\tau_f = \lfloor Tr_f \rfloor$ with an expansion rate determined by the mildly explosive coefficient $\delta_T = 1 + cT^{-\alpha}$ with $c > 0$ and $\alpha \in (0, 1)$. When the bubble terminates, the process collapses to a value $X_{\tau_f}^*$ which equals X_{τ_e} plus an $O_p(1)$ perturbation (i.e. $X_{\tau_f}^* = X_{\tau_e} + X^*$) at period $\tau_f + 1$, which represents a re-initialization of the process to a level that relates to the last pre-bubble observation X_{τ_e}. The pre-bubble period $N_0 = [1, \tau_e)$ and post-bubble period $N_1 = (\tau_f, \tau_T]$ are assumed to follow a pure random walk process.

The model is readily extended to include multiple bubble episodes. Suppose there are K bubble episodes in the sample period, represented in terms of sample fraction intervals as $B_i = [\tau_{ie}, \tau_{if}]$ for $i = 1, 2, \ldots, K$. The shifting dynamics of X_t are then given by the model

$$\begin{aligned} X_t &= (X_{t-1} + \varepsilon_t) 1\{t \in N_0\} + (\delta_T X_{t-1} + \varepsilon_t) 1\{t \in B_i\} \\ &+ \sum_{i=1}^{K} \left(\sum_{l=\tau_{if}+1}^{t} \varepsilon_l + X_{\tau_{if}}^* \right) 1\{t \in N_i\}, \end{aligned} \tag{5}$$

where $X_{\tau_{if}}^* = X_{\tau_{ie}} + X^*$ and the intervening subperiods $N_0 = [1, \tau_{1e})$, $N_j = (\tau_{j-1f}, \tau_{je})$ with $j = 1, \ldots, K-1$, and $N_K = (\tau_{Kf}, \tau_T]$ are 'normal' intervals of pure random walk (or more generally martingale) evolution.

The dating algorithms studied here are implemented repeatedly for observations starting from some initialization $\lfloor Tr_0 \rfloor$, where r_0 is the minimum window size required to initiate the regression. For each individual observation $t = \lfloor Tr \rfloor$, we suppose that interest centres on whether this particular observation comes from a bubble realization or an interval of normal martingale behavior. Both the PWY and PSY algorithms use data from the same information set that starts from the first observation and goes up to the observation of interest (i.e. $I_r = \{1, 2, \ldots, \lfloor Tr \rfloor\}$).

PWY conduct recursive right tailed unit root tests with sample data running from the first observation to the current observation $t = \lfloor Tr \rfloor$. The corresponding unit root t statistic at $t = \lfloor Tr \rfloor$ is denoted DF_r. PSY conduct recursive right tailed unit root tests repeatedly on a sequence of (backward expanding from observation t) windows of data and perform inference based on the sup value of this t statistic sequence. Let r_1 and r_2 denote the start and end points of the regression. The regression window width r_w then equals $r_2 - r_1$. With the end point of the regressions r_2 fixed at r (so that the test refers to the state of the process at the current observation $t = \lfloor Tr \rfloor$) and $r_1 \geq 0$, the backward expanding sample sequence extends the window size r_w from r_0 to r_2 (which is equivalent to varying r_1 from 0 to $r_2 - r_0$). The corresponding unit root test sequence is denoted by $\left\{DF_{r_1}^{r_2}\right\}_{r_1 \in [0, r_2 - r_0]}$. The sup value of the test statistic sequence is called the backward SDF statistic and is defined as

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

The origination and termination dates of any bubbles that are detected are calculated using the first crossing principle. Specifically, in the single bubble scenario, the origination (termination) date of the bubble is the first chronological observation whose DF or BSDF statistic exceeds (goes below) its corresponding critical value. The duration of a bubble is restricted to be longer than a slowly varying (at infinity) quantity such as $L_T = \delta \log(T)/T$, where δ is a frequency dependent parameter – see PSY for further discussion. The origination and termination estimators are calculated as the crossing times

$$PWY : \hat{r}^e = \inf_{r \in [r_0, 1]} \left\{r : DF_r > cv^{\beta_T}\right\} \text{ and } \hat{r}^f = \inf_{r \in [\hat{r}^e + L_T, 1]} \left\{r : DF_r < cv^{\beta_T}\right\}, \tag{6}$$

$$PSY : \hat{r}^e = \inf_{r \in [r_0, 1]} \left\{r : BSDF_r(r_0) > scv^{\beta_T}\right\} \text{ and } \hat{r}^f = \inf_{r \in [\hat{r}^e + L_T, 1]} \left\{r : BSDF_r(r_0) < scv^{\beta_T}\right\},$$

(7)

where cv^{β_T} and scv^{β_T} are the $100\,(1-\beta_T)\,\%$ critical values of the DF and BSDF statistics.

In the multiple bubbles scenario, estimators associated with the first bubble are defined as in equation (6) and (7), and denoted by $\hat{r}^{1e}$ and $\hat{r}^{1f}$. The origination (termination) of bubble i (for $i \geq 2$) is the first chronological observation after $\hat{r}^{i-1f}$ whose DF or BSDF statistic exceeds (goes below) its corresponding critical value. Structurally,

$$PWY: \ \hat{r}^{ie} = \inf_{r\in[\hat{r}^{i-1f},1]} \left\{ r: DF_r > cv^{\beta_T} \right\} \text{ and } \hat{r}^{if} = \inf_{r\in[\hat{r}^{ie}+L_T,1]} \left\{ r: DF_r < cv^{\beta_T} \right\} \tag{8}$$

$$PSY: \ \hat{r}^{ie} = \inf_{r\in[\hat{r}^{i-1f},1]} \left\{ r: BSDF_r\left(r_0\right) > scv^{\beta_T} \right\} \text{ and } \hat{r}^{if} = \inf_{r\in[\hat{r}^{ie}+L_T,1]} \left\{ r: BSDF_r\left(r_0\right) < scv^{\beta_T} \right\}. \tag{9}$$

For the sequential PWY procedure, the dating criteria of the first bubble remains the same (i.e. equation (6)). For all subsequent bubbles, the sequential procedure uses information starting from the termination of the previous bubble and ending at the current observation, i.e. $I_{i,r} = \left\{ \left\lfloor T\hat{r}^{i-1f} \right\rfloor + 1, \ldots, \lfloor Tr \rfloor \right\}$ for $i \geq 2$. Importantly, note that the distance between r and $\hat{r}^{i-1f}$ needs to be greater than the minimum regression window r_0, which restricts the capability of this sequential procedure to detect bubble activity in the intervening period $(\hat{r}^{i-1f}, r_0)$. The origination and termination dates of bubble i is then calculated as

$$Seq\_PWY: \hat{r}^{ie} = \inf_{r\in[\hat{r}^{i-1f}+r_0,1]} \left\{ r:_{\hat{r}^{i-1f}} DF_r > cv^{\beta_T} \right\} \text{ and } \hat{r}^{if} = \inf_{r\in[\hat{r}^{ie}+L_T,1]} \left\{ r:_{\hat{r}^{i-1f}} DF_r < cv^{\beta_T} \right\}, \tag{10}$$

where ${}_{\hat{r}^{i-1f}}DF_r$ is the DF statistic calculated over $\left(\hat{r}^{i-1f}, r\right]$.

3 Asymptotic Properties of the Detectors

The asymptotic performance of the dating algorithms is examined in this section. Under the null hypothesis of no bubble episodes, the limit distributions of the DF and BSDF statistics follow from PSY (Theorem 1). Both the DF and BSDF statistics are special cases of the GSADF statistic introduced in PSY. For the DF statistic, the start point of the regression is $r_1 = 0$ and the end point r_2 is fixed at r. Therefore, the limit distribution of the DF statistic under the null hypothesis is

$$F_r(W) := \frac{\frac{1}{2}r\left[W(r)^2 - r\right] - \int_0^r W(s)\,dsW(r)}{r^{1/2}\left\{r\int_0^r W(s)^2\,ds - \left[\int_0^r W(s)\,ds\right]^2\right\}^{1/2}}, \tag{11}$$

where W is a standard Wiener process. For the BSDF statistic, the end point r_2 is fixed at r and the start point r_1 varies from 0 to $r - r_0$. The limit distribution of the BSDF statistic is

$$F_r(W, r_0) := \sup_{\substack{r_1\in[0,r-r_0]\\ r_w=r-r_1}}\left\{\frac{\frac{1}{2}r_w\left[W(r)^2 - W(r_1)^2 - r_w\right] - \int_{r_1}^r W(s)\,ds\,[W(r) - W(r_1)]}{r_w^{1/2}\left\{r_w\int_{r_1}^r W(s)^2\,ds - \left[\int_{r_1}^r W(s)\,ds\right]^2\right\}^{1/2}}\right\}. \tag{12}$$

The asymptotic critical values cv^{β_T} and scv^{β_T} are defined as the $100(1-\beta_T)\%$ quantiles of $F_r(W)$ and $F_r(W, r_0)$, respectively. Notice that the significance level β_T depends on the sample size T and it is assumed that $\beta_T \to 0$ as $T \to \infty$. This control ensures that cv^{β_T} and scv^{β_T} diverge to infinity and thereby under the null hypothesis the probabilities of (falsely) detecting a bubble using the DF and BSDF statistics, (6) - (10), tend to zero as $T \to \infty$.

We next derive the limit distributions under mildly explosive alternatives. We consider the case of a single bubble and multiple bubbles separately as the properties of some of the detectors differ markedly in the case of multiple bubbles. The derivations require some careful calculations involving weak convergence arguments and mildly explosive limit theory, paying attention to some subtleties in the orders of magnitude of the various components of the test statistics. The details are provided in the Appendix, together with the technical supplement to the paper (Phillips, Shi and Yu, 2013c).

Single Bubble Alternative

Theorem 1. *Under the data generating process (4), the asymptotic distributions of the DF_r and $BSDF_r(r_0)$ statistics are as follows:*

$$DF_r \sim_a \begin{cases} F_r(W) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_e}\frac{r_w^{3/2}B(r_e)}{2(r_e-r_1)\int_{r_1}^{r_e}B(s)ds} & \text{if } r \in B \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \end{cases} \tag{13}$$

$$BSDF_r(r_0) \quad \sim_a \begin{cases} F_r(W, r_0) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_e} \sup_{r_1 \in [0, r-r_0]} \left\{ \frac{r_w^{3/2} B(r_e)}{2(r_e - r_1) \int_{r_1}^{r_e} B(s) ds} \right\} & \text{if } r \in B \\ -T^{(1-\alpha)/2} \sup_{r_1 \in [0, r-r_0]} \left\{ \left(\frac{1}{2} c r_w\right)^{1/2} \right\} & \text{if } r \in N_1 \end{cases}, \tag{14}$$

where $B(r) \equiv \sigma W(r)$.

Evidently, for all three cases the order magnitudes of the DF and BSDF statistics are the same. Specifically, the test statistics diverge to positive infinity when the current observation falls in the explosive bubble period and to negative infinity when it is in a bubble collapsing period. Based on these limit distributions, we have the following consistency results for the date detectors.

Theorem 2 (PWY detector)**.** *Suppose* $\hat{r}_e$ *and* $\hat{r}_f$ *are the date estimates obtained from the DF t statistic crossing times (6). Under the alternative hypothesis of mildly explosive behavior in model (4), if*

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}} \to 0, \tag{15}$$

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

Theorem 3 (PSY detector)**.** *Suppose* $\hat{r}_e$ *and* $\hat{r}_f$ *are the date estimates obtained from the backward sup DF statistic crossing times (7). Under the alternative hypothesis of mildly explosive behavior in model (4), if*

$$\frac{1}{scv^{\beta_T}} + \frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}} \to 0, \tag{16}$$

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. The regularity conditions in Theorems 2 and 3 imply that the orders of magnitude of the critical value expansion rates need to be smaller than $T^{1/2}\delta_T^{\tau-\tau_e}$ to deliver consistency of $\hat{r}_e$ and $\hat{r}_f$. In effect, for consistent estimation of r_e the critical value sequence needs to pass to infinity but not too fast – otherwise the signal from the mildly explosive period under the alternative is not strong enough to ensure that the critical value is exceeded. The first condition ($cv^{\beta_T}, scv^{\beta_T} \to \infty$) ensures that there are no false positives prior to the origination date r_e. The second condition ($\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}}, \frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}} \to 0$) ensures that the signal from the data during the mildly explosive period dominates that from the earlier unit root period leading to identifying information that there is now exuberance in the data.

An implicit restriction in these two results is that the minimum window size r_0 is smaller than the origination date of the bubble r_e (i.e. $r_0 < r_e$) so that the recursive regressions provide information for some $r \in N_0$ for comparison to identify the origination point. This requirement is also implicit in what follows, in particular in later proofs of consistency of the first bubble orgination date in the multiple bubbles scenario as discussed below. In the event that $r_0 \in (r_e, r_f)$, then the results given in the second panels of (13) and (14) are relevant and the origination date of the first bubble is determined to be r_0, so r_e is estimated with delay.

For consistent estimation of r_f, both conditions again come into play. The second condition ($\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}}, \frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_e}} \to 0$) ensures that there is no underestimation of r_f asymptotically because for $r \leq r_f$ the signal from the data during the mildly explosive period continues to dominate. When $r > r_f$, the autoregressive estimate is calculated from data that involves the explosive episode as well as post explosive ($r > r_f$) data, which makes the post-collapse data look mean reverting and, as shown in the proofs of Theorems 2 and 3, the test statistics become negative. The expansion condition ($cv^{\beta_T}, scv^{\beta_T} \to \infty$) then ensures that there is no overestimation of r_f asymptotically.

Multiple Bubble Alternatives

The limit behavior of the recursive DF and BSDF statistics in the presence of multiple bubbles is much more complicated. The strengths and weaknesses of the various detectors are well illustrated by considering a mildly explosive process with two bubble episodes. We therefore confine much of our discussion here to the case of model (5) with $K = 2$. Even in this case, as shown below, there are several possibilities depending on the respective durations of the bubbles.

We start with the case where the duration of the first bubble exceeds that of the second bubble. Also, to obtain the BSDF asymptotics in Theorems 4 and 5, it is assumed that the distance separating the termination dates of the first and second bubbles exceeds the minimum window size (i.e. $r_{2e} - r_{1f} > r_0$). This requirement seems a natural condition to achieve identification of the second bubble. The effect of its relaxation is considered later.

Theorem 4. *Under the data generating process of (5) with $K = 2$ and $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$, the limit behavior of the recursive statistics DF_r, $BSDF_r(r_0)$ and ${}_{\hat{r}_{1f}}DF_r$ is given by:*

$$DF_r \sim_a \begin{cases} F_r(W) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_{1e}} \frac{r_w^{3/2} B(r_{1e})}{2(r_{1e}-r_1)\int_{r_1}^{r_{1e}} B(s)ds} & \text{if } r \in B_1 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \cup B_2 \cup N_2 \end{cases} \tag{17}$$

$$BSDF_r(r_0) \sim_a \begin{cases} F_r(W, r_0) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_{ie}} \sup\limits_{r_1 \in [0, r-r_0]} \left\{ \frac{r_w^{3/2} B(r_{ie})}{2(r_{ie}-r_1)\int_{r_1}^{r_{ie}} B(s)ds} \right\} & \text{if } r \in B_i \text{ with } i = 1, 2 \\ -T^{(1-\alpha)/2} \sup\limits_{r_1 \in [0, r-r_0]} \left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \cup N_2 \end{cases} \tag{18}$$

$${}_{\hat{r}_{1f}}DF_r \sim_a \begin{cases} F_r(W) & \text{if } r \in N_1 \\ T^{1/2}\delta_T^{\tau-\tau_{2e}} \frac{r_w^{3/2} B(r_{2e})}{2(r_{2e}-r_1)\int_{r_1}^{r_{2e}} B(s)ds} & \text{if } r \in B_2 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_2 \end{cases} \tag{19}$$

Evidently from the first panel (17), it is clear that when the duration of the first bubble exceeds that of the second bubble, the DF statistic diverges to positive infinity when $r \in B_1$, whereas for $r \in N_1 \cup B_2 \cup N_2$, the statistic is asymptotically equivalent to $-T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2}$ and tends to negative infinity as $T \to \infty$. Importantly, therefore, the behavior of the DF statistic during the second (shorter) bubble B_2 is the same as it is for the normal martingale periods N_1 and N_2. Hence, the DF statistic does not have discriminatory power for second bubble detection when the duration of the second bubble is less than that of the first bubble.

From the second panel (18), the behavior of the BSDF statistic in both bubble periods B_1 and B_2 is the same and is distinct from that of the normal periods N_0, N_1 and N_2. Unlike the DF statistic, the BSDF statistic therefore has discriminatory power in detecting both bubbles. From the final panel (19), it is clear that the limit behavior of the sequential DF statistic ${}_{\hat{r}_{1f}}DF_r$ is the same as that of the BSDF statistic for $r \in B_2$ and $r \in N_2$. Hence, like BSDF, the sequential DF statistic has discriminatory power for the two bubble periods.

Next consider the case where the duration of the second bubble exceeds that of the first bubble.

Theorem 5. *Under the data generating process of (5) with $K = 2$ and $\tau_{1f} - \tau_{1e} \leq \tau_{2f} - \tau_{2e}$, the limit behavior of the recursive statistics DF_r, $BSDF_r(r_0)$ and ${}_{\hat{r}_{1f}}DF_r$ is as follows:*

$$DF_r \sim_a \begin{cases} F_r(W) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_{1e}} \frac{r_w^{3/2} B(r_{1e})}{2(r_{1e}-r_1)\int_{r_1}^{r_{1e}} B(s)ds} & \text{if } r \in B_1 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \cup N_2 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in B_2 \text{ and } r_{1f} - r_{1e} > r - r_{2e} \\ T^{1-\alpha/2}\left[\frac{cr_w^3}{2\left(r_{1e}-r_1+r_{2e}-r_{1f}\right)}\right]^{1/2} & \text{if } r \in B_2 \text{ and } r_{1f} - r_{1e} \leq r - r_{2e} \end{cases} \quad (20)$$

$$BSDF_r(r_0) \sim_a \begin{cases} F_r(W, r_0) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_{ie}} \sup_{r_1 \in [0, r-r_0]} \left\{\frac{r_w^{3/2} B(r_{ie})}{2(r_{ie}-r_1)\int_{r_1}^{r_{ie}} B(s)ds}\right\} & \text{if } r \in B_1 \cup B_2 \\ -T^{(1-\alpha)/2} \sup_{r_1 \in [0, r_2-r_0]} \left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \cup N_2 \end{cases} \quad (21)$$

$${}_{\hat{r}_{1f}}DF_r \sim_a \begin{cases} F_r(W) & \text{if } r \in N_1 \\ T^{1/2}\delta_T^{\tau-\tau_{2e}} \frac{r_w^{3/2} B(r_{ie})}{2(r_{ie}-r_1)\int_{r_1}^{r_{ie}} B(s)ds} & \text{if } r \in B_2 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_2 \end{cases}. \quad (22)$$

As is evident in panels (21) and (22) of this theorem, the limit behaviors of the BSDF statistic and sequential DF statistic are identical to those that apply in the earlier case where $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$. Thus both procedures have the same discriminatory capability for identifying bubble episodes in the data. Again, results are very different for the DF statistic where the behavior of the statistic during the second bubble ($r \in B_2$) is contingent on the timing of latest date (r) in the recursion. In particular, when $r \in B_2$, the limit behavior of the DF statistic depends on the relative length of $r_{1f} - r_{1e}$ (the duration of the first bubble) and $r - r_{2e}$ (the segment of the second bubble that is included in data used in the recursion). When $r_{1f} - r_{1e}$ exceeds $r - r_{2e}$, the statistic diverges to negative infinity, just as for the case where $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$. Thus, in this case there is insufficient data to identify the second bubble period. However, as is clear from the final asymptotic expression in (20), behavior changes dramatically as soon as there is more data. Specifically, when the segment of the second bubble included in the recursive regression exceeds the duration of the first bubble (i.e., when $r - r_{2e} \geq r_{1f} - r_{1e}$) the sign in the limit behavior of the DF statistic changes and the statistic now diverges to positive infinity rather than negative infinity. The order of the magnitude in the divergence also rises (from $T^{(1-\alpha)/2}$ to $T^{1-\alpha/2}$). It follows that the DF statistic has discriminatory power once there is sufficient data for this test to identify a second bubble - that is, as soon as data from the second bubble dominates that of the first bubble.

With the limit behavior of the recursive tests in hand, results on the consistency properties of the bubble date detectors now follow. It is convenient to separate the results according to each of the recursive tests and contingent conditions regarding duration of the bubbles.

Theorem 6 (PWY detector)**.** *Suppose $\hat{r}_{1e}$, $\hat{r}_{1f}$, $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are obtained from the DF test based on the t statistic (8). Given the alternative hypothesis of mildly explosive behavior in model (5) with $K = 2$ and durations satisfying $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$, if*

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0,$$

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 7 (PWY detector)**.** *Suppose $\hat{r}_{1e}$, $\hat{r}_{1f}$, $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are obtained from the DF test based on the t statistic (8). Given the alternative hypothesis of mildly explosive behavior in model (5) with $K = 2$ and durations satisfying $\tau_{1f} - \tau_{1e} \leq \tau_{2f} - \tau_{2e}$, if*

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0,$$

we have $\hat{r}_{1e} \xrightarrow{p} r_{1e}$ and $\hat{r}_{1f} \xrightarrow{p} r_{1f}$; if

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1-\alpha/2}} \to 0$$

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

Theorem 8 (PSY detector)**.** *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 (9). Given the alternative hypothesis of mildly explosive behavior* *in model (5) with $K = 2$, if*

$$\frac{1}{scv^{\beta_T}} + \frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{ie}}} \to 0 \text{ with } i = 1, 2,$$

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 9 (Sequential PWY detector)**.** *Suppose $\hat{r}_{1e}$, $\hat{r}_{1f}$, $\hat{r}_{2e}$ and $\hat{r}_{2f}$ are obtained from sequential application of the DF test based on the t statistics (6) and (10). Given the alternative hypothesis of mildly explosive behavior in model (5) with $K = 2$, if*

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{ie}}} \to 0,$$

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* $T \to \infty$.

Theorems 6 - 9 characterize the consistency properties of the detectors when there are two bubble episodes in the observed data. The results depend on the detector and certain side conditions regarding the duration of the bubbles. Importantly, the PWY strategy consistently estimates the origination and termination of the first bubble but not the second bubble. When the duration of the first bubble exceeds that of the second bubble, the PWY strategy fails to detect the second bubble. When the duration of the second bubble exceeds the first, the PWY recursion detects the presence of a second bubble but with a delay measured by the duration of the first bubble ($r_{1f} - r_{1e}$). The PWY detector is therefore inconsistent in date stamping the second bubble even when the conditions favor its detection. In contrast, the PSY and sequential PWY recursions are both consistent date detectors for the origination and termination of the two bubbles irrespective of their relative durations. These procedures are therefore robust to bubble duration.

Theorems 6 - 9 can be extended to scenarios with multiple bubbles ($K > 2$). In this case, if the duration of bubble $i+1$ is less than that of bubble i for some $i \in \{1, 2, \cdots, K-1\}$, then the PWY recursion may, under certain conditions such as increasing duration up to bubble i, detect the presence of bubble i, but it will not detect bubble $i+1$. In contrast, the PSY and sequential PWY strategies detect each of the K bubbles, with fully consistent date detection by the PSY recursion.

We now consider the extreme scenario, mentioned earlier, where the minimum window length r_0 exceeds the distance between the termination dates of the two bubbles. Suppose $K = 2$. For the sequential PWY procedure, the first regression after re-initialization from the end point of the first bubble now runs from period N_1 directly to N_2, so this procedure completely passes over the second bubble and is unable to detect it. Somewhat remarkably however, the PSY strategy still has some detective capability for the second bubble depending on the relative length of $\tau_{1f} - \tau_{1e}$ and $\tau_2 - \tau_{2e}$. Specifically, for observations in the second bubble episode (i.e. $r \in B_2$), their backward expanding regression sample sequences does not include the case of $\tau_1 \in N_1$ and $\tau_2 \in B_2$ when $r_0 > r_{2f} - r_{1f}$. Hence, the limit behavior of $BSDF_r(r_0)$ under the two-bubble data generating process is

$$BSDF_r(r_0) \sim_a \begin{cases} -T^{(1-\alpha)/2} \sup\limits_{r_1 \in [0, r_2 - r_0]} \left(\frac{1}{2} c r_w\right)^{1/2} & \text{if } r \in B_2 \text{ and } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ T^{1-\alpha/2} \left(\frac{cr^3}{2\left(r_{1e} + r_{2e} - r_{1f}\right)}\right)^{1/2} & \text{if } r \in B_2 \text{ and } \tau_{1f} - \tau_{1e} \leq \tau_2 - \tau_{2e} \end{cases}. \quad (23)$$

Then, if $\tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e}$, the limit behavior of $BSDF_r(r_0)$ at $r \in B_2$ is the same as when $r \in N_1 \cup N_2$, so in that event the PSY strategy also cannot detect the second bubble. But when $\tau_{1f} - \tau_{1e} \leq \tau_2 - \tau_{2e}$, the limit behavior of $BSDF_r(r_0)$ at $r \in B_2$ is divergent with an order magnitude of $T^{1-\alpha/2}$. Hence, even though $r_0 > r_{2f} - r_{1f}$, the PSY strategy is still able to detect the second bubble (with a delay of $r_{1f} - r_{1e}$ in the estimated origination date) as long as the duration of the second bubble exceeds the first bubble.

A less extreme scenario is the case where $r_{2e} - r_{1f} < r_0 \leq r_{2f} - r_{1f}$. That is, the minimum window size exceeds the distance separating the two bubbles but does not exceed the distance between the termination dates of these two bubbles. In this circumstance, the limit behaviors of $BSDF_r(r_0)$ and ${}_{\hat{r}_{1f}}DF_r$ remain the same as in (21) and (22) for $r_{1f} + r_0 \leq r \leq r_{2f}$ (the later segment of B_2). However, for observations prior to that in B_2, the ${}_{\hat{r}_{1f}}DF_r$ statistic does not exist by construction and the BSDF statistic follows the limit behavior of (23). Therefore, there will be delay in estimates of the second bubble origination date using both the PSY and sequential PWY strategies. However, the delay is potentially smaller using the PSY strategy due to the last panel of (23).

The advantage of the PSY strategy over the sequential PWY procedure is revealed in the simulations reported below which consider some less extreme cases. For instance, when $r_{3e} - r_{2f} < r_0 < r_{3f} - r_{2f}$ (i.e. $0.05 < 0.12 < 0.15$) as in the first panel of Table 10, the detection rate of the sequential PWY strategy is zero as oppose to 62% for the PSY strategy.

4 Simulation Evidence

This section reports simulations to explore the finite sample performance of the PSY, PWY, sequential PWY, and CUSUM procedures for bubble detection. These simulations focus on detection rates and estimation accuracy of the dating algorithms of these procedures. They complement the findings reported in PSY and examine performance characteristics in systems with many bubbles.

Experiments are conducted with generating models that involve up to three separate bubbles. The generating system for single, dual and three bubbles are as in (4) and (5). The parameter settings follow those used in PSY, so that $y_0 = 100$, $\sigma = 6.79$, $c = 1$ and $T = 100$. In the single bubble setting, we explore the sensitivities of the dating strategies to the parameters determining the magnitude of the bubbles (the bubble expansion rate α and the bubble duration $d_T = \tau_f - \tau_e$), the bubble location parameter τ_e and the sample size T. We focus our attention on the impact of bubble durations in the two bubble and three bubble settings. For each parameter constellation, 5,000 replications were used. Bubbles were identified using respective finite sample 95% quantiles, obtained from simulations with 5,000 replications. The minimum window size has 12 observations.

We report the proportion of samples in which a bubble was successfully detected, along with the empirical mean and standard deviation (in parentheses) of the estimated origination and termination dates. Successful detection of a bubble is defined as an outcome where the estimated origination date is greater than or equal to the true origination date and smaller than the true termination date of that particular bubble (i.e. $r_{ie} \leq \hat{r}_{ie} < r_{if}$).

4.1 A Single Bubble

In Tables 1 and 2, the bubble expansion rate α and bubble duration d_T can each take three values: specifically, the expansion rate $\alpha \in \{0.60, 0.55, 0.50\}$ with corresponding autoregressive coefficient $\delta_T \in \{1.04, 1.05, 1.07\}$ when $T = 100$; and duration is $d_T \in \{\lfloor 0.10T \rfloor, \lfloor 0.15T \rfloor, \lfloor 0.20T \rfloor\}$. Evidently for all algorithms the bubble detection rate increases with the value of the autoregressive coefficient δ_T and the bubble duration d_T. Moreover, a higher autoregressive coefficient results in more timely detection of the bubble, whereas longer bubble duration is associated with longer delay (i.e. $\hat{r}_e - r_e$). For instance, the delay in the PSY estimate reduces from 0.05 to 0.03 when δ_T increases from 1.04 to 1.07 and the delay increases from 0.04 to 0.06 when the bubble duration extends from $\lfloor 0.10T \rfloor$ to $\lfloor 0.20T \rfloor$.

Table 1: Detection rate and estimation of the origination and termination dates under single bubble DGP and different bubble expansion rates. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \tau_e = \lfloor 0.4T \rfloor, \tau_f - \tau_e = \lfloor 0.15T \rfloor, T = 100$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\alpha = 0.60, \delta_T = 1.04$				
Detection Rate	0.78	0.86	0.80	0.86
$r_e = 0.40$	0.46 (0.04)	0.45 (0.03)	0.46 (0.03)	0.46 (0.03)
$r_f = 0.55$	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)
$\alpha = 0.55, \delta_T = 1.05$				
Detection Rate	0.85	0.91	0.86	0.91
$r_e = 0.40$	0.45 (0.03)	0.44 (0.03)	0.45 (0.03)	0.45 (0.03)
$r_f = 0.55$	0.55 (0.00)	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)
$\alpha = 0.50, \delta_T = 1.07$				
Detection Rate	0.90	0.94	0.91	0.93
$r_e = 0.40$	0.45 (0.03)	0.43 (0.03)	0.45 (0.03)	0.44 (0.03)
$r_f = 0.55$	0.55 (0.00)	0.55 (0.00)	0.55 (0.00)	0.55 (0.01)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

In Table 3, the location parameter τ_e varies from $\lfloor 0.2T \rfloor$ to $\lfloor 0.6T \rfloor$. When the bubble originates at a later stage of the sample, the bubble detection rates of all strategies are lower. Table 4 monitors the effects of increasing the sample size from 100 to 400. Evidently, the bubble

Table 2: Detection rate and estimation of the origination and termination dates under single bubble DGP and different bubble durations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, \tau_e = \lfloor 0.4T \rfloor, T = 100$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_f - \tau_e = \lfloor 0.10T \rfloor$				
Detection Rate	0.57	0.71	0.57	0.69
$r_e = 0.40$	0.44 (0.02)	0.44 (0.02)	0.44 (0.02)	0.44 (0.02)
$r_f = 0.50$	0.50 (0.00)	0.50 (0.00)	0.50 (0.01)	0.50 (0.01)
$\tau_f - \tau_e = \lfloor 0.15T \rfloor$				
Detection Rate	0.78	0.86	0.80	0.86
$r_e = 0.40$	0.46 (0.04)	0.45 (0.03)	0.46 (0.03)	0.46 (0.03)
$r_f = 0.55$	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)
$\tau_f - \tau_e = \lfloor 0.20T \rfloor$				
Detection Rate	0.87	0.93	0.88	0.92
$r_e = 0.40$	0.47 (0.04)	0.46 (0.04)	0.47 (0.04)	0.46 (0.04)
$r_f = 0.60$	0.60 (0.01)	0.60 (0.01)	0.60 (0.01)	0.60 (0.01)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

detection rate increases with the sample size as expected. But the time needed to detect bubbles in all algorithms is largely unaffected by the location of the bubble and the sample size.

The most striking finding in Tables 1 - 3 is the superiority of the PSY strategy relative to the other algorithms in the single bubble case. The PSY strategy has a higher rate of bubble detection and provides a more accurate estimate of the origination date. All strategies deliver a good detection rate of the termination date of the bubble, which is no doubt associated with the sharp collapse specification in the model formulation.

4.2 Two Bubbles

Two duration scenarios feature in the dual bubble simulations. In one the first bubble has longer duration (Table 5), while in the other the second bubble has longer duration (Table 6). The bubbles originate 20% and 60% into the sample and the expansion rate of the two bubbles is 1.04 (i.e. $\alpha = 0.6$).

In Table 5, the duration of the first bubble is 20% of the total sample. The duration of the

Table 3: Detection rate and estimation of the origination and termination dates under single bubble DGP and different bubble locations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, \tau_f - \tau_e = \lfloor 0.15T \rfloor, T = 100$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_e = \lfloor 0.2T \rfloor$				
Detection Rate	0.88	0.91	0.87	0.87
$r_e = 0.20$	0.26 (0.03)	0.25 (0.03)	0.26 (0.03)	0.26 (0.03)
$r_f = 0.35$	0.35 (0.01)	0.35 (0.01)	0.35 (0.01)	0.35 (0.01)
$\tau_e = \lfloor 0.4T \rfloor$				
Detection Rate	0.78	0.86	0.80	0.86
$r_e = 0.40$	0.46 (0.04)	0.45 (0.03)	0.46 (0.03)	0.46 (0.03)
$r_f = 0.55$	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)
$\tau_e = \lfloor 0.6T \rfloor$				
Detection Rate	0.72	0.83	0.74	0.82
$r_e = 0.60$	0.66 (0.03)	0.65 (0.03)	0.66 (0.03)	0.65 (0.03)
$r_f = 0.75$	0.75 (0.01)	0.75 (0.01)	0.75 (0.01)	0.75 (0.01)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

second bubble is shorter than the first one, taking values $d_T = \tau_{2f} - \tau_{2e} = \lfloor 0.10T \rfloor, \lfloor 0.15T \rfloor$. As anticipated from asymptotic theory, PWY fails to detect the second bubble in this duration scenario. For instance, when $d_T = \lfloor 0.10T \rfloor$, the proportion of samples where the second bubble is detected using PWY is negligible (around 0.01). Noticeably, all algorithms perform well in identifying the first bubble. The average delay in detecting this bubble is four to five observations.

The opposite setting is considered in the simulations reported in Table 6. Here the duration of the first bubble is fixed at $\lfloor 0.10T \rfloor$ and the duration of the second bubble varies from $\lfloor 0.10T \rfloor$ to $\lfloor 0.20T \rfloor$. Several results emerge from the table. First, there is no dramatic performance difference in identifying the first bubble among the dating algorithms. It is interesting to note that, due to its shorter bubble duration, the detection rates for the first bubble are lower than those in Table 5. Second, we observe a significant boost in the second bubble detection rate for the PWY strategy. In particular, when the duration of the second bubble is twice as long as the first, the detection rates of the PWY strategy is 76%. This outcome contrasts sharply with the

Table 4: Detection rate and estimation of the origination and termination dates under single bubble DGP and different sample sizes. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.60, \tau_e = \lfloor 0.4T \rfloor, \tau_f - \tau_e = \lfloor 0.15T \rfloor, \tau_f - \tau_e = \lfloor 0.15T \rfloor$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$T = 100$				
Detection Rate	0.78	0.86	0.80	0.86
$r_e = 0.40$	0.46 (0.04)	0.45 (0.03)	0.46 (0.03)	0.46 (0.03)
$r_f = 0.55$	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)	0.55 (0.01)
$T = 200$				
Detection Rate	0.80	0.93	0.83	0.89
$r_e = 0.40$	0.46 (0.04)	0.45 (0.04)	0.46 (0.04)	0.45 (0.03)
$r_f = 0.55$	0.55 (0.01)	0.54 (0.02)	0.55 (0.02)	0.55 (0.02)
$T = 400$				
Detection Rate	0.86	0.99	0.89	0.86
$r_e = 0.40$	0.46 (0.04)	0.45 (0.04)	0.46 (0.04)	0.45 (0.03)
$r_f = 0.55$	0.55 (0.02)	0.54 (0.04)	0.54 (0.02)	0.54 (0.03)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

PWY detection rates for the second bubble displayed in Table 5. Third, there are relatively long delays in PWY detection of the second bubble. As a case in the point, when the duration of the second bubble is $\lfloor 0.20T \rfloor$, the PWY estimate of the origination date of the second bubble is 0.71 with a delay of 11 observations (nearly twice as long as the delay in detection of 6 observations when using PSY). Those findings corroborate closely the asymptotic theory, which shows how the PWY detector consistently estimates the first bubble but only identifies the second bubble with some delay when $\tau_{2f} - \tau_{2e} > \tau_{1f} - \tau_{1e}$.

In both experiments (Tables 5 and 6), the performance of the CUSUM procedure follows closely that of the PWY procedure. The PSY and the sequential PWY detectors are much more reliable in all cases, as shown in their higher detection rates and more timely detection of both bubbles. Overall, the findings indicate that the PSY strategy provides the best performance when there are two bubbles in the time series.

Table 5: Detection rate and estimation of the origination and termination dates under two bubble DGP with shorter second bubble durations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, \tau_{1e} = \lfloor 0.20T \rfloor, \tau_{2e} = \lfloor 0.60T \rfloor, \tau_{1f} - \tau_{1e} = \lfloor 0.20T \rfloor, T = 100$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_{2f} - \tau_{2e} = \lfloor 0.10T \rfloor$				
Detection Rate (1)	0.93	0.97	0.93	0.95
$r_{1e} = 0.20$	0.26 (0.04)	0.26 (0.04)	0.26 (0.04)	0.27 (0.04)
$r_{1f} = 0.40$	0.40 (0.01)	0.40 (0.01)	0.40 (0.01)	0.40 (0.01)
Detection Rate (2)	0.01	0.73	0.67	0.03
$r_{2e} = 0.60$	0.67 (0.02)	0.64 (0.02)	0.64 (0.02)	0.66 (0.02)
$r_{2f} = 0.70$	0.70 (0.00)	0.70 (0.00)	0.70 (0.00)	0.70 (0.00)
$\tau_{2f} - \tau_{2e} = \lfloor 0.15T \rfloor$				
Detection Rate (1)	0.93	0.97	0.93	0.95
$r_{1e} = 0.20$	0.26 (0.04)	0.26 (0.04)	0.26 (0.04)	0.27 (0.04)
$r_{1f} = 0.40$	0.40 (0.01)	0.40 (0.01)	0.40 (0.01)	0.40 (0.02)
Detection Rate (2)	0.05	0.89	0.83	0.13
$r_{2e} = 0.60$	0.70 (0.03)	0.65 (0.03)	0.65 (0.03)	0.70 (0.03)
$r_{2f} = 0.75$	0.75 (0.00)	0.75 (0.01)	0.75 (0.01)	0.75 (0.01)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

4.3 Three bubbles

Table 7 - 10 report findings for the three bubble case. In Tables 7 - 9, we adjust the duration of one bubble to $d_T \in \{\lfloor 0.10T \rfloor, \lfloor 0.20T \rfloor\}$ and fix the durations of the other two bubbles. The bubbles originate 15%, 45% and 75% into the sample and the bubble expansion rate is 1.04 in each case.

Results are similar to the two bubble case and are consistent with asymptotic theory in the more complex scenarios of multiple bubbles. First, when the duration of bubble i (for $i = 1, 2$) is longer than bubble $i + 1$, theory indicates that the PWY strategy is not capable of detecting the presence of bubble $i + 1$. The simulation findings in Table 7 show that, due to the longer duration of the second bubble where $d_T = \lfloor 0.20T \rfloor$, the PWY detection rate is zero for the third bubble, whose duration is $d_T = \lfloor 0.10T \rfloor$. Similar results are found in Table 9 where the duration of the first bubble is longer than the second. An interesting feature of the PWY outcomes is that the presence of a long duration bubble causes weak identification of all subsequent bubbles. In particular, when the first bubble lasts longer than the second and third bubbles (the first panel of Table 9), the PWY detection rates of these two bubbles are 0.00 and 0.01.

Second, the simulations confirm that when the duration of bubble i is shorter than that of bubble $i+1$, the PWY strategy detects the existence of both bubbles but with a delay in the identification of bubble $i+1$. A case in point occurs in the first panel of Table 8 where the duration of the second bubble is shorter than that of the third bubble. The detection rate of the third bubble using the PWY strategy is 0.68 and the length of the delay in the detection of this bubble is $\lfloor 0.13T \rfloor$, more than twice the delay incurred by the PSY detector. Third, just as for the two bubble case, the behaviour of the CUSUM detector resembles that of PWY.

Fourth, the performances of PSY and sequential PWY are invariant to the relative durations among the bubbles. In other words, the frequency of detecting bubble i and the time needed to detect this bubble depend on the duration of this particular bubble, not on the duration of bubble j (for $j \neq i$).

Overall best performance is delivered by the PSY algorithm, followed by the sequential PWY strategy. Notice that when the duration of bubble i is twice as long as the duration of bubble $i+1$, the sequential PWY detection rate of bubble $i+1$ rises to a higher level than PSY. For example, in the first panel of Table 7 where $\tau_{2f} - \tau_{2e} = \lfloor 0.20T \rfloor$ and $\tau_{3f} - \tau_{3e} = \lfloor 0.10T \rfloor$, the third bubble detection rate of sequential PWY is 0.81, exceeding that of PSY at 0.73. This is due to the fact that the sequential procedure re-initializes after the collapse of the second bubble and the first regression following re-initialization already covers several observations of the third bubble episode. This situation resembles the case of bubbles occurring at the beginning of the sample, which increases the bubble detection rate as shown in Table 3.

In extreme cases when the first regression after re-initialization covers most observations of the particular bubble episode, the sequential PWY procedure may fail to detect this bubble. Table 10 gives examples that forcefully illustrate this point. In the first panel of the table, the sequential PWY procedure re-initiates at $\lfloor 0.65T \rfloor$ and the undetectable period (due to the minimum regression window requirement of 12 observations) following this re-initialization is over the period $\lfloor 0.65T \rfloor$ to $\lfloor 0.77T \rfloor$ and covers most of the third bubble episode. As a result, the detection rate of the third bubble episode using the sequential PWY procedure is zero, whereas the detection rate of the third bubble using PSY is 62%. A further example occurs in the bottom panel of the same table. For the same reason, the sequential procedure fails to detect the second bubble episode in 94% of cases – the detection rate reported in the table is only 6%. Noticeably, the unsuccessful detection of the second bubble also leads to a low detection rate for the third bubble, which may be partly explained by the fact that the remaining sample period includes two bubble episodes. In all of these cases the PSY detector works well with a high average detection rate (94%, 62% and 76% for bubbles 1, 2, and 3 respectively) and an average delay of 4-7 observations in detection.

5 Conclusions

We develop limit theory for real time dating of the origination and termination of mildly explosive periods using detectors based on the PWY, PSY, and sequential PWY algorithms. All three strategies rely on recursive right tailed unit root tests but involve different types of recursion. The asymptotic performance of the detectors are evaluated using the extended PWY bubble model where mildly explosive bubble episodes are embedded within a longer period of normal stochastic trend behavior.

The PWY date estimates are shown to depend on the number of bubble episodes within the sample period and the relative durations of the bubbles when there are multiple bubble episodes. Specifically, in the single bubble case, the PWY estimators are consistent under some mild regularity conditions. When the sample period includes two bubble episodes, the PWY approach can consistently estimate the first bubble but not the second. The dating accuracy of the second bubble is related to the relative duration of the two bubbles. If the first bubble lasts longer than the second, the PWY strategy cannot detect occurrence of the second bubble. Alternatively, if the duration of the second bubble exceeds the first, the PWY detector finds the second bubble but with some delay even asymptotically. In contrast, the PSY approach and a sequential implementation of the PWY strategy both provide consistent estimators of all bubbles regardless the number of bubble episodes occurring in the sample period and their relative duration.

Finite sample simulation are strongly confirmative of the asymptotics, indicating that the PSY algorithm is much more reliable as a detector than the PWY strategy. The second best procedure is the sequential PWY strategy. The performance of the CUSUM procedure resembles that of the PWY strategy and has similar disadvantages in multiple bubble cases.

The results obtained here require some detailed and complex calculations to obtain the limit theory of the various recursive detection algorithms. While these results are specific to the bubble model context under study, the methods should be useful in other recursive regression contexts. Also, with some modifications, the results continue to hold under more general conditions on the innovations than those used here. The main requirements are that the weak convergence (2) applies under normal periods and the limit theory for mildly explosive periods applies as it is known to do under general forms of weak dependence (Phillips and Magdalinos, 2007b).

6 References

Bohl, M. T., Kaufmann, P., and Stephan, P. M., 2013, From hero to zero: Evidence of performance reversal and speculative bubbles in german renewable energy stocks. *Energy Economics,* 37: 40-51.

Das, S., Gupta, R., and Kanda, P. T., 2011, International articles: Bubbles in south African house prices and their impact on consumption. *Journal of Real Estate Literature*, 19(1):69-91.

Etienne, X., Irwin, S. H., and Garcia, P., 2013, Bubbles in food commodity markets: Four decades of evidence. *Working Paper.*

Gilbert, C. L., 2010, Speculative influences on commodity futures prices 2006 – 2008, Working paper 197, United Nations Conference on Trade and Development (UNCTAD).

Gutierrez, L., 2013, Speculative bubbles in agricultural commodity markets. *European Review of Agricultural Economics*, 40(2):217-238.

Homm, U., and Breitung, J., 2012, Testing for speculative bubbles in stock markets: a comparison of alternative methods. *Journal of Financial Econometrics*, 10(1):198-231.

Masters, M. W., 2008. Testimony before the United States Senate Committee of Homeland Security and Government Affairs. Washington, DC, 20 May.

Meyer, G., 2013, Commodity booms, busts and bubbles. *Financial Times*, March 25.

Phillips, P. C. B., 1987, Time series regression with a unit root. *Econometrica: Journal of the Econometric Society*, 55(2):277-301.

Phillips, P. C. B., and Magdalinos, T., 2007a, Limit theory for moderate deviations from a unit root. *Journal of Econometrics*, 136(1):115-130.

Phillips, P. C. B. and T. Magdalinos., 2007b, Limit Theory for Moderate Deviations from Unity under Weak Dependence" in G. D. A. Phillips and E. Tzavalis (Eds.) *The Refinement of Econometric Estimation and Test Procedures: Finite Sample and Asymptotic Analysis.* Cambridge: Cambridge University Press, 2007, pp.123-162.

Phillips, P. C. B., and Magdalinos, T., 2009. Unit root and cointegrating limit theory when initialization is in the in nite past. *Econometric Theory*, 25, 1682-1715.

Phillips, P. C. B., and Perron, P., 1988, Testing for a unit root in time series regression. *Biometrika*, 75(2):335-346.

Phillips, P. C. B., Shi, S., and Yu, J., 2013a, Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics*, forthcoming.

Phillips, P. C. B., Shi, S., and Yu, J., 2013b, Testing for multiple bubbles 1: Historical episodes of exuberance and collapse in the S&P 500. *Working paper*.

Phillips, P. C. B., Shi, S., and Yu, J., 2013c, Technical Supplement to the Paper: Testing for Multiple Bubbles 2: Limit Theory of Real Time Detectors. *Manuscrip,* 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(1):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. *Unpublished manuscrip, Sim Kee Boon Institute for Financial Economics, Singapore Management University.*

Phillips, P. C. B, and Yu, J., 2011a, Warning Signs of Future Asset Bubbles, *The Straits Times*, Singapore, April 26.

Phillips, P. C. B., and Yu, J., 2011b, Dating the timeline of nancial bubbles during the subprime crisis. *Quantitative Economics*, 2(3):455-491.

Phillips, P. C. B, and Yu, J., 2013, Bubble or Roller Coaster in World Stock Markets, *The Business Times*, Singapore, June 28.

Yiu, M., Yu, J., and Jin, L., 2012, Detecting Bubbles in the Hong Kong Residential Property Market: An Explosive-Pattern Approach. Working Paper 012012, Hong Kong Institute for Monetary Research.

APPENDIX A. The Dating Algorithms (a single bubble)

Section A.1 provides some useful preliminary results than characterize the limit behavior of the regression components over the various subperiods of the data. Section A.2 provides test asymptotics and gives proofs of Theorems 1-3 which describe the consistency properties of the PWY and PSY dating strategies.

A.1: 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 $N_1 = [\tau_f + 1, \tau_T]$, 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 $t = \lfloor Tp \rfloor$.
- $B(.) \equiv \sigma W(.)$, where W is standard Brownian motion.

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

Lemma A1. *Under the data generating process,*

(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 $T^{-1/2}X_{t=\lfloor Tp \rfloor} \Rightarrow B(p)$ as $T \to \infty$. (2) For $t \in B$, the data generating process

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

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

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

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

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

This implies that the first term has a higher order than the second term. 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 $t \in N_1$,

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

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

□

Lemma A2. *Under the data generating process,*

(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* $\tau_2 \in N_1$,

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

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

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

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

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

The second term is

$$\begin{aligned}\frac{1}{\tau_w}\sum_{j=\tau_e}^{\tau_2} X_j &= \frac{X_{\tau_e}}{\tau_w}\sum_{j=\tau_e}^{\tau_2}\delta_T^{j-\tau_e}\left\{1+o_p(1)\right\} \text{ from Lemma A1}\\ &= \frac{X_{\tau_e}}{\tau_w}\frac{\delta_T^{\tau_2-\tau_e+1}-1}{\delta_T-1}\left\{1+o_p(1)\right\}\\ &= X_{\tau_e}\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}+c\delta_T^{\tau_2-\tau_e}-T^{\alpha}}{\tau_w c}\left\{1+o_p(1)\right\}\\ &= X_{\tau_e}\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c}\left\{1+o_p(1)\right\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_2-\tau_e}\frac{1}{r_w c}B(r_e). \qquad (26)\end{aligned}$$

Furthermore, we have

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

This implies that $\tau_w^{-1}\sum_{j=\tau_e}^{\tau_2} X_j$ has a higher order than $\tau_w^{-1}\sum_{j=\tau_1}^{\tau_e-1} X_j$. Hence,

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_2} X_j = \frac{1}{\tau_w}\sum_{j=\tau_e}^{\tau_2} X_j\left\{1+o_p(1)\right\}$$

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

(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

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_f} X_j = \frac{X_{\tau_e}}{\tau_w}\sum_{j=\tau_1}^{\tau_f}\delta_T^{j-\tau_e}\left\{1+o_p(1)\right\} \text{ from Lemma A1}$$
$$= \frac{X_{\tau_e}}{\tau_w}\frac{\delta_T^{\tau_f-\tau_1+1}-1}{\delta_T-1}\left\{1+o_p(1)\right\}$$
$$= \frac{X_{\tau_e}}{\tau_w}\frac{T^{\alpha}\delta_T^{\tau_f-\tau_1}+c\delta_T^{\tau_f-\tau_1}-T^{\alpha}}{c}\left\{1+o_p(1)\right\}$$
$$= \frac{T^{\alpha}\delta_T^{\tau_f-\tau_1}}{\tau_w c}X_{\tau_e}\left\{1+o_p(1)\right\}$$
$$\sim_a T^{\alpha-1/2}\delta_T^{\tau_f-\tau_1}\frac{1}{r_w c}B(r_e).$$

The second term is

$$\frac{1}{\tau_w}\sum_{j=\tau_f+1}^{\tau_2} X_j$$
$$= \frac{1}{\tau_w}\sum_{j=\tau_f+1}^{\tau_2}\left[\sum_{k=\tau_f+1}^{j}\varepsilon_k + X_{\tau_e}\right] \tag{27}$$
$$= T^{1/2}\frac{\tau_2-\tau_f}{\tau_w}\left[\frac{1}{\tau_2-\tau_f}\sum_{j=\tau_f+1}^{\tau_2}\left(T^{-1/2}\sum_{k=\tau_f+1}^{j}\varepsilon_k\right)\right] + T^{1/2}\frac{\tau_2-\tau_f}{\tau_w}\left(T^{-1/2}X_{\tau_e}\right)$$
$$\sim_a T^{1/2}\frac{r_2-r_f}{r_w}\int_{r_f}^{r_2}\left[B(s)-B(r_f)\right]ds + T^{1/2}\frac{r_2-r_f}{r_w}B(r_e)$$
$$= T^{1/2}\frac{r_2-r_f}{r_w}\left\{\int_{r_f}^{r_2}\left[B(s)-B(r_f)\right]ds - B(r_e)\right\}. \tag{28}$$

Furthermore, we have

$$\frac{T^{\alpha-1/2}\delta_T^{\tau_f-\tau_1}}{T^{1/2}} = \frac{\delta_T^{\tau_f-\tau_1}}{T^{1-\alpha}} = \frac{e^{c\left(r_f-r_1\right)T^{1-\alpha}}}{T^{1-\alpha}} > 1.$$

This implies that $\tau_w^{-1}\sum_{j=\tau_1}^{\tau_f} X_j$ has a higher order than $\tau_w^{-1}\sum_{j=\tau_f+1}^{\tau_2} X_j$. Hence,

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

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

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

The first term is

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

The second term is

$$\begin{aligned}
\frac{1}{\tau_w}\sum_{j=\tau_e}^{\tau_f} X_j &= \frac{X_{\tau_e}}{\tau_w}\sum_{j=\tau_e}^{\tau_f}\delta_T^{j-\tau_e}\left\{1+o_p(1)\right\} \text{ from Lemma A1} \\
&= \frac{X_{\tau_e}}{\tau_w}\frac{\delta_T^{\tau_f-\tau_e+1}-1}{\delta_T-1}\left\{1+o_p(1)\right\} \\
&= \frac{X_{\tau_e}}{\tau_w c}\left(T^\alpha\delta_T^{\tau_f-\tau_e}+c\delta_T^{\tau_f-\tau_e}-T^\alpha\right)\left\{1+o_p(1)\right\} \\
&= \frac{T^\alpha\delta_T^{\tau_f-\tau_e}}{\tau_w c}X_{\tau_e}\left\{1+o_p(1)\right\} \qquad (29)\\
&\sim_a T^{\alpha-1/2}\delta_T^{\tau_f-\tau_e}\frac{1}{r_w c}B(r_e).
\end{aligned}$$

The third term is

$$\frac{1}{\tau_w}\sum_{j=\tau_f+1}^{\tau_2} X_j \sim_a T^{1/2}\frac{r_2-r_f}{r_w}\left\{\int_{r_f}^{r_2}\left[B(s)-B(r_f)\right]ds - B(r_e)\right\} \text{ from equation (28).}$$

Furthermore, we know

$$\frac{T^{\alpha-1/2}\delta_T^{\tau_f-\tau_e}}{T^{1/2}} = \frac{e^{c\left(r_f-r_e\right)T^{1-\alpha}}}{T^{1-\alpha}} > 1.$$

This implies 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 $\tau_w^{-1}\sum_{j=\tau_f+1}^{\tau_2} X_j$. Therefore,

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

□

Lemma A3. *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* $\tau_2 \in B$,

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

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

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

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

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

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

$$\tilde{X}_t = X_t - \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} X_j = -\frac{T^\alpha \delta_T^{\tau_2-\tau_e}}{\tau_w c} X_{\tau_e} \{1 + o_p(1)\}, \tag{30}$$

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

$$T^{-1/2} X_t \sim_a B(p) \text{ from Lemma A1}$$
$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j \sim_a T^{\alpha-1/2} \delta_T^{\tau_2-\tau_e} \frac{1}{r_w c} B(r_e) \text{ from Lemma A2}$$

and

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

If $t \in B$,

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

(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 $t \in N_1$,

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

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

$$X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} \left[B(p) - B(r_f) + B(r_e) \right] \text{ from Lemma A1}$$
$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j \sim_a T^{\alpha - 1/2} \delta_T^{\tau_f - \tau_1} \frac{1}{r_w c} B(r_e) \text{ from Lemma A2}$$

and

$$\frac{T^{\alpha-1/2} \delta_T^{\tau_f - \tau_1}}{T^{1/2}} = \frac{\delta_T^{\tau_f - \tau_1}}{T^{1-\alpha}} = \frac{e^{c\left(r_f - r_1\right) T^{1-\alpha}}}{T^{1-\alpha}} > 1.$$

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

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

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

$$X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} B(p) \text{ from Lemma A1}$$
$$\frac{1}{\tau_w} \sum_{j=\tau_1}^{\tau_2} X_j \sim_a T^{\alpha - 1/2} \delta_T^{\tau_f - \tau_e} \frac{1}{r_w c} B(r_e) \text{ from Lemma A2}$$

and

$$\frac{T^{\alpha-1/2} \delta_T^{\tau_f - \tau_e}}{T^{1/2}} > 1.$$

If $t \in B$,

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

If $t \in N_1$,

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

since $X_{t=\lfloor Tp \rfloor} \sim_a T^{1/2} \left[B(p) - B(r_f) + B(r_e) \right]$ (from Lemma A1). $\square$

Lemma A4. *The sample variance terms involving $\tilde{X}_t$ behave as follows.*

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

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

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

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

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

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

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

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

The first term is

$$\begin{aligned}
\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2 &= \sum_{j=\tau_1}^{\tau_e-1} \frac{T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)}}{\tau_w^2 c^2} X_{\tau_e}^2 \left\{1 + o_p(1)\right\} \text{ from Lemma A3} \\
&= \frac{\tau_e-\tau_1}{\tau_w^2 c^2} T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)} X_{\tau_e}^2 \left\{1 + o_p(1)\right\} \\
&\sim_a \frac{r_e-r_1}{r_w^2 c} T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)} B(r_e) .
\end{aligned}$$

Given that

$$\begin{aligned}
\sum_{j=\tau_e}^{\tau_2} \delta_T^{2(j-1-\tau_e)} &= \frac{\delta_T^{2(\tau_2-\tau_e)} - \delta_T^{-2}}{\delta_T^2 - 1} = \frac{T^{\alpha}\delta_T^{2(\tau_2-\tau_e)}}{2c} \left\{1 + o_p(1)\right\} \\
\sum_{j=\tau_e}^{\tau_2} \delta_T^{j-1-\tau_e} &= \frac{\delta_T^{\tau_2-\tau_e} - \delta_T^{-1}}{\delta_T - 1} = \frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{c} \left\{1 + o_p(1)\right\} ,
\end{aligned}$$

the second term

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

$$
\begin{aligned}
&= \sum_{j=\tau_e}^{\tau_2} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_2-\tau_e}}{\tau_w c} \right]^2 X_{\tau_e}^2 \{1 + o_p(1)\} \\
&= \sum_{j=\tau e}^{\tau_2} \left[\delta_T^{2(j-1-\tau_e)} - 2\delta_T^{j-1-\tau_e} \frac{T^\alpha \delta_T^{\tau_2-\tau_e}}{\tau_w c} + \frac{T^{2\alpha} \delta_T^{2(\tau_2-\tau_e)}}{\tau_w^2 c^2} \right] X_{\tau_e}^2 \{1 + o_p(1)\} \\
&= \left[\frac{T^\alpha \delta_T^{2(\tau_2-\tau_e)}}{2c} - 2\frac{T^{2\alpha-1} \delta_T^{2(\tau_2-\tau_e)}}{r_w c^2} + \frac{r_2 - r_e + \frac{1}{T}}{r_w^2 c^2} T^{2\alpha-1} \delta_T^{2(\tau_2-\tau_e)} \right] X_{\tau_e}^2 \{1 + o_p(1)\} \\
&= \frac{T^\alpha \delta_T^{2(\tau_2-\tau_e)}}{2c} X_{\tau_e}^2 \{1 + o_p(1)\} \ \text{(since } \alpha > 2\alpha - 1) \\
&\sim_a \frac{T^{1+\alpha} \delta_T^{2(\tau_2-\tau_e)}}{2c} B(r_e)^2 .
\end{aligned}
$$

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

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

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

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

Given that

$$
\begin{aligned}
\sum_{j=\tau_1}^{\tau_f} \delta_T^{2(j-1-\tau_e)} &= \frac{T^\alpha \left[\delta_T^{2\left(\tau_f-\tau_e\right)} - \delta_T^{2(\tau_1-\tau_e-1)} \right]}{2c + c^2 T^{-\alpha}} = \frac{T^\alpha \delta_T^{2\left(\tau_f-\tau_e\right)}}{2c} \{1 + o_p(1)\} \\
\sum_{j=\tau_1}^{\tau_f} \delta_T^{j-1-\tau_e} &= \frac{T^\alpha \left[\delta_T^{\tau_f-\tau_e} - \delta_T^{\tau_1-\tau_e-1} \right]}{c} = \frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{c} \{1 + o_p(1)\} ,
\end{aligned}
$$

the first term is

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

$$= \left[\frac{T^{\alpha}\delta_T^{2(\tau_f - \tau_e)}}{2c} - 2\frac{T^{\alpha}\delta_T^{\tau_f - \tau_1}}{\tau_w c}\frac{T^{\alpha}\delta_T^{\tau_f - \tau_e}}{c} + \frac{\tau_f - \tau_1 + 1}{\tau_w^2 c^2} T^{2\alpha}\delta_T^{2(\tau_f - \tau_1)} \right] X_{\tau_e}^2 \{1 + o_p(1)\}$$

$$= \left[\frac{T^{\alpha}\delta_T^{2(\tau_f - \tau_e)}}{2c} - 2\frac{\delta_T^{(\tau_f - \tau_1)+(\tau_f - \tau_e)}}{T^{1-2\alpha} r_w c^2} + \frac{r_f - r_1 + \frac{1}{T}}{T^{1-2\alpha} r_w^2 c^2}\delta_T^{2(\tau_f - \tau_1)} \right] X_{\tau_e}^2 \{1 + o_p(1)\}$$

$$= \frac{T^{\alpha}\delta_T^{2(\tau_f - \tau_e)}}{2c} X_{\tau_e}^2 \{1 + o_p(1)\} \text{ (since } \alpha > 2\alpha - 1 \text{ and } \tau_f - \tau_e > \tau_f - \tau_1)$$

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

The second term is

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

$$= \frac{\tau_2 - \tau_f}{\tau_w^2 c^2} T^{2\alpha}\delta_T^{2(\tau_f - \tau_1)} X_{\tau_e}^2 \{1 + o_p(1)\}$$

$$\sim_a \frac{r_2 - r_f}{r_w^2 c^2} T^{2\alpha}\delta_T^{2(\tau_f - \tau_1)} B(r_e)^2 .$$

Since $1 + \alpha > 2\alpha$, $\sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2$ dominates $\sum_{j=\tau_f+1}^{\tau_2} \tilde{X}_{j-1}^2$. Therefore,

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

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

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

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

The first term is

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

$$= \frac{\tau_e - \tau_1}{\tau_w^2 c^2} T^{2\alpha}\delta_T^{2(\tau_f - \tau_e)} X_{\tau_e}^2 \{1 + o_p(1)\}$$

$$\sim_a \frac{r_e - r_1}{r_w^2 c^2} T^{2\alpha} \delta_T^{2\left(\tau_f - \tau_e\right)} B\left(r_e\right)^2 .$$

Given that

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

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

the second term

$$\begin{aligned}
&\sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2 \\
&= \sum_{j=\tau_e}^{\tau_f} \left[\delta_T^{j-1-\tau_e} - \frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{\tau_w c}\right]^2 X_{\tau_e}^2 \left\{1 + o_p(1)\right\} \\
&= \left[\frac{T^\alpha \delta_T^{2\left(\tau_f-\tau_e\right)}}{2c} - 2\frac{\delta_T^{2\left(\tau_f-\tau_e\right)}}{T^{1-2\alpha} r_w c^2} + \frac{r_f - r_e + \frac{1}{T}}{T^{1-2\alpha} r_w^2 c^2} \delta_T^{2\left(\tau_f-\tau_e\right)}\right] X_{\tau_e}^2 \left\{1 + o_p(1)\right\} \\
&= \frac{T^\alpha \delta_T^{2\left(\tau_f-\tau_e\right)}}{2c} X_{\tau_e}^2 \left\{1 + o_p(1)\right\} \ \text{(since } \alpha > 2\alpha - 1\text{)} \\
&\sim_a \frac{T^{\alpha+1} \delta_T^{2\left(\tau_f-\tau_e\right)}}{2c} B\left(r_e\right)^2 .
\end{aligned}$$

The third term is

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

Since $1 + \alpha > 2\alpha$, $\sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2$ dominates the other two terms. Therefore,

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

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

$\square$

Lemma A5. *The sample covariance of $\tilde{X}_t$ and ε_t behaves as follows.*

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

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

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

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

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

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

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

$$\begin{aligned}
\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\varepsilon_j &= \sum_{j=\tau_1}^{\tau_e-1} -\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c} X_{\tau_e}\varepsilon_j \left\{1 + o_p(1)\right\} \\
&= -\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c} X_{\tau_e} \sum_{j=\tau_1}^{\tau_e-1} \varepsilon_j \left\{1 + o_p(1)\right\} \\
&= -\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{r_w c} \left(T^{-1/2} X_{\tau_e}\right) \left(T^{-1/2} \sum_{j=\tau_1}^{\tau_e-1} \varepsilon_j\right) \left\{1 + o_p(1)\right\} \\
&\sim_a -\frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{r_w c} B\left(r_e\right)\left[B\left(r_e\right) - B\left(r_1\right)\right].
\end{aligned}$$

The second term is

$$
\begin{aligned}
&\sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \\
&= \sum_{j=\tau_e}^{\tau_2} \left[\delta_T^{j-1-\tau_e} - \frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c}\right] X_{\tau_e}\varepsilon_j \left\{1+o_p(1)\right\} \\
&= \left[\sum_{j=\tau_e}^{\tau_2} \delta_T^{j-1-\tau_e}\varepsilon_j - \frac{T^{\alpha}\delta_T^{\tau_2-\tau_e}}{\tau_w c}\sum_{j=\tau_e}^{\tau_2}\varepsilon_j\right] X_{\tau_e}\left\{1+o_p(1)\right\} \\
&= \left[T^{\alpha/2}\delta_T^{\tau_2-\tau_e}\left(\frac{1}{T^{\alpha/2}}\sum_{j=\tau_e}^{\tau_2}\delta_T^{-(\tau_2-j+1)}\varepsilon_j\right) - \frac{\delta_T^{\tau_2-\tau_e}}{T^{1/2-\alpha}r_w c}\left(\frac{1}{\sqrt{T}}\sum_{j=\tau_e}^{\tau_2}\varepsilon_j\right)\right] X_{\tau_e}\left\{1+o_p(1)\right\} \\
&= T^{\alpha/2}\delta_T^{\tau_2-\tau_e}\left(T^{-\alpha/2}\sum_{j=\tau_e}^{\tau_2}\delta_T^{-(\tau_2-j+1)}\varepsilon_j\right) X_{\tau_e}\left\{1+o_p(1)\right\} \text{ (since } \alpha/2>\alpha-1/2) \\
&\sim_a T^{(\alpha+1)/2}\delta_T^{\tau_2-\tau_e}X_c B(r_e).
\end{aligned}
$$

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

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

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

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

The first term is

$$
\begin{aligned}
&\sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j \\
&= \sum_{j=\tau_1}^{\tau_f} \left[\delta_T^{j-1-\tau_e} - \frac{T^{\alpha}\delta_T^{\tau_f-\tau_1}}{\tau_w c}\right] X_{\tau_e}\varepsilon_j \left\{1+o_p(1)\right\} \\
&= \left[\sum_{j=\tau_1}^{\tau_f} \delta_T^{j-1-\tau_e}\varepsilon_j - \frac{T^{\alpha}\delta_T^{\tau_f-\tau_1}}{\tau_w c}\sum_{j=\tau_1}^{\tau_f}\varepsilon_j\right] X_{\tau_e}\left\{1+o_p(1)\right\} \\
&= \left[T^{\alpha/2}\delta_T^{\tau_f-\tau_e}\left(\frac{1}{T^{\alpha/2}}\sum_{j=\tau_1}^{\tau_f}\delta_T^{-\left(\tau_f-j+1\right)}\varepsilon_j\right) - \frac{T^{\alpha+1/2}\delta_T^{\tau_f-\tau_1}}{\tau_w c}\left(\frac{1}{\sqrt{T}}\sum_{j=\tau_1}^{\tau_f}\varepsilon_j\right)\right] X_{\tau_e}\left\{1+o_p(1)\right\}
\end{aligned}
$$

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

The second term is

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

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

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

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

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

The first term is

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

The second term is

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

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

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

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

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

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

The third term is

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

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

$$\sim_a -\frac{T^\alpha \delta_T^{\tau_f-\tau_e}}{r_w c} B(r_e) \left[B(r_2) - B(r_f) \right].$$

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

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

□

Lemma A6. *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 $\tau_2 \in B$,*

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

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

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

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

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

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

$$\begin{aligned}
\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\left(X_j - \delta_T X_{j-1}\right) &= \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\left(X_j - X_{j-1} + X_{j-1} - \delta_T X_{j-1}\right) \\
&= \sum_{j=\tau_e}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\left(\varepsilon_j - \frac{c}{T^\alpha} X_{j-1}\right) \\
&= \sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j - \frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}X_{j-1}. \qquad (31)
\end{aligned}$$

The first term is

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_2-\tau_e} X_c B\left(r_e\right) \text{ (from Lemma A5).}$$

The second term is

$$\begin{aligned}
&\frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}X_{j-1} \\
&= \frac{c}{T^\alpha}\sum_{j=\tau_1}^{\tau_e-1} -\frac{T^\alpha \delta_T^{\tau_2-\tau_e}}{\tau_w c} X_{\tau_e} X_{j-1}\left\{1 + o_p(1)\right\} \\
&= -\frac{\delta_T^{\tau_2-\tau_e}}{\tau_w} X_{\tau_e} \sum_{j=\tau_1}^{\tau_e-1} X_{j-1}\left\{1 + o_p(1)\right\} \\
&= -\frac{\tau_e-\tau_1}{\tau_w} T\delta_T^{\tau_2-\tau_e}\left(T^{-1/2}X_{\tau_e}\right)\left[\frac{1}{\tau_e-\tau_1}\sum_{j=\tau_1}^{\tau_e-1}\left(T^{-1/2}X_{j-1}\right)\right]\left\{1 + o_p(1)\right\} \\
&\sim_a -\frac{r_e - r_1}{r_w} T\delta_T^{\tau_2-\tau_e} B\left(r_e\right)\int_{r_1}^{r_e} B(s)\,ds.
\end{aligned}$$

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

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

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

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

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

The first term is

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a T^{(\alpha+1)/2}\delta_T^{\tau_f-\tau_e} X_c B(r_e) \text{ (from Lemma A5).}$$

The second term is

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

due to the fact that

$$\frac{\delta_T^{\tau_f-\tau_e}}{T^{\alpha-1}\delta_T^{\tau_f-\tau_1}} = T^{1-\alpha}\delta_T^{\tau_1-\tau_e} > 1.$$

The third term is

$$\begin{aligned}
&\frac{c}{T^\alpha}\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}X_{j-1} \\
&= \frac{c}{T^\alpha}\sum_{j=\tau_f+2}^{\tau_2} -\frac{T^\alpha \delta_T^{\tau_f-\tau_1}}{\tau_w c} X_{\tau_e} X_{j-1}\{1 + o_p(1)\} \\
&= -\frac{\delta_T^{\tau_f-\tau_1}}{\tau_w} X_{\tau_e} \sum_{j=\tau_f+2}^{\tau_2} X_{j-1}\{1 + o_p(1)\} \\
&= -\frac{\tau_2 - \tau_f - 1}{\tau_w} T\delta_T^{\tau_f-\tau_1}\left(T^{-1/2}X_{\tau_e}\right)\left(\frac{1}{\tau_2 - \tau_f - 1}\sum_{j=\tau_f+2}^{\tau_2} T^{-1/2}X_{j-1}\right)\{1 + o_p(1)\}
\end{aligned}$$

$$\sim_a -\frac{r_2 - r_f}{r_w} T\delta_T^{\tau_f - \tau_1} B(r_e) \int_{r_f}^{r_2} B(s)\, ds.$$

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

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

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

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

The first term is

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a T^{(\alpha+1)/2} \delta_T^{\tau_f - \tau_e} X_c B(r_e) \text{ (from Lemma A5).}$$

The second term is

$$\begin{aligned}
&\frac{c}{T^\alpha} \sum_{j=\tau_1}^{\tau_e - 1} \tilde{X}_{j-1} X_{j-1} \\
&= \frac{c}{T^\alpha} \sum_{j=\tau_1}^{\tau_e - 1} -\frac{T^\alpha \delta_T^{\tau_f - \tau_e}}{\tau_w c} X_{\tau_e} X_{j-1} \{1 + o_p(1)\} \\
&= -\frac{\delta_T^{\tau_f - \tau_e}}{\tau_w} X_{\tau_e} \sum_{j=\tau_1}^{\tau_e - 1} X_{j-1} \{1 + o_p(1)\}
\end{aligned}$$

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

$$\sim_a -\frac{r_e - r_1}{r_w} T\delta_T^{\tau_f - \tau_e} B(r_e) \int_{r_1}^{r_e} B(s)\, ds$$

The third term is

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

due to the fact that

$$\frac{\delta_T^{\tau_f - \tau_e}}{T^{\alpha - 1} \delta_T^{\tau_f - \tau_e}} = T^{1-\alpha} > 1.$$

The fourth term is

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

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

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

□

A.2: Test Asymptotics and Proofs of Theorems 1-3.

The fitted regression model for the subperiod unit root test is

$$X_t = \hat{\alpha}_{r_1,r_2} + \hat{\delta}_{r_1,r_2} X_{t-1} + \hat{\varepsilon}_t, \quad t \in [\lfloor Tr_1 \rfloor, \lfloor Tr_2 \rfloor] \tag{32}$$

The intercept $\hat{\alpha}_{r_1,r_2}$ and slope coefficient $\hat{\delta}_{r_1,r_2}$ are obtained using data over the subperiod $[r_1, r_2]$. We calculate the asymptotic distribution of the unit root statistic under the alternative hypothesis. Based on Lemma A4 and Lemma A6, we can obtain the limit distribution of $\hat{\delta}_{r_1,r_2} - \delta_T$ using

$$\hat{\delta}_T - \delta_T = \frac{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \left(X_j - \delta_T X_{j-1} \right)}{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2}.$$

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

$$\hat{\delta}_{r_1,r_2} - \delta_T \sim_a T^{-\alpha} \delta_T^{-(\tau_2 - \tau_{ie})} \frac{\frac{r_e - r_1}{r_w} \int_{r_1}^{r_e} B(s)\, ds}{B(r_e)};$$

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

$$\hat{\delta}_{r_1,r_2} - \delta_T \sim_a -2T^{-\alpha} c;$$

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

$$\hat{\delta}_{r_1,r_2} - \delta_T \sim_a -2T^{-\alpha} c.$$

A.2.1: Limit Behavior of the recursive unit root statistics

The asymptotic distributions of the unit root coefficient Z-statistics are as follows: (1) When $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\begin{aligned}
DF_{r_1,r_2}^z &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1 \right) = \tau_w \left(\delta_T - 1 \right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T \right) \\
&= \tau_w \left(\delta_T - 1 \right) + o_p \left(r_w \frac{T^{1-\alpha}}{\delta_T^{\tau_2 - \tau_e}} \right) \\
&= \frac{\tau_w c}{T^\alpha} + o_p \left(r_w \frac{T^{1-\alpha}}{e^{c(r_2 - r_e) T^{1-\alpha}}} \right) \\
&= r_w c T^{1-\alpha} + o_p(1) \to \infty;
\end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= \frac{\tau_w c}{T^\alpha} + o_p\left(\frac{\tau_w}{T^\alpha}\right) \\ &= (c - 2c)\frac{\tau_w}{T^\alpha} \\ &= -c r_w T^{1-\alpha} \to -\infty; \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= \frac{\tau_w c}{T^\alpha} + o_p\left(\frac{\tau_w}{T^\alpha}\right) \\ &= (c - 2c)\frac{\tau_w}{T^\alpha} \\ &= -c r_w T^{1-\alpha} \to -\infty. \end{aligned}$$

This implies that when $\tau_1 \in N_0$ and $\tau_2 \in B$,

$$\hat{\delta}_{r_1,r_2} - 1 \sim_a T^{-\alpha} c \text{ and } T^\alpha \left(\hat{\delta}_{r_1,r_2} - 1\right) \xrightarrow{L} c;$$

and for the other two cases,

$$\hat{\delta}_{r_1,r_2} - 1 \sim_a -T^{-\alpha} c \text{ and } T^\alpha \left(\hat{\delta}_{r_1,r_2} - 1\right) \xrightarrow{L} -c.$$

To obtain the asymptotic distributions of the unit root t-statistics, we need first to estimate the standard error of $\hat{\delta}_{r_1,r_2}$. (1)When $\tau_1 \in N_0$ and $\tau_2 \in B$,

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

$$
\begin{aligned}
&- 2\left(\hat{\delta}_{r_1,r_2} - 1\right)\tau_w^{-1}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}\varepsilon_j - 2\left(\hat{\delta}_{r_1,r_2} - \delta_T\right)\tau_w^{-1}\sum_{j=\tau_e}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j \\
&= \left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2\tau_w^{-1}\sum_{j=\tau_e}^{\tau_2}\tilde{X}_{j-1}^2 \\
&\sim_a \frac{2c}{T^\alpha}\frac{(r_e - r_1)^2}{r_w^3}\left[\int_{r_1}^{r_e} B(s)\,ds\right]^2.
\end{aligned}
$$

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

$$
\begin{aligned}
\left(\hat{\delta}_{r_1,r_2} - 1\right)^2{\tau_w}^{-1}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}^2 &= O_p\left(T^{-2\alpha}\right)O_p\left(T^{2\alpha-1}\delta_T^{2(\tau_2-\tau_e)}\right) = O_p\left(T^{-1}\delta_T^{2(\tau_2-\tau_e)}\right), \\
\left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2\tau_w^{-1}\sum_{j=\tau_e}^{\tau_2}\tilde{X}_{j-1}^2 &= O_p\left(\frac{1}{T^{2\alpha}\delta_T^{2(\tau_2-\tau_e)}}\right)O_p\left(T^\alpha\delta_T^{2(\tau_2-\tau_e)}\right) = O_p\left(T^{-\alpha}\right), \\
2\left(\hat{\delta}_{r_1,r_2} - 1\right)\tau_w^{-1}\sum_{j=\tau_1}^{\tau_e-1}\tilde{X}_{j-1}\varepsilon_j &= O_p\left(T^{-\alpha}\right)O_p\left(\frac{\delta_T^{\tau_2-\tau_e}}{T^{1-\alpha}}\right) = O_p\left(T^{-1}\delta_T^{\tau_2-\tau_e}\right), \\
2\left(\hat{\delta}_{r_1,r_2} - \delta_T\right)\tau_w^{-1}\sum_{j=\tau_e}^{\tau_2}\tilde{X}_{j-1}\varepsilon_j &= O_p\left(\frac{1}{T^\alpha\delta_T^{\tau_2-\tau_e}}\right)O_p\left(\frac{\delta_T^{\tau_2-\tau_e}}{T^{(1-\alpha)/2}}\right) = O_p\left(T^{-(1+3\alpha)/2}\right).
\end{aligned}
$$

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

$$
\begin{aligned}
&\tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2}\tilde{X}_{\tau_f} \\
&= \frac{\delta_T^{\tau_f-\tau_1}}{r_w c T^{1-\alpha}}X_{\tau_e} - \tilde{X}_{\tau_f} - \left[\hat{\delta}_{r_1,r_2} - 1\right]\tilde{X}_{\tau_f} \\
&= O_p\left(T^{\alpha-1/2}\delta_T^{\tau_f-\tau_1}\right) - O_p\left(T^{1/2}\delta_T^{\tau_f-\tau_e}\right) - O_p\left(T^{-\alpha}\right)O_p\left(T^{1/2}\delta_T^{\tau_f-\tau_e}\right) \\
&= -\tilde{X}_{\tau_f} = -\delta_T^{\tau_f-\tau_e}X_{\tau_e}\{1 + o_p(1)\},
\end{aligned}
$$

using the fact that

$$
\tilde{X}_{\tau_f} = \left[\delta_T^{\tau_f-\tau_e} - \frac{\delta_T^{\tau_f-\tau_1}}{r_w c T^{1-\alpha}}\right]X_{\tau_e}\{1 + o_p(1)\} = \delta_T^{\tau_f-\tau_e}X_{\tau_e}\{1 + o_p(1)\}.
$$

Therefore,

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

$$
\begin{aligned}
&= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \left(\tilde{X}_j - \hat{\delta}_{r_1,r_2} \tilde{X}_{j-1} \right)^2 \\
&= \tau_w^{-1} \left\{ \sum_{j=\tau_f+2}^{\tau_2} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - 1 \right) \tilde{X}_{j-1} \right]^2 + \sum_{j=\tau_1}^{\tau_f} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - \delta_T \right) \tilde{X}_{j-1} \right]^2 \right. \\
&\left. + \left[\tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2} \tilde{X}_{\tau_f}^2 - \varepsilon_{\tau_f+1} + \varepsilon_{\tau_f+1} \right]^2 \right\} \\
&= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \varepsilon_j^2 + \left(\hat{\delta}_{r_1,r_2} - 1 \right)^2 \tau_w{}^{-1} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}^2 + \left(\hat{\delta}_{r_1,r_2} - \delta_T \right)^2 \tau_w^{-1} \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1}^2 \\
&- 2 \left(\hat{\delta}_{r_1,r_2} - 1 \right) \tau_w^{-1} \sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1} \varepsilon_j - 2 \left(\hat{\delta}_{r_1,r_2} - \delta_T \right) \tau_w^{-1} \sum_{j=\tau_1}^{\tau_f} \tilde{X}_{j-1} \varepsilon_j + \tau_w^{-1} \tilde{X}_{\tau_f}^2 \\
&= \tau_w^{-1} \tilde{X}_{\tau_f}^2 = \tau_w^{-1} \delta_T^{2\left(\tau_f - \tau_e\right)} X_{\tau_e}^2 \left\{ 1 + o_p\left(1\right) \right\} \\
&\sim_a \frac{1}{r_w} \delta_T^{2\left(\tau_f - \tau_e\right)} B\left(r_e\right)^2 .
\end{aligned}
$$

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

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

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

$$
\begin{aligned}
&\tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2} \tilde{X}_{\tau_f} - \varepsilon_{\tau_f+1} \\
&= -\frac{\delta_T^{\tau_f-\tau_e}}{r_w c T^{1-\alpha}} X_{\tau_e} - \tilde{X}_{\tau_f} - \left[\hat{\delta}_{r_1,r_2} - 1 \right] \tilde{X}_{\tau_f} \\
&= -O_p\left(T^{\alpha-1/2} \delta_T^{\tau_f-\tau_e} \right) - O_p\left(T^{1/2} \delta_T^{\tau_f-\tau_e} \right) - O_p\left(T^{-\alpha}\right) O_p\left(T^{1/2} \delta_T^{\tau_f-\tau_e} \right)
\end{aligned}
$$

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

using the fact that

$$\tilde{X}_{\tau_f} = \left[\delta_T^{\tau_f-\tau_e} - \frac{\delta_T^{\tau_f-\tau_e}}{r_w c T^{1-\alpha}}\right] X_{\tau_e} \{1 + o_p(1)\} = \delta_T^{\tau_f-\tau_e} X_{\tau_e} \{1 + o_p(1)\}.$$

$$\begin{aligned}
&Var\left(\hat{\delta}_{r_1,r_2}\right) \\
&= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \left(\tilde{X}_j - \hat{\delta}_{r_1,r_2}\tilde{X}_{j-1}\right)^2 \\
&= \tau_w^{-1} \left\{ \sum_{j=\tau_f+2}^{\tau_2} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - 1\right)\tilde{X}_{j-1}\right]^2 + \sum_{j=\tau_1}^{\tau_e-1} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - 1\right)\tilde{X}_{j-1}\right]^2 \right. \\
&\left. + \sum_{j=\tau_e}^{\tau_f} \left[\varepsilon_j - \left(\hat{\delta}_{r_1,r_2} - \delta_T\right)\tilde{X}_{j-1}\right]^2 + \tilde{X}_{\tau_f+1} - \hat{\delta}_{r_1,r_2}\tilde{X}_{\tau_f}^2 \right\}^2 \\
&= \tau_w^{-1} \sum_{j=\tau_1}^{\tau_2} \varepsilon_j^2 + \left(\hat{\delta}_{r_1,r_2} - 1\right)^2 \tau_w{}^{-1} \left[\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}^2 + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}^2\right] + \left(\hat{\delta}_{r_1,r_2} - \delta_T\right)^2 \tau_w^{-1} \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}^2 \\
&- 2\left(\hat{\delta}_{r_1,r_2} - 1\right)\tau_w^{-1} \left[\sum_{j=\tau_f+2}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j + \sum_{j=\tau_1}^{\tau_e-1} \tilde{X}_{j-1}\varepsilon_j\right] - 2\left(\hat{\delta}_{r_1,r_2} - \delta_T\right)\tau_w^{-1} \sum_{j=\tau_e}^{\tau_f} \tilde{X}_{j-1}\varepsilon_j + \tau_w^{-1}\tau_f^2 \\
&= \tau_w^{-1}\tilde{X}_{\tau_f}^2 = \frac{\delta_T^{2(\tau_f-\tau_e)}}{\tau_w} X_{\tau_e}^2 \{1 + o_p(1)\} \sim_a \frac{\delta_T^{2(\tau_f-\tau_e)}}{r_w} B(r_e)^2.
\end{aligned}$$

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

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

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

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

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

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

The asymptotic distributions of the t-statistic

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

can be calculated as follows:

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

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

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

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

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

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

Taken together with (11) and (12), these results establish the limit behavior of the unit root statistics DF_r and $BSDF_r\left(r_0\right)$ in Theorem 1 (see also (33) below).

A.2.2: The PWY strategy

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

$$\hat{r}^e = \inf_{r\in[r_0,1]}\left\{r : DF_r > cv^{\beta_T}\right\} \text{ and } \hat{r}^f = \inf_{r\in[\hat{r}_e+L_T,1]}\left\{r : DF_r < cv^{\beta_T}\right\}.$$

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

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

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

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

$$\lim_{T\to\infty} \Pr\left\{DF_r > cv^{\beta_T}\right\} = \Pr\left\{F_r(W) = \infty\right\} = 0.$$

If $r \in B$, $\lim_{T\to\infty} \Pr\left\{DF_r > cv^{\beta_T}\right\} = 1$ provided that $\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_e}} \to 0$. If $r \in N_1$, $\lim_{T\to\infty} \Pr\left\{DF_r < cv_r^{\beta_T}\right\} = \lim_{T\to\infty} \Pr\left\{-T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} < cv^{\beta_T}\right\} = 1$.

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 $\Pr\left\{DF_{r_e+a_\eta} > cv^{\beta_T}\right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{DF_{r_f-a_\gamma} > cv^{\beta_T}\right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary and $\Pr\left\{\hat{r}_e < r_e\right\} \to 0$ 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

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_e}} \to 0.$$

Therefore, the PWY date detectors $\hat{r}_e$ and $\hat{r}_f$ are consistent estimators of r_e and r_f. This proves Theorem 2.

A.2.3: The PSY algorithm

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

$$\hat{r}^e = \inf_{r\in[r_0,1]} \left\{r : BSDF_r(r_0) > scv^{\beta_T}\right\} \text{ and } \hat{r}^f = \inf_{r\in[\hat{r}_e+L_T,1]} \left\{r_2 : BSDF_r(r_0) < scv^{\beta_T}\right\}.$$

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

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

$$BSDF_r(r_0) \sim_a \begin{cases} F_r(W, r_0) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_e} \sup_{r_1\in[0,r-r_0]} \left\{\frac{r_w^{3/2} B(r_e)}{2(r_e-r_1)\int_{r_1}^{r_e} B(s)ds}\right\} & \text{if } r \in B \\ -T^{(1-\alpha)/2} \sup_{r_1\in[0,r-r_0]} \left\{\left(\frac{1}{2}cr_w\right)^{1/2}\right\} \to \infty & \text{if } r \in N_1 \end{cases}. \quad (33)$$

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

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

If $r \in B$, $\lim_{T\to\infty} \Pr\left\{BSDF_r(r_0) > scv^{\beta_T}\right\} = 1$ provided that $\frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_e}} \to 0$. If $r \in N_1$, $\lim_{T\to\infty} \Pr\left\{BSDF_r(r_0) < scv_r^{\beta_T}\right\} = 1$.

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

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

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

$$\frac{1}{scv^{\beta_T}} + \frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau_2-\tau_e}} \to 0.$$

Therefore, the PSY date detectors $\hat{r}_e$ and $\hat{r}_f$ are consistent estimators of r_e and r_f. This proves Theorem 3.

APPENDIX B. The Dating Algorithms (two bubbles)

Section B.1 provides preliminary results that characterize the limit behavior of the regression components over subperiods of the data. Section B.2 provides test asymptotics and gives proofs of Theorems 4-9 which describe the consistency properties of the PWY, PSY and sequential PWY dating strategies.

B.1: 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 $N_0 = [1, \tau_{1e})$, $N_1 = (\tau_{1f}, \tau_{2e})$, $N_2 = (\tau_{2f}, \tau_T]$, where $\tau = \lfloor Tr \rfloor$ is the last observation of the sample.

We use the data generating process

$$X_t = \begin{cases} X_{t-1} + \varepsilon_t & \text{for } t \in N_0 \\ \delta_T X_{t-1} + \varepsilon_t & \text{for } t \in B_i \text{ with } i = 1, 2 \\ X^*_{\tau_{if}} + \sum_{k=\tau_{if}+1}^{t} \varepsilon_k & \text{for } t \in N_i \text{ with } i = 1, 2 \end{cases}, \tag{34}$$

where $\delta_T = 1 + cT^{-\alpha}$ with $c > 0$ and $\alpha \in (0, 1)$, $\varepsilon_t \overset{iid}{\sim} \left(0, \sigma^2\right)$ and $X^*_{\tau_{if}} = X_{\tau_{ie}} + X^*$ with $X^* = O_p(1)$ for $i = 1, 2$. We state the following lemmas whose proofs follow arguments closely related to those given in the proofs of Lemmas A1-A6. They are provided in full in the technical supplement (Phillips, Shi and Yu, 2013c; lemmas S1-S6).

Lemma A7. *Under the data generating process,*

(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 A8. *Under the data generating process,*

(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* $\tau_2 \in N_2$, *if* $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$

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

and if $\tau_{1f} - \tau_{1e} \le \tau_{2f} - \tau_{2e}$

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

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

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

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

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

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

$$\frac{1}{\tau_w}\sum_{j=\tau_1}^{\tau_2} X_j = \frac{T^{\alpha}\delta_T^{\tau_{1f}-\tau_1}}{\tau_w c} X_{\tau_{1e}}\left\{1 + o_p(1)\right\} \sim_a T^{\alpha-1/2}\delta_T^{\tau_{1f}-\tau_1}\frac{1}{r_w c}B(r_1)$$

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

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

(7) *For* $\tau_1 \in N_0$ *and* $\tau_2 \in B_2$, *if* $\tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e}$,

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

and if $\tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e}$,

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

Lemma A9. *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* $i = 1, 2$,

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

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

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

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

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

(4) *For* $\tau_1 \in N_0$ *and* $\tau_2 \in N_2$*, if* $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$

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

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

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

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

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

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

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

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

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

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

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

(7) *For* $\tau_1 \in N_0$ *and* $\tau_2 \in B_2$, *if* $\tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e}$

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

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

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

Lemma A10. *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* $i = 1, 2$,

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

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

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

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

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

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

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

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

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \begin{cases} \frac{T^\alpha \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}}{2c} X_{\tau_{1e}}^2 \left\{1+o_p\left(1\right)\right\} \sim_a \frac{T^{\alpha+1}\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}}{2c} B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_2-\tau_{2e} \\ \frac{T^\alpha \delta_T^{2(\tau_2-\tau_{2e})}}{2c} X_{\tau_{2e}}^2 \left\{1+o_p\left(1\right)\right\} \sim_a T^{\alpha+1}\delta_T^{2(\tau_2-\tau_{2e})}\frac{1}{2c} B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_2-\tau_{2e} \end{cases}.$$

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

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \begin{cases} \frac{T^\alpha \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}}{2c} X_{\tau_{1e}}^2 \left\{1+o_p\left(1\right)\right\} \sim_a \frac{T^{\alpha+1}\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}}{2c} B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_{2f}-\tau_{2e} \\ \frac{T^\alpha \delta_T^{2\left(\tau_{2f}-\tau_{2e}\right)}}{2c} X_{\tau_{2e}}^2 \left\{1+o_p\left(1\right)\right\} \sim_a \frac{T^{\alpha+1}\delta_T^{2\left(\tau_{2f}-\tau_{2e}\right)}}{2c} B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_{2f}-\tau_{2e} \end{cases}.$$

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

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2 = \begin{cases} \frac{T^\alpha \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}}{2c} X_{\tau_{1e}}^2 \left\{1+o_p\left(1\right)\right\} \sim_a \frac{T^{\alpha+1}\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}}{2c} B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_2-\tau_{2e} \\ \frac{T^\alpha \delta_T^{2(\tau_2-\tau_{2e})}}{2c} X_{\tau_{2e}}^2 \left\{1+o_p\left(1\right)\right\} \sim_a \frac{T^{\alpha+1}\delta_T^{2(\tau_2-\tau_{2e})}}{2c} B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_2-\tau_{2e} \end{cases}.$$

Lemma A11. *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* $i = 1, 2$,

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

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

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

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

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

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

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a \begin{cases} T^{(1+\alpha)/2}\delta_T^{\tau_{1f}-\tau_{1e}} X_c B\left(r_{1e}\right) & \text{if } \tau_{1f}-\tau_{1e} > \tau_{2f}-\tau_{2e} \\ T^{(1+\alpha)/2}\delta_T^{\tau_{2f}-\tau_{2e}} X_c B\left(r_{2e}\right) & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_{2f}-\tau_{2e} \end{cases}.$$

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

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}\varepsilon_j \sim_a \begin{cases} T^{(\alpha+1)/2}\delta_T^{\tau_{1f}-\tau_{1e}} X_c B\left(r_{1e}\right) & \text{if } \tau_{1f}-\tau_{1e} > \tau_2-\tau_{2e} \\ T^{(\alpha+1)/2}\delta_T^{\tau_2-\tau_{2e}} X_c B\left(r_{2e}\right) & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_2-\tau_{2e} \end{cases}.$$

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

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

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

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

Lemma A12. *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* $\tau_2 \in N_2$,

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}(X_j - \delta_T X_{j-1}) \sim_a \begin{cases} -T\delta_T^{2(\tau_{1f}-\tau_{1e})} B(r_{1e})^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_{2f}-\tau_{2e} \\ -T\delta_T^{2(\tau_{2f}-\tau_{2e})} B(r_{2e})^2 & \text{if } \tau_{1f}-\tau_{1e} \le \tau_{2f}-\tau_{2e} \end{cases}.$$

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

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

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

$$\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}(X_j - \delta_T X_{j-1}) \sim_a \begin{cases} -T\delta_T^{2(\tau_{1f}-\tau_{1e})} B(r_{1e})^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_{2f}-\tau_{2e} \\ -T\delta_T^{2(\tau_{2f}-\tau_{2e})} B(r_{2e})^2 & \text{if } \tau_{1f}-\tau_{1e} \le \tau_{2f}-\tau_{2e} \end{cases}.$$

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

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

B.2: Test asymptotics and Proofs of Theorems 4-9

The fitted regression model for the recursive unit root tests is

$$X_t = \hat{\alpha}_{r_1,r_2} + \hat{\delta}_{r_1,r_2} X_{t-1} + \hat{\varepsilon}_t,$$

where as in (32) above the intercept $\hat{\alpha}_{r_1,r_2}$ and slope coefficient $\hat{\delta}_{r_1,r_2}$ are obtained using data over the subperiod $[r_1, r_2]$. First, we calculate the asymptotic distribution of the unit root statistic under the alternative hypothesis. Based on Lemma A10 and Lemma A12, we can obtain the limit distribution of $\hat{\delta}_{r_1,r_2} - \delta_T$ using

$$\hat{\delta}_T - \delta_T = \frac{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1} \left(X_j - \delta_T X_{j-1}\right)}{\sum_{j=\tau_1}^{\tau_2} \tilde{X}_{j-1}^2}.$$

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

$$\hat{\delta}_T - \delta_T \sim_a T^{-\alpha} \delta_T^{-(\tau_2 - \tau_{ie})} \frac{\frac{r_{ie} - r_1}{r_w} \int_{r_1}^{r_{ie}} B(s)\, ds}{B(r_{ie})};$$

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

$$\hat{\delta}_T - \delta_T \sim_a -2T^{-\alpha} c;$$

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

$$\hat{\delta}_T - \delta_T \sim_a -2T^{-\alpha} c;$$

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

$$\hat{\delta}_T - \delta_T \sim_a -2T^{-\alpha} c;$$

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

$$\hat{\delta}_T - \delta_T \sim_a \begin{cases} -2T^{-\alpha} c & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ T^{-1} \delta_T^{-(\tau_2 - \tau_{2e}) + (\tau_{1f} - \tau_{1e})} \frac{2B(r_{1e})}{r_w B(r_{2e})} & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases};$$

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

$$\hat{\delta}_T - \delta_T \sim_a -2T^{-\alpha} c;$$

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

$$\hat{\delta}_T - \delta_T \sim_a \begin{cases} -2T^{-\alpha}c & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ T^{-1}\delta_T^{-(\tau_2-\tau_{2e})+\left(\tau_{1f}-\tau_{1e}\right)} \frac{2B(r_{1e})}{r_w B(r_{2e})} & \text{if } \tau_{1f} - \tau_{1e} \leq \tau_2 - \tau_{2e} \end{cases}.$$

The asymptotic distributions of the unit root coefficient Z-statistics are as follows: (1) When $\tau_1 \in N_{i-1}$ and $\tau_2 \in B_i$ with $i = 1, 2$,

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= \tau_w \left(\delta_T - 1\right) + o_p \left(r_w \frac{T^{1-\alpha}}{\delta_T^{\tau_2-\tau_{ie}}} \right) \\ &= \frac{\tau_w c}{T^\alpha} + o_p \left(r_w \frac{T^{1-\alpha}}{e^{c(r_2-r_{ie})T^{1-\alpha}}} \right) \\ &= r_w c T^{1-\alpha} + o_p (1) \to \infty. \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= c r_w T^{1-\alpha} + o_p \left(r_w T^{1-\alpha}\right) \\ &= -c r_w T^{1-\alpha} \to -\infty. \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= c r_w T^{1-\alpha} + o_p \left(r_w T^{1-\alpha}\right) \\ &= -c r_w T^{1-\alpha} \to -\infty. \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= c r_w T^{1-\alpha} + o_p \left(r_w T^{1-\alpha}\right) \\ &= -c r_w T^{1-\alpha} \to -\infty. \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= \begin{cases} c r_w T^{1-\alpha} + o_p\left(r_w T^{1-\alpha}\right) & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ c r_w T^{1-\alpha} + o_p\left(\frac{r_w}{\delta_T^{(\tau_2-\tau_{2e})-\left(\tau_{1f}-\tau_{1e}\right)}}\right) & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases} \\ &= \begin{cases} -c r_w T^{1-\alpha} \to -\infty & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ c r_w T^{1-\alpha} \to \infty & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases}. \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= c r_w T^{1-\alpha} + o_p\left(r_w T^{1-\alpha}\right) \\ &= -c r_w T^{1-\alpha} \to -\infty. \end{aligned}$$

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

$$\begin{aligned} DF^z_{r_1,r_2} &= \tau_w \left(\hat{\delta}_{r_1,r_2} - 1\right) = \tau_w \left(\delta_T - 1\right) + \tau_w \left(\hat{\delta}_{r_1,r_2} - \delta_T\right) \\ &= \begin{cases} c r_w T^{1-\alpha} + o_p\left(r_w T^{1-\alpha}\right) & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ c r_w T^{1-\alpha} + o_p\left(\frac{r_w}{\delta_T^{(\tau_2-\tau_{2e})-\left(\tau_{1f}-\tau_{1e}\right)}}\right) & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases} \\ &= \begin{cases} -c r_w T^{1-\alpha} \to -\infty & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ c r_w T^{1-\alpha} \to \infty & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases}. \end{aligned}$$

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

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

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

$$Var\left(\hat{\delta}_{r_1,r_2}\right) = \tau_w^{-1} \tilde{X}^2_{\tau_{if}} \sim_a \frac{1}{r_w} \delta_T^{2\left(\tau_{if}-\tau_{ie}\right)} B\left(r_{ie}\right)^2.$$

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

$$Var\left(\hat{\delta}_{r_1,r_2}\right) = \tau_w^{-1} \tilde{X}_{\tau_{if}} \sim_a \frac{\delta_T^{2\left(\tau_{if}-\tau_{ie}\right)}}{r_w} B\left(r_{ie}\right)^2.$$

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

$$Var\left(\hat{\delta}_{r_1,r_2}\right) = \begin{cases} \tau_w^{-1}\tilde{X}^2_{\tau_{1f}} \sim_a r_w^{-1}\delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_{2f}-\tau_{2e} \\ \tau_w^{-1}\tilde{X}^2_{\tau_{2f}} \sim_a r_w^{-1}\delta_T^{2\left(\tau_{2f}-\tau_{2e}\right)}B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_{2f}-\tau_{2e} \end{cases}.$$

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

$$Var\left(\hat{\delta}_{r_1,r_2}\right) = \begin{cases} \tau_w^{-1}\tilde{X}^2_{\tau_{1f}} \sim_a \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}r_w^{-1}B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_2-\tau_{2e} \\ \left(\hat{\delta}_{r_1,r_2}-1\right)^2\tau_w^{-1}\sum_{j=\tau_{1f}+2}^{\tau_{2e}-1}\tilde{X}^2_{j-1} \sim_a T^{-1}\delta_T^{2(\tau_2-\tau_{2e})}\frac{r_{2e}-r_{1f}}{r_w^3}B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_2-\tau_{2e} \end{cases}.$$

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

$$Var\left(\hat{\delta}_{r_1,r_2}\right) = \begin{cases} \tau_w^{-1}\tilde{X}^2_{\tau_{1f}} \sim_a \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}\frac{1}{r_w}B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_{2f}-\tau_{2e} \\ \tau_w^{-1}\tilde{X}^2_{\tau_{2f}} \sim_a \delta_T^{2\left(\tau_{2f}-\tau_{2e}\right)}\frac{1}{r_w}B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_{2f}-\tau_{2e} \end{cases}.$$

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

$$\begin{aligned} Var\left(\hat{\delta}_{r_1,r_2}\right) &= \begin{cases} \tau_w^{-1}\tilde{X}^2_{\tau_{1f}} & \text{if } \tau_{1f}-\tau_{1e} > \tau_2-\tau_{2e} \\ \left(\hat{\delta}_{r_1,r_2}-1\right)^2\tau_w^{-1}\left[\sum_{j=\tau_1}^{\tau_{1e}-1}\tilde{X}^2_{j-1}+\sum_{j=\tau_{1f}+2}^{\tau_{2e}-1}\tilde{X}^2_{j-1}\right] & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_2-\tau_{2e} \end{cases} \\ &\sim_a \begin{cases} \delta_T^{2\left(\tau_{1f}-\tau_{1e}\right)}\frac{1}{r_w}B\left(r_{1e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} > \tau_2-\tau_{2e} \\ T^{-1}\delta_T^{2(\tau_2-\tau_{2e})}\frac{r_{1e}-r_1+r_{2e}-r_{1f}}{r_w^3}B\left(r_{2e}\right)^2 & \text{if } \tau_{1f}-\tau_{1e} \leq \tau_2-\tau_{2e} \end{cases}. \end{aligned}$$

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

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

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

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

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

$$DF^t_{r_1,r_2} \sim_a -\left(\frac{1}{2}cr_w\right)^{1/2}T^{(1-\alpha)/2} \to -\infty;$$

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

$$DF^t_{r_1,r_2} \sim_a -\left(\frac{1}{2}cr_w\right)^{1/2}T^{(1-\alpha)/2} \to -\infty;$$

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

$$DF^t_{r_1,r_2} \sim_a -\left(\frac{1}{2}cr_w\right)^{1/2} T^{(1-\alpha)/2} \to -\infty;$$

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

$$DF^t_{r_1,r_2} \sim_a \begin{cases} -\left(\frac{1}{2}cr_w\right)^{1/2} T^{(1-\alpha)/2} \to -\infty & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ \left[\frac{cr_w^3}{2\left(r_{2c}-r_{1f}\right)}\right]^{1/2} T^{1-\alpha/2} \to \infty & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases};$$

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

$$DF^t_{r_1,r_2} \sim_a -\left(\frac{1}{2}cr_w\right)^{1/2} T^{(1-\alpha)/2} \to -\infty;$$

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

$$DF^t_{r_1,r_2} \sim_a \begin{cases} -\left(\frac{1}{2}cr_w\right)^{1/2} T^{(1-\alpha)/2} \to -\infty & \text{if } \tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e} \\ \left[\frac{cr_w^3}{2\left(r_{1e}-r_1+r_{2e}-r_{1f}\right)}\right]^{1/2} T^{1-\alpha/2} \to \infty & \text{if } \tau_{1f} - \tau_{1e} \le \tau_2 - \tau_{2e} \end{cases}.$$

Taken together with (11) and (12), these results establish the limit behavior of the unit root statistics DF_r and $BSDF_r(r_0)$ in the two cases considered in theorems 4 and 5 (see also (36) below).

B.2.1: The PWY 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 DF test are identified as

$$\hat{r}_{1e} = \inf_{r\in[r_0,1]} \left\{r_2 : DF_r > cv^{\beta_T}\right\} \text{ and } \hat{r}_{1f} = \inf_{r\in[\hat{r}_{1e}+L_T,1]} \left\{r_2 : DF_r < cv^{\beta_T}\right\},$$
$$\hat{r}_{2e} = \inf_{r\in\left(\hat{r}_{1f},1\right]} \left\{r_2 : DF_r > cv^{\beta_T}\right\} \text{ and } \hat{r}_{2f} = \inf_{r\in[\hat{r}_{2e}+L_T,1]} \left\{r_2 : DF_r < cv^{\beta_T}\right\}.$$

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

Case I Suppose $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$. The asymptotic distributions of the DF statistic under the alternative hypothesis are

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

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

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

If $r \in B_1$, $\lim_{T\to\infty} \Pr\left\{DF_r > cv^{\beta_T}\right\} = 1$ provided that $\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0$. If $r \in N_1$, $\lim_{T\to\infty} \Pr\left\{DF_r < cv^{\beta_T}\right\} = 1$.

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

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

due to the fact that $\Pr\left\{DF_{r_{1e}+a_\eta} > cv^{\beta_T}\right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{DF_{r_{1f}-a_\gamma} > cv^{\beta_T}\right\} \to 1$ for all $0 < a_\gamma < \gamma$. Since $\eta, \gamma > 0$ is arbitrary, $\Pr\left\{\hat{r}_{1e} < r_{1e}\right\} \to 0$ 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

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0.$$

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

Since $\lim_{T\to\infty} \Pr\left\{DF_r < cv^{\beta_T}\right\} = 1$ when $r \in N_1 \cup B_2 \cup N_2$, the strategy **cannot** estimate r_{2e} and r_{2f} consistently when $\tau_{1f} - \tau_{1e} > \tau_{2f} - \tau_{2e}$. This proves Theorem 6.

Case II Suppose $\tau_{1f} - \tau_{1e} \leq \tau_{2f} - \tau_{2e}$. The asymptotic distributions of the DF statistic under the alternative hypothesis are

$$DF_r \sim_a \begin{cases} F_r(W) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_{1e}} \frac{r_w^{3/2} B(r_{1e})}{2(r_{1e}-r_1)\int_{r_1}^{r_{1e}} B(s)ds} & \text{if } r \in B_1 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \cup N_2 \\ -T^{(1-\alpha)/2}\left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in B_2 \text{ and } \tau_{1f} - \tau_{1e} > \tau - \tau_{2e} \\ T^{1-\alpha/2}\left[\frac{cr_w^3}{2\left(r_{1e}-r_1+r_{2e}-r_{1f}\right)}\right]^{1/2} & \text{if } r \in B_2 \text{ and } \tau_{1f} - \tau_{1e} \leq \tau - \tau_{2e} \end{cases}. \quad (35)$$

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

$$\lim_{T\to\infty} \Pr\left\{DF_{r_2} > cv^{\beta_T}\right\} = \Pr\left\{F_r(W) = \infty\right\} = 0.$$

If $r \in B_1$, $\lim_{T\to\infty} \Pr\left\{DF_r > cv^{\beta_T}\right\} = 1$ provided that $\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0$. If $r \in N_1$, $\lim_{T\to\infty} \Pr\left\{DF_r < cv^{\beta_T}\right\} = 1$.

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

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

due to the fact that $\Pr\left\{BDF_{r_{1e}+a_\eta} > cv^{\beta_T}\right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{DF_{r_{1f}-a_\gamma} > cv^{\beta_T}\right\} \to 1$ 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

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0.$$

The strategy therefore consistently estimates r_{1e} and r_{1f}.

If $r \in B_2$ and $\tau_{1f} - \tau_{1e} > \tau - \tau_{2e}$, $\lim_{T\to\infty} \Pr\left\{DF_r < cv^{\beta_T}\right\} = 1$ since $cv^{\beta_T} \to \infty$. If $r \in B_2$ and $\tau_{1f} - \tau_{1e} \leq \tau - \tau_{2e}$, $\lim_{T\to\infty} \Pr\left\{DF_r > cv^{\beta_T}\right\} = 1$ provided that $\frac{cv^{\beta_T}}{T^{1-\alpha/2}} \to 0$ in view of the final panel entry of (35). If $r \in N_1$, $\lim_{T\to\infty} \Pr\left\{DF_r < cv^{\beta_T}\right\} = 1$. This implies that the strategy cannot identify the second bubble when $\tau_{1f} - \tau_{1e} > \tau_2 - \tau_{2e}$. However, when $\tau_{1f} - \tau_{1e} \leq \tau_2 - \tau_{2e}$ it can identify the second bubble provided that

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1-\alpha/2}} \to 0.$$

This suggests that estimated second bubble origination date $\hat{r}_{2e}$ will be biased, taking values of $r_{2e} + r_{1f} - r_{1e}$ (in view of the condition $\tau_{1f} - \tau_{1e} \leq \tau - \tau_{2e}$ under which the final panel entry of (35) holds). The termination point r_{2f} can be consistently estimated. This proves Theorem 7.

B.2.2: The PSY algorithm

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 follows:

$$\hat{r}_{1e} = \inf_{r\in[r_0,1]} \left\{r : BSDF_r(r_0) > scv^{\beta_T}\right\} \text{ and } \hat{r}_{1f} = \inf_{r\in[\hat{r}_{1e}+L_T,1]} \left\{r : BSDF_r(r_0) < scv^{\beta_T}\right\},$$

$$\hat{r}_{2e} = \inf_{r\in\left(\hat{r}_{1f},1\right]} \left\{ r : BSDF_r\left(r_0\right) > scv^{\beta_T} \right\} \text{ and } \hat{r}_{2f} = \inf_{r\in\left[\hat{r}_{2e}+L_T,1\right]} \left\{ r : BSDF_r\left(r_0\right) < scv^{\beta_T} \right\}.$$

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

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

$$BSDF_r\left(r_0\right) \sim_a \begin{cases} F_r\left(W, r_0\right) & \text{if } r \in N_0 \\ T^{1/2}\delta_T^{\tau-\tau_{ie}} \sup_{r_1\in[0,r-r_0]} \left\{ \frac{r_w^{3/2} B(r_{ie})}{2(r_{ie}-r_1)\int_{r_1}^{r_{ie}} B(s)ds} \right\} & \text{if } r \in B_i \\ -T^{(1-\alpha)/2} \sup_{r_1\in[0,r-r_0]} \left(\frac{1}{2}cr_w\right)^{1/2} & \text{if } r \in N_1 \cup N_2 \end{cases}. \tag{36}$$

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

$$\lim_{T\to\infty} \Pr\left\{ BSDF_r\left(r_0\right) > scv^{\beta_T} \right\} = \Pr\left\{ F_r\left(W, r_0\right) = \infty \right\} = 0.$$

If $r \in B_i$ with $i = 1, 2$, $\lim_{T\to\infty} \Pr\left\{ BSDF_r\left(r_0\right) > scv^{\beta_T} \right\} = 1$ provided that $\frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{ie}}} \to 0$. If $r \in N_i$ with $i = 1, 2$, $\lim_{T\to\infty} \Pr\left\{ BSDF_r\left(r_0\right) < scv^{\beta_T} \right\} = 1$.

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

$$\frac{1}{scv^{\beta_T}} + \frac{scv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{ie}}} \to 0\ .$$

Therefore, the date-stamping strategy based on the backward sup ADF test can consistently estimate r_{1e}, r_{1f}, r_{2e} and r_{2f}. This proves Theorem 8.

B.2.3: The sequential PWY procedure

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

$$\hat{r}_{1e} = \inf_{r\in[r_0,1]} \left\{ r : DF_r > cv^{\beta_T} \right\} \text{ and } \hat{r}_{1f} = \inf_{r\in\left[\hat{r}_{1e}+L_T,1\right]} \left\{ r : DF_r < cv^{\beta_T} \right\},$$

$$\hat{r}_{2e} = \inf_{r \in \left(\hat{r}_{1f}+r_0,1\right]} \left\{ r :_{\hat{r}_{1f}} DF_r > cv^{\beta_T} \right\} \text{ and } \hat{r}_{2f} = \inf_{r \in [\hat{r}_{2e}+L_T,1]} \left\{ r :_{\hat{r}_{1f}} DF_r < cv^{\beta_T} \right\}.$$

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

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

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

and

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

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

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

If $r \in B_1$, $\lim_{T\to\infty} \Pr\left\{DF_r > cv^{\beta_T}\right\} = 1$ provided that $\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0$. If $r \in N_1$, $\lim_{T\to\infty} \Pr\left\{DF_r < cv^{\beta_T}\right\} = 1$ and $\lim_{T\to\infty} \Pr\left\{{}_{\hat{r}_{1f}}DF_r > cv^{\beta_T}\right\} = \Pr\left\{F_r(W) = \infty\right\} = 0$. If $r \in B_2$, $\lim_{T\to\infty} \Pr\left\{{}_{\hat{r}_{1f}}DF_r > cv^{\beta_T}\right\} = 1$ provided that $\frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{2e}}} \to 0$. This implies that provided that $\frac{cv^{\beta_T}}{T^{1/2}} \to 0$, $\lim_{T\to\infty} \Pr\left\{{}_{\hat{r}_{1f}}DF_r > cv^{\beta_T}\right\} = 1$ for any $r \in B_2$. If $r \in N_2$, $\lim_{T\to\infty} \Pr\left\{{}_{\hat{r}_{1f}}DF_r < cv^{\beta_T}\right\} = 1$.

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 $\Pr\left\{DF_{r_{1e}+a_\eta} > cv^{\beta_T}\right\} \to 1$ for all $0 < a_\eta < \eta$ and $\Pr\left\{DF_{r_{1f}-a_\gamma} > cv^{\beta_T}\right\} \to 1$ 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

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{1e}}} \to 0.$$

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

$$\frac{1}{cv^{\beta_T}} + \frac{cv^{\beta_T}}{T^{1/2}\delta_T^{\tau-\tau_{2e}}} \to 0.$$

Therefore, the alternative sequential implementation of the PWY procedure consistently estimates r_{2e} and r_{2f}. This proves Theorem 9.

Table 6: Detection rate and estimation of the origination and termination dates under two bubble DGP with longer second bubble durations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, \tau_{1e} = \lfloor 0.20T \rfloor, \tau_{2e} = \lfloor 0.60T \rfloor, \tau_{1f} - \tau_{1e} = \lfloor 0.10T \rfloor, T = 100$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_{2f} - \tau_{2e} = \lfloor 0.10T \rfloor$				
Detection Rate (1)	0.70	0.76	0.68	0.65
$r_{1e} = 0.20$	0.24 (0.02)	0.24 (0.02)	0.24 (0.02)	0.24 (0.02)
$r_{1f} = 0.30$	0.30 (0.00)	0.30 (0.00)	0.30 (0.01)	0.30 (0.01)
Detection Rate (2)	0.21	0.71	0.59	0.45
$r_{2e} = 0.60$	0.66 (0.02)	0.64 (0.02)	0.64 (0.02)	0.66 (0.02)
$r_{2f} = 0.70$	0.70 (0.00)	0.70 (0.01)	0.70 (0.01)	0.71 (0.00)
$\tau_{2f} - \tau_{2e} = \lfloor 0.15T \rfloor$				
Detection Rate (1)	0.70	0.76	0.68	0.65
$r_{1e} = 0.20$	0.24 (0.02)	0.24 (0.02)	0.24 (0.02)	0.24 (0.02)
$r_{1f} = 0.30$	0.30 (0.00)	0.30 (0.00)	0.30 (0.02)	0.30 (0.01)
Detection Rate (2)	0.53	0.87	0.78	0.77
$r_{2e} = 0.60$	0.69 (0.03)	0.65 (0.03)	0.66 (0.03)	0.68 (0.03)
$r_{2f} = 0.75$	0.75 (0.00)	0.75 (0.01)	0.75 (0.01)	0.75 (0.00)
$\tau_{2f} - \tau_{2e} = \lfloor 0.20T \rfloor$				
Detection Rate (1)	0.70	0.76	0.68	0.65
$r_{1e} = 0.20$	0.24 (0.02)	0.24 (0.02)	0.24 (0.02)	0.24 (0.02)
$r_{1f} = 0.30$	0.30 (0.00)	0.30 (0.00)	0.30 (0.01)	0.30 (0.01)
Detection Rate (2)	0.76	0.93	0.87	0.90
$r_{2e} = 0.60$	0.71 (0.04)	0.66 (0.04)	0.67 (0.04)	0.69 (0.04)
$r_{2f} = 0.80$	0.80 (0.00)	0.80 (0.02)	0.80 (0.01)	0.80 (0.01)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

Table 7: Detection rate and estimates of the origination and termination dates under three bubble DGP with different first bubble durations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, T = 100, \tau_{1e} = \lfloor 0.15T \rfloor, \tau_{2e} = \lfloor 0.45T \rfloor, \tau_{3e} = \lfloor 0.75T \rfloor$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_{1f} - \tau_{1e} = \lfloor 0.1T \rfloor$, $\tau_{2f} - \tau_{2e} = \lfloor 0.2T \rfloor$, $\tau_{3f} - \tau_{3e} = \lfloor 0.1T \rfloor$				
Detection Rate (1)	0.71	0.73	0.68	0.68
$r_{1e} = 0.15$	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)
$r_{1f} = 0.25$	0.25 (0.00)	0.25 (0.00)	0.25 (0.00)	0.25 (0.02)
Detection Rate (2)	0.79	0.96	0.92	0.92
$r_{2e} = 0.45$	0.57 (0.04)	0.51 (0.04)	0.52 (0.04)	0.55 (0.04)
$r_{2f} = 0.65$	0.65 (0.00)	0.65 (0.01)	0.65 (0.01)	0.65 (0.01)
Detection Rate (3)	0.00	0.73	0.81	0.01
$r_{3e} = 0.75$	-	0.79 (0.02)	0.79 (0.02)	0.80 (0.03)
$r_{3f} = 0.85$	-	0.85 (0.00)	0.85 (0.00)	0.85 (0.01)
$\tau_{1f} - \tau_{1e} = \lfloor 0.2T \rfloor$, $\tau_{2f} - \tau_{2e} = \lfloor 0.2T \rfloor$, $\tau_{3f} - \tau_{3e} = \lfloor 0.1T \rfloor$				
Detection Rate (1)	0.92	0.94	0.88	0.93
$r_{1e} = 0.15$	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)
$r_{1f} = 0.35$	0.35 (0.01)	0.35 (0.01)	0.35 (0.01)	0.35 (0.02)
Detection Rate (2)	0.13	1.00	0.95	0.27
$r_{2e} = 0.45$	0.60 (0.03)	0.51 (0.04)	0.50 (0.04)	0.60 (0.03)
$r_{2f} = 0.65$	0.65 (0.00)	0.65 (0.01)	0.65 (0.01)	0.65 (0.00)
Detection Rate (3)	0.00	0.75	0.83	0.00
$r_{3e} = 0.75$	-	0.79 (0.02)	0.78 (0.02)	0.81 (0.02)
$r_{3f} = 0.85$	-	0.85 (0.00)	0.85 (0.00)	0.85 (0.00)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

Table 8: Detection rate and estimates of the origination and termination dates under three bubble DGP with different second bubble durations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, T = 100, \tau_{1e} = \lfloor 0.15T \rfloor, \tau_{2e} = \lfloor 0.45T \rfloor, \tau_{3e} = \lfloor 0.75T \rfloor$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_{1f} - \tau_{1e} = \lfloor 0.1T \rfloor$, $\tau_{2f} - \tau_{2e} = \lfloor 0.1T \rfloor$, $\tau_{3f} - \tau_{3e} = \lfloor 0.2T \rfloor$				
Detection Rate (1)	0.71	0.73	0.68	0.68
$r_{1e} = 0.15$	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)
$r_{1f} = 0.25$	0.25 (0.00)	0.25 (0.00)	0.25 (0.00)	0.25 (0.01)
Detection Rate (2)	0.17	0.74	0.64	0.37
$r_{2e} = 0.45$	0.51 (0.02)	0.49 (0.02)	0.49 (0.02)	0.51 (0.02)
$r_{2f} = 0.55$	0.55 (0.00)	0.55 (0.00)	0.55 (0.01)	0.55 (0.01)
Detection Rate (3)	0.68	0.94	0.86	0.87
$r_{3e} = 0.75$	0.88 (0.03)	0.81 (0.04)	0.81 (0.05)	0.86 (0.04)
$r_{3f} = 0.95$	0.95 (0.00)	0.95 (0.01)	0.95 (0.01)	0.95 (0.01)
$\tau_{1f} - \tau_{1e} = \lfloor 0.1T \rfloor$, $\tau_{2f} - \tau_{2e} = \lfloor 0.2T \rfloor$, $\tau_{3f} - \tau_{3e} = \lfloor 0.2T \rfloor$				
Detection Rate (1)	0.71	0.73	0.68	0.68
$r_{1e} = 0.15$	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)
$r_{1f} = 0.25$	0.25 (0.00)	0.25 (0.00)	0.25 (0.00)	0.25 (0.02)
Detection Rate (2)	0.79	0.96	0.90	0.92
$r_{2e} = 0.45$	0.57 (0.04)	0.51 (0.04)	0.52 (0.04)	0.54 (0.04)
$r_{2f} = 0.65$	0.65 (0.00)	0.65 (0.01)	0.65 (0.01)	0.65 (0.01)
Detection Rate (3)	0.13	0.96	0.92	0.22
$r_{3e} = 0.75$	0.91 (0.02)	0.81 (0.04)	0.80 (0.04)	0.90 (0.03)
$r_{3f} = 0.95$	0.95 (0.00)	0.95 (0.02)	0.95 (0.01)	0.95 (0.01)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

Table 9: Detection rate and estimation of the origination and termination dates under three bubble DGP with different third bubble durations. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, T = 100, \tau_{1e} = \lfloor 0.15T \rfloor, \tau_{2e} = \lfloor 0.45T \rfloor, \tau_{3e} = \lfloor 0.75T \rfloor$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_{1f} - \tau_{1e} = \lfloor 0.2T \rfloor$, $\tau_{2f} - \tau_{2e} = \lfloor 0.1T \rfloor$, $\tau_{3f} - \tau_{3e} = \lfloor 0.1T \rfloor$				
Detection Rate (1)	0.92	0.94	0.88	0.93
$r_{1e} = 0.15$	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)
$r_{1f} = 0.35$	0.35 (0.00)	0.35 (0.01)	0.35 (0.01)	0.35 (0.01)
Detection Rate (2)	0.00	0.75	0.84	0.01
$r_{2e} = 0.45$	-	0.49 (0.02)	0.49 (0.02)	0.51 (0.02)
$r_{2f} = 0.55$	-	0.55 (0.00)	0.55 (0.01)	0.55 (0.01)
Detection Rate (3)	0.01	0.76	0.68	0.05
$r_{3e} = 0.75$	0.82 (0.02)	0.79 (0.02)	0.79 (0.02)	0.81 (0.02)
$r_{3f} = 0.85$	0.85 (0.00)	0.85 (0.00)	0.85 (0.00)	0.85 (0.00)
$\tau_{1f} - \tau_{1e} = \lfloor 0.2T \rfloor$, $\tau_{2f} - \tau_{2e} = \lfloor 0.1T \rfloor$, $\tau_{3f} - \tau_{3e} = \lfloor 0.2T \rfloor$				
Detection Rate (1)	0.92	0.94	0.88	0.93
$r_{1e} = 0.15$	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)
$r_{1f} = 0.35$	0.35 (0.00)	0.35 (0.01)	0.35 (0.01)	0.35 (0.01)
Detection Rate (2)	0.00	0.62	0.06	0.01
$r_{2e} = 0.40$	-	0.47 (0.00)	0.47 (0.01)	0.45 (0.03)
$r_{2f} = 0.50$	-	0.50 (0.00)	0.50 (0.01)	0.50 (0.01)
Detection Rate (3)	0.01	0.76	0.17	0.06
$r_{3e} = 0.75$	0.82 (0.02)	0.79 (0.02)	0.81 (0.02)	0.81 (0.02)
$r_{3f} = 0.85$	0.85 (0.00)	0.85 (0.00)	0.55 (0.00)	0.85 (0.00)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.

Table 10: Detection rate and estimates of the origination and termination dates under three bubble DGP and special examples. Parameters are set to: $y_0 = 100, c = 1, \sigma = 6.79, \alpha = 0.6, T = 100, \tau_{1e} = \lfloor 0.15T \rfloor$. Figures in parentheses are standard deviations.

	PWY	PSY	Seq	CUSUM
$\tau_{1f} - \tau_{1e} = \lfloor 0.1T \rfloor, \tau_{2f} - \tau_{2e} = \lfloor 0.2T \rfloor, \tau_{3f} - \tau_{3e} = \lfloor 0.10T \rfloor, \tau_{2e} = \lfloor 0.45T \rfloor, \tau_{3e} = \lfloor 0.70T \rfloor$				
Detection Rate (1)	0.72	0.74	0.68	0.69
$r_{1e} = 0.15$	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)	0.19 (0.02)
$r_{1f} = 0.25$	0.25 (0.00)	0.25 (0.00)	0.25 (0.00)	0.25 (0.02)
Detection Rate (2)	0.79	0.95	0.90	0.91
$r_{2e} = 0.45$	0.57 (0.04)	0.51 (0.04)	0.51 (0.04)	0.55 (0.04)
$r_{2f} = 0.65$	0.65 (0.00)	0.65 (0.01)	0.65 (0.01)	0.65 (0.01)
Detection Rate (3)	0.00	0.62	0.00	0.01
$r_{3e} = 0.70$	-	0.77 (0.00)	-	0.75 (0.03)
$r_{3f} = 0.80$	-	0.80 (0.00)	-	0.80 (0.01)
$\tau_{1f} - \tau_{1e} = \lfloor 0.2T \rfloor, \tau_{2f} - \tau_{2e} = \lfloor 0.1T \rfloor, \tau_{3f} - \tau_{3e} = \lfloor 0.10T \rfloor, \tau_{2e} = \lfloor 0.40T \rfloor, \tau_{3e} = \lfloor 0.75T \rfloor$				
Detection Rate (1)	0.72	0.94	0.88	0.93
$r_{1e} = 0.15$	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)	0.21 (0.04)
$r_{1f} = 0.35$	0.35 (0.00)	0.35 (0.01)	0.35 (0.01)	0.35 (0.01)
Detection Rate (2)	0.00	0.62	0.06	0.01
$r_{2e} = 0.40$	-	0.47 (0.00)	0.47 (0.01)	0.45 (0.03)
$r_{2f} = 0.50$	-	0.50 (0.00)	0.50 (0.01)	0.50 (0.01)
Detection Rate (3)	0.01	0.76	0.17	0.06
$r_{3e} = 0.75$	0.82 (0.02)	0.79 (0.02)	0.81 (0.02)	0.81 (0.02)
$r_{3f} = 0.85$	0.85 (0.00)	0.85 (0.00)	0.85 (0.00)	0.85 (0.00)

Note: Calculations are based on 5,000 replications. The minimum window has 12 observations.