Sign-based Unit Root Tests for Explosive Financial Bubbles in the Presence of Deterministically Time-Varying Volatility*

David I. Harvey, Stephen J. Leybourne and Yang Zu
School of Economics, University of Nottingham

January 2019

Abstract

This paper considers the problem of testing for an explosive bubble in financial data in the presence of time-varying volatility. We propose a sign-based variant of the Phillips, Shi and Yu (2015) test. Unlike the original test, the sign-based test does not require bootstrap-type methods to control size in the presence of time-varying volatility. Under a locally explosive alternative, the sign-based test delivers higher power than the original test for many time-varying volatility and bubble specifications. However, since the original test can still outperform the sign-based one for some specifications, we also propose a union of rejections procedure that combines the original and sign-based tests, employing a wild bootstrap to control size. This is shown to capture most of the power available from the better performing of the two tests. We also show how a sign-based statistic can be used to date the bubble start and end points. An empirical illustration using Bitcoin price data is provided.

Keywords: Rational bubble; Explosive autoregression; Time-varying volatility; Right-tailed unit root testing; Sign-based test.

JEL Classification: C12, C32.

*We are grateful to the Editor, Peter Phillips, the Co-Editor, Anna Mikusheva, and three anonymous referees for their very helpful and constructive comments. Correspondence to: David Harvey, School of Economics, University of Nottingham, University Park, Nottingham, NG7 2RD, UK. Email: dave.harvey@nottingham.ac.uk

1 Introduction

Empirical identification of explosive behaviour in financial asset price series is closely related to the study of rational bubbles, with a rational bubble deemed to have occurred if explosive characteristics are manifest in the time path of prices, but not for the dividends. Consequently, methods for testing for explosive time series behaviour have been a focus of much recent research. Phillips et al. (2015) [PSY] model potential bubble behaviour using a time-varying autoregressive specification, which allows an explosive autoregressive regime over a subset of the data, and suggest testing for such a property using a double supremum of forward and backward recursive right-tailed Dickey-Fuller (DF) unit root tests, a generalisation of the original and widely-used Phillips et al. (2011) [PWY] test that employed a single supremum of forward-only recursive DF tests.

These papers assume constant unconditional volatility in the underlying error process, yet, in practice, time-varying volatility is a well-known stylised fact observed in empirical financial data (see, for example, Rapach *et al.* (2008)). Harvey et al. (2016) [HLST] demonstrate that the asymptotic null distribution of the PWY test depends on the nature of the volatility, so if the test is compared to critical values derived under a homoskedastic error assumption, its size is not controlled under time-varying volatility. This lack of size control typically leads to serious over-sizing, and consequently frequent spurious identification of a bubble. HLST propose a wild bootstrap method to provide critical values for the PWY test, which delivers correct asymptotic size in the presence of time-varying volatility (while retaining the same local asymptotic power as the original PWY test, were it infeasibly size-corrected to account for the time-varying volatility). An entirely similar bootstrap approach can be applied to the PSY test.[1]

In this paper we suggest a new approach to obtaining heteroskedasticity-robust inference in the presence of a bubble. Instead of calculating the PSY statistic (denoted PSY) directly from the observed data series, y_t say, we calculate it from the series of cumulated *signs* of the first differences of the data, i.e. a cumulation of $sign(\Delta y_t) = \Delta y_t / |\Delta y_t|$ (for nonzero Δy_t),which is clearly invariant to the variance of Δy_t (assuming a zero mean for Δy_t). As a direct consequence of this, the sign-based PSY statistic (denoted $sPSY$) is then exact invariant to the pattern of time-varying volatility and therefore, unlike PSY, requires no wild bootstrap procedure to control size. Sign-based approaches to testing for unit roots against stationary autoregressive models have been considered by, *inter alios*, Campbell and Dufour (1995) and So and Shin (2001), although these are not based on the cumulations of $sign(\Delta y_t)$. Unit root testing using cumulated standardised differences is also considered by Beare (2018) (in a context of full sample testing against a stationary alternative), but our method is quite distinct in that we standardise by $|\Delta y_t|$, rather than using a nonparametric estimator of the spot standard deviation, resulting in the sign of Δy_t.

We derive the asymptotic distribution of the sign-based test under the unit root null and alternative of a local to unit root explosive regime. Here, we derive a stochastic

[1]Relatedly, a wild bootstrap methodology is employed in the context of recursive testing in Shi et al. (2018b) and Phillips and Shi (2018b) to address the additional size control issues that arise under multiple testing.

expansion of the partial sum process (PSP) of signed first differences, allowing for time-varying volatility and a time-varying autoregressive coefficient, thereby extending the results of Boldin (2013) for a homoskedastic constant coefficient model. We then use this result to establish the asymptotic properties of our test statistics.

Using a number of different specifications for the bubble process and pattern of time-varying volatility, we show that the local asymptotic power of $sPSY$ compares very well with that of PSY and it is, rather more often than not, the more powerful of the two test procedures, sometimes by a significant margin. Although $sPSY$ has a good deal of merit as a stand alone test, because for some bubble process and volatility pattern settings the power of PSY is higher than that of $sPSY$, we then proceed to consider a union of rejections approach (cf. Harvey et al. (2009)), whereby the null is rejected in favour of explosive behaviour if either PSY or $sPSY$ rejects. We find that the union of rejections testing strategy performs very well across the full range of volatility and bubble specifications that we consider, capturing much of the power available from either test. In common with PSY, a feasible variant of the union test does require a (joint) wild bootstrap to ensure asymptotic size control. In the paper we refer mainly to the test of PSY and its sign-based counterpart, but we simultaneously consider variants appropriate for the original test of PWY.

We then move on to consider how a modified variant of our sign-based statistic can be used to date the start and end of a bubble. We propose a new dating strategy based on maximising a dating statistic. Under a mildly explosive assumption for the bubble magnitude, we show that our proposed dating strategy is consistent for estimating the start and end of the bubble. This, of course, is a relevant property to establish from the viewpoint of an applied researcher who is interested in characterising the timeline of a historical bubble episode in relation to, say, economic or financial events that are known to have occurred.

The rest of the paper is organised as follows. Section 2 outlines the heteroskedastic bubble model, describes the PSY testing approach and introduces our sign-based version of the PSY test. Here we also establish the limit distributions of these tests under local bubble alternatives. Asymptotic size (where relevant) and local powers are compared in section 3. The union of rejections procedure and the associated wild bootstrap method are outlined in section 4. Finite sample properties of the tests are explored in section 5. Our sign-based dating methodology is described and its consistency properties shown in section 6. A generalisation of our sign-based test to account for possible asymmetry of the innovation distribution is given in section 7. Section 8 briefly discusses extensions to the basic model. An empirical illustration of our new testing and dating procedures, using Bitcoin price data, is provided in section 9, with section 10 concluding the paper. Proofs of our asymptotic results are provided in an appendix.

We use the following notation: $\mathbf{1}(.)$ denotes the indicator function; $\lfloor \cdot \rfloor$ the integer part; $\Rightarrow$ weak convergence; $\overset{p}{\Rightarrow}$ weak convergence in probability, and $\xrightarrow{p}$ convergence in probability. $\mathcal{D} = D[0,1]$ denotes the space of right continuous with left limit (càdlàg) processes on $[0,1]$. Finally,'$x := y$' ('$x =: y$') indicates that x (y) is defined by y (x). In this paper, we study two types of models for the explosive behaviour in data, and we use the following terminology: (i) *locally explosive* refers to the alternative where the autoregressive root is $1 + cT^{-1}$, with c a positive constant and T the sample size; (ii) *mildly explosive* refers to the alternative specified in Phillips and Magdalinos (2007), where the root is $1 + cT^{-\alpha}$ with $\alpha \in (0, 1)$. Formally, our sign function is defined as $\text{sign}(x) = -2\mathbf{1}(x \leqslant 0) + 1$.

2 The heteroskedastic stochastic bubble model: PSY and sign-based PSY tests

We will consider the time series process $\{y_t\}$ generated according to the following DGP (cf. HLST, Phillips and Shi, 2018a)

$$\begin{aligned} y_t &= \mu + u_t \qquad (1) \\ u_t &= \begin{cases} u_{t-1} + \varepsilon_t, & t = 2, ..., \lfloor \tau_1 T \rfloor, \\ (1 + \delta_{1,T}) u_{t-1} + \varepsilon_t, & t = \lfloor \tau_1 T \rfloor + 1, ..., \lfloor \tau_2 T \rfloor, \\ (1 - \delta_{2,T}) u_{t-1} + \varepsilon_t, & t = \lfloor \tau_2 T \rfloor + 1, ..., \lfloor \tau_3 T \rfloor, \\ u_{t-1} + \varepsilon_t, & t = \lfloor \tau_3 T \rfloor + 1, ..., T \end{cases} \end{aligned}$$

where $\delta_{1,T} \geq 0$ and $\delta_{2,T} \geq 0$. We assume that the initial condition u_1 is such that $u_1 = o_p(T^{1/2})$. Here ε_t is a zero-mean, (possibly) heteroskedastic innovation process whose precise assumptions are detailed later.

The DGP imposes a unit root on y_t up to time $\lfloor \tau_1 T \rfloor$, after which y_t is explosive (when $\delta_{1,T} > 0$) until time $\lfloor \tau_2 T \rfloor$.[2] If $\tau_2 < 1$, the explosive regime then terminates at some in-sample date, at which point the model permits a possible collapse, where $\delta_{2,T} > 0$ and stationary mean-reverting behaviour acts to proxy the collapse regime. The null hypothesis, H_0, is that no bubble is present in the series and y_t follows a unit root process throughout the sample period i.e. $H_0 : \delta_{i.T} = 0$, $i = 1, 2$ (equivalently, $H_0 : \tau_1 = 1$). The alternative hypothesis, $H_1 : \delta_{1,T} > 0$ and $\delta_{2,T} > 0$, comprises any one of the following four scenarios for the behaviour of y_t:

DGP 1 $\quad 0 < \tau_1 < 1$, $\tau_2 = \tau_3 = 1$
(unit root, then bubble to sample end)

DGP 2 $\quad 0 < \tau_1 < \tau_2 < 1$, $\tau_2 = \tau_3$
(unit root, then bubble, then unit root to sample end)

DGP 3 $\quad 0 < \tau_1 < \tau_2 < 1$, $\tau_3 = 1$
(unit root, then bubble, then collapse to sample end)

DGP 4 $\quad 0 < \tau_1 < \tau_2 < \tau_3 < 1$
(unit root, then bubble, then collapse, then unit root to sample end)

For the majority of our analysis, under H_1 we will consider locally explosive alternatives (and collapses) of the form $\delta_{i.T} = c_i T^{-1}$, $c_i > 0$, $i = 1, 2$; the scaling by T^{-1} providing the appropriate Pitman drift for asymptotic power comparisons of the tests.

[2] Note that while we assume the presence of a unit root regime at the beginning of the sample prior to any explosive behaviour in keeping with much of the recent literature on bubble testing, this is not critical for our analysis and the explosive regime could originate at the sample start date.

In section 6 below, when we consider dating the start and end points of the bubble, a slightly stronger, mildly explosive, bubble magnitude will be assumed for $\delta_{1,T}$. For the innovation process ε_t we make the following assumptions:

A1 *$\varepsilon_t = \sigma_t z_t$ where $z_t \sim IID$ with $E(z_t) = 0$, $E(z_t^2) = 1$ and $E(|z_t|^r) < \infty$ for some $r \geq 4$.*

A2 *The volatility term σ_t satisfies $\sigma_t = \sigma(t/T)$, where $\sigma(\cdot) \in \mathcal{D}$ is non-stochastic and strictly positive.*

A3 *The CDF of z_t, denoted $F(z)$, is such that $f(z) = F'(z)$ is continuous at 0 and satisfies $f(0) > 0$ with $\sup_z f(z) < \infty$.*

A4 *$F(0) = 1/2$.*

Under Assumption A2, the innovation variance is non-stochastic, bounded and displays a countable number of jumps. It also allows for variance processes displaying (possibly) multiple one-time volatility shifts (which need not be located at the same point in the sample as the putative regimes associated with bubble behaviour), polynomially (possibly piecewise) trending volatility and smooth transition variance breaks, among other things. The conventional homoskedasticity assumption, that $\sigma_t = \sigma$ for all t, is also permitted, since here $\sigma(s) = \sigma$ for all s. Assumption A3 ensures $F(z)$ is continuously differentiable in a small neighbourhood around zero, and that the density $f(z)$ exists, is strictly positive, and is bounded from above. Assumption A4 implies that $E(sign(z_t)) = 0$, which is necessary for the invariance principle of the partial sum of the signs to hold. Assumption A4 also implies the median of z_t is zero, in addition to the zero mean assumption from A1; the imposed distributional assumption on z_t is only slightly weaker than assuming the distribution of z_t is symmetric about zero. Note that monthly financial returns, which are often used in a bubble testing context, are usually found to be symmetric about zero; see, for example, Tsay (2010, Table 1.2), Christoffersen (2012, Section 2). In section 7 below we will consider relaxing this symmetry assumption.

Under Assumptions A1-A2, the following invariance principle holds for the PSP of ε_t:

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \varepsilon_t \Rightarrow \int_0^r \sigma(h) dW(h) =: W_\sigma(r).$$

Under Assumptions A1-A2 and A4, we have the following invariance principle for the PSP of $sign(\varepsilon_t) := -2\mathbf{1}(\varepsilon_t \leq 0) + 1$:

$$T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} sign(\varepsilon_t) = T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} sign(z_t) \Rightarrow W^s(r). \tag{2}$$

Here $W(r)$ and $W^s(r)$ are standard Brownian motion processes, with the correlation coefficient being the constant $-2E\{\mathbf{1}(z_t \leq 0)z_t\}$. Notice that $W_\sigma(r)$ is a stochastic integral dependent on the volatility function $\sigma(s)$. Also note that $sign(\varepsilon_t)$ is exact invariant to σ_t.

2.1 The PSY test

The PSY statistic is used to test H_0 against H_1, the alternative being that y_t behaves as an explosive AR(1) process for at least some sub-period of the sample. In this context, and in the absence of knowledge concerning the timing of any potential explosive behaviour, PSY propose a test based on the double-supremum of recursive right-tailed DF tests. Specifically, the statistic is given by

$$PSY := \sup_{\lambda_1 \in [0,1-\pi]} \sup_{\lambda_2 \in [\lambda_1+\pi,1]} DF(\lambda_1, \lambda_2)$$

where $DF(\lambda_1, \lambda_2)$ denotes the standard DF test, that is the t-ratio for $\hat{\phi}(\lambda_1, \lambda_2)$ in the fitted ordinary least squares (OLS) regression

$$\Delta y_t = \hat{\alpha}(\lambda_1, \lambda_2) + \hat{\phi}(\lambda_1, \lambda_2) y_{t-1} + \hat{\varepsilon}_t \tag{3}$$

calculated over the sub-sample period $t = \lfloor \lambda_1 T \rfloor, ..., \lfloor \lambda_2 T \rfloor$. That is

$$DF(\lambda_1, \lambda_2) := \frac{\hat{\phi}(\lambda_1, \lambda_2)}{\sqrt{\hat{\sigma}^2(\lambda_1, \lambda_2) / \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} (y_{t-1} - \bar{y})^2}}$$

where $\bar{y} := (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} y_{t-1}$ and $\hat{\sigma}^2(\lambda_1, \lambda_2) := (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 2)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} \hat{\varepsilon}_t^2$. The PSY statistic is therefore the supremum of a double sequence of statistics with minimum sample length $\lfloor \pi T \rfloor$. We assume that $\tau_1 \geq \pi$, such that the onset of a bubble regime (should one occur), begins after the shortest sub-sample considered. The single-supremum statistic of PWY arises as a special case of the PSY statistic: $PWY := \sup_{\lambda_2 \in [\pi,1]} DF(0, \lambda_2)$.

We now state the large sample behaviour of PSY under a locally explosive H_1 for DGP 4. Its behaviour under DGPs 1-3, and under H_0, arise as special cases.

Theorem 1 *For model (1), under H_1 with $\delta_{i.T} = c_i T^{-1}$, $c_i > 0$, $i = 1, 2$ and Assumptions A1-A2,*

$$PSY \Rightarrow \sup_{\lambda_1 \in [0,1-\pi]} \sup_{\lambda_2 \in [\lambda_1+\pi,1]} L_{c_1,c_2}(\lambda_1, \lambda_2) =: MM_{c_1,c_2}$$

where

$$L_{c_1,c_2}(\lambda_1, \lambda_2) := \frac{\tilde{U}(\lambda_2)^2 - \tilde{U}(\lambda_1)^2 - \int_{\lambda_1}^{\lambda_2} \sigma(r)^2 dr}{2\sqrt{(\lambda_2 - \lambda_1)^{-1} \int_{\lambda_1}^{\lambda_2} \sigma(r)^2 dr \int_{\lambda_1}^{\lambda_2} \tilde{U}(r)^2 dr}}$$

and

$$\tilde{U}(r) := U(r) - (\lambda_2 - \lambda_1)^{-1} \int_{\lambda_1}^{\lambda_2} U(h) dh$$

with

$$U(r) := \begin{cases} W_\sigma(r) & r \leq \tau_1 \\ V_1(r) & \tau_1 < r \leq \tau_2 \\ V_2(r) & \tau_2 < r \leq \tau_3 \\ V_2(\tau_3) + W_\sigma(r) - W_\sigma(\tau_3) & r > \tau_3 \end{cases}$$

where

$$\begin{aligned} V_1(r) &:= e^{(r-\tau_1)c_1} W_\sigma(\tau_1) + \int_{\tau_1}^{r} e^{(r-h)c_1} dW_\sigma(h) \\ V_2(r) &:= e^{-(r-\tau_2)c_2} V_1(\tau_2) + \int_{\tau_2}^{r} e^{-(r-h)c_2} dW_\sigma(h). \end{aligned}$$

Corresponding limiting distributions under DGP 1, DGP 2 or DGP 3 are obtained by imposing the relevant restrictions on τ_2 and τ_3. The limit distribution of PSY under the null hypothesis H_0 is given by $MM_{0,0}$ (or, equivalently, on setting $\tau_1 = 1$ such that $U(r) = W_\sigma(r)$). The limit of the PWY test is $\sup_{\lambda_2 \in [\pi,1]} L_{c_1,c_2}(0,\lambda_2) =: M_{c_1,c_2}$, with distribution $M_{0,0}$ under H_0. The limits of both PSY and PWY are dependent on the (limit) form of heteroskedasticity $\sigma(s)$ under the null and alternative hypotheses.

2.2 The sign-based PSY test

Let C_t be the cumulated sum of signs $C_t := \sum_{i=2}^{t} sign(\Delta y_t)$, $t = 2, ..., T$. The sign-based analogue of (3) is then given by

$$sPSY := \sup_{\lambda_1 \in [0,1-\pi]} \sup_{\lambda_2 \in [\lambda_1+\pi,1]} sDF(\lambda_1, \lambda_2) \tag{4}$$

where $sDF(\lambda_1, \lambda_2)$ denotes the t-ratio for $\hat{\rho}(\lambda_1, \lambda_2)$ in the fitted (without intercept) OLS regression

$$\Delta C_t = \hat{\rho}(\lambda_1, \lambda_2) C_{t-1} + e_t$$

calculated over the period $t = \lfloor \lambda_1 T \rfloor, ..., \lfloor \lambda_2 T \rfloor$, i.e.

$$sDF(\lambda_1, \lambda_2) := \frac{\hat{\rho}(\lambda_1, \lambda_2)}{\sqrt{\hat{s}^2(\lambda_1, \lambda_2) / \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} C_{t-1}^2}}$$

where $\hat{s}^2(\lambda_1, \lambda_2) := (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 1)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} e_t^2$. The sign-based analogue of the PWY test arises as a special case of the $sPSY$ test: $sPWY := \sup_{\lambda_2 \in [\pi,1]} sDF(0, \lambda_2)$. Under the null hypothesis, since $sign(\Delta y_t) = sign(z_t)$, these tests are exact invariant to the pattern of heteroskedasticity σ_t.

For DGP 4, the next Theorem gives the large sample behaviour of $sPSY$ under a locally explosive H_1.

Theorem 2 *For model (1), under H_1 with $\delta_{i.T} = c_i T^{-1}$, $c_i > 0$, $i = 1, 2$ and Assumptions A1-A4,*

$$sPSY \Rightarrow \sup_{\lambda_1 \in [0,1-\pi]} \sup_{\lambda_2 \in [\lambda_1+\pi,1]} L^s_{c_1,c_2}(\lambda_1, \lambda_2) =: MM^s_{c_1,c_2} \tag{5}$$

where

$$L^s_{c_1,c_2}(\lambda_1, \lambda_2) := \frac{U^s(\lambda_2)^2 - U^s(\lambda_1)^2 - (\lambda_2 - \lambda_1)}{2\sqrt{\int_{\lambda_1}^{\lambda_2} U^s(r)^2 dr}}$$

with

$$U^s(r) := \begin{cases} W^s(r) & r \leq \tau_1 \\ W^s(r) + 2f(0)X_1(r) & \tau_1 < r \leq \tau_2 \\ W^s(r) + 2f(0)X_2(r) & \tau_2 < r \leq \tau_3 \\ W^s(r) + 2f(0)X_2(\tau_3) & r > \tau_3 \end{cases} \tag{6}$$

and

$$\begin{aligned} X_1(r) &:= c_1 \int_{\tau_1}^{r} \{V_1(h)/\sigma(h)\}\, dh \\ X_2(r) &:= X_1(\tau_2) - c_2 \int_{\tau_2}^{r} \{V_2(h)/\sigma(h)\}\, dh \end{aligned}$$

where $V_1(r)$ and $V_2(r)$ are as defined in Theorem 1.

Once more the corresponding limiting distributions under DGP 1, DGP 2 or DGP 3 obtain by imposing the relevant restrictions on τ_2 and τ_3, with the limit distributions of $sPSY$ under the null hypothesis H_0 being given by $MM^s_{0,0}$ (or on setting $\tau_1 = 1$ so that $U^s(r) = W^s(r)$). The limit of the PWY sign test, $sPWY$, is given by $\sup_{\lambda_2 \in [\pi,1]} L^s_{c_1,c_2}(0, \lambda_2) =: M^s_{c_1,c_2}$, with distribution $M^s_{0,0}$ under H_0. Note that $MM^s_{0,0}$ and $M^s_{0,0}$ are invariant to $\sigma(s)$, while under the alternative hypothesis, the limits of $sPSY$ and $sPWY$ depend on the pattern of heteroskedasticity $\sigma(s)$, and also the density of z_t via the appearance of $f(0)$.

For $\pi = 0.1$, limit null critical values for $sPSY$ and $sPWY$, for the standard significance levels, are given in Table 1 under "$T = \infty$". These are computed using direct simulation of the limiting functionals of Theorem 2, using 2000 Monte Carlo replications, and approximating the Brownian motion process involved using $NIID(0, 1)$ random variates, with the integrals approximated by normalized sums of 1000 steps. Also shown in Table 1 are finite sample critical values for $sPSY$ and $sPWY$ based on generating ε_t as $NIID(0, 1)$ (with $u_1 = \varepsilon_1$) for $T = 100$, 200 and 400. It is clear that convergence of the finite sample critical values to their asymptotic counterparts is fairly slow (particularly for $sPSY$), but this is not uncommon for extremum statistics based on sub-samples.

Remark 1 By construction, both the original PSY and PWY statistics, and the sign-based variants $sPSY$ and $sPWY$, are numerically invariant to the nuisance parameter μ in the DGP (1). For PSY and PWY this follows due to the inclusion of an intercept term in the Dickey-Fuller regressions (3), while for $sPSY$ and $sPWY$ the statistics only make use of C_t, which, being based on the (sign of) Δy_t, does not depend on μ. Hence the finite sample and limit distributions of these statistics, and consequently their finite sample and asymptotic sizes and local powers, do not depend on μ. One could also envisage tests of the form of PSY and PWY but based on Dickey-Fuller regressions that exclude an intercept term. Such tests would have finite sample distributions that depend on the nuisance parameter μ under both the null and alternative, while it can easily be shown that their asymptotic null and local alternative distributions would be invariant to μ provided $\mu = o(T^{1/2})$.

3 Asymptotic size and power of the tests

We now consider the asymptotic size and power of the PSY and PWY tests, and asymptotic powers of $sPSY$ and $sPWY$ tests. The sizes and powers are computed via direct simulation of the limiting functionals in Theorems 1 and 2, again using 2000 Monte Carlo replications.

3.1 Size

Sizes for PSY and PWY are examined in the case of volatility shifts of the form

$$\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_{\sigma 1}) + \sigma_1 \mathbf{1}(\tau_{\sigma 1} < s \leq \tau_{\sigma 2}) + \mathbf{1}(\tau_{\sigma 2} < s \leq 1).$$

We simulate the asymptotic sizes of upper-tail nominal 0.05-level tests, and use the limit null critical value which would be obtained under homoskedasticity, i.e. from the distributions $MM_{0,0}$ and $M_{0,0}$ evaluated assuming $\sigma(s) = 1$, which is akin to ignoring any possibility of heteroskedasticity. We consider the range of values $\sigma_1 \in \{1, 1/6, 1/3, 3, 6\}$. The results are given in Table 2. We do not show size results for $sPSY$ and $sPWY$ as they are always correctly sized asymptotically.

Panel (a) of Table 2 sets $\tau_{\sigma 1} \in \{0.4, 0.8\}$ and $\tau_{\sigma 2} = 1$. This represents a single volatility shift at time fraction $\tau_{\sigma 1}$, which might be thought of as being akin to DGP 1 with the bubble episode being replaced by a heteroskedastic one. It is evident that PSY and PWY are both badly oversized when $\sigma_1 > 1$, this oversize being particularly serious for $\sigma_1 = 6$. Comparing PSY and PWY, we see that the length of the heteroskedastic episode, as measured by $\tau_{\sigma 1}$, actually has little effect on the degree of oversize present in PSY, while for PWY we see a modest decrease in size with increasing $\tau_{\sigma 1}$.

In Panel (b) we set $\tau_{\sigma 1} \in \{0.1, 0.5\}$ and $\tau_{\sigma 2} = 0.7$. Here there is a change in volatility between time fractions $\tau_{\sigma 1}$ and $\tau_{\sigma 2}$, which is now akin to DGP 2 with the bubble episode being replaced by a heteroskedastic one. Here we see that PSY is badly oversized for *both* $\sigma_1 < 1$ and $\sigma_1 > 1$, and for all $\tau_{\sigma 1}$. While PWY is similarly oversized for $\sigma_1 > 1$, for $\sigma_1 < 1$ (modest) oversize is only evident when $\tau_{\sigma 1} = 0.1$. This represents something of a departure in behaviour between the two tests, indicating that the size of PSY is more sensitive to the presence of heteroskedasticity.

3.2 Power

We now examine the asymptotic power of the tests under a locally explosive H_1, for both a benchmark case of homoskedasticity, and also in the presence of heteroskedasticity. We do this in the context of DGP 1, DGP 2 and a representative DGP involving a collapse regime (specifically, DGP 4), noting that we find the specification of the collapse regime to have relatively little bearing on the powers of the tests. We simulate the asymptotic powers of upper-tail nominal 0.05-level tests. For PSY and PWY, we infeasibly size-correct when a particular pattern of heteroskedasticity is present by taking critical values from the $\sigma(s)$-dependent $MM_{0,0}$ and $M_{0,0}$ limit distributions. For $sPSY$ and $sPWY$, the critical values are the limit ones from Table 1. To evaluate the powers of these tests we (implicitly) assume that $z_t \sim NIID(0,1)$ and correspondingly set $f(0) = 1/\sqrt{2\pi} = 0.399$.

The model parameter settings we consider are as follows:

DGP 1 $\tau_1 \in \{0.4, 0.8\}$
$\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq 1)$
(unit root, then bubble to sample end)

DGP 2 $\tau_1 \in \{0.1, 0.5\}$, $\tau_2 = 0.7$
$\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq \tau_2) + \mathbf{1}(\tau_2 < s \leq 1)$
(unit root, then bubble, then unit root to sample end)

DGP 4 $\tau_1 \in \{0.1, 0.5\}$, $\tau_2 = 0.7$, $\tau_3 = 0.8$, $c_2 = c_1$
$\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq \tau_3) + \mathbf{1}(\tau_3 < s \leq 1)$
(unit root, then bubble, then collapse, then unit root to sample end)

We set $c_1 \in \{2, 4, 6, 8\}$ and $\sigma_1 \in \{1, 1/6, 1/3, 3, 6\}$; $\sigma_1 = 1$ representing the benchmark homoskedastic case. In each DGP, the heteroskedastic episode is made coincident with the bubble (or bubble and collapse) regime(s), which seems a reasonable restriction to impose and it limits the number of cases to consider. The results are given in Table 3. As there are still a large number of table entries, as a simple gauge of the broad relative power performance of PSY compared with $sPSY$, and PWY with $sPWY$, entries where the power of one test exceeds that of the other by at least 0.04 are underlined.

The results for DGP 1 appear in Table 3(a). Considering $\tau_1 = 0.4$, it is fairly evident that, outside of the homoskedastic case, PSY is generally less powerful than $sPSY$, and PWY is less powerful than $sPWY$. It is also evident that PWY can perform very poorly compared to $sPWY$ when $\sigma_1 < 1$. When $\tau_1 = 0.8$, PSY is dominated by $sPSY$ for $\sigma_1 < 1$ but now shows some gains when $\sigma_1 > 1$. PWY is generally now more powerful than $sPWY$ unless $\sigma_1 < 1$, where $sPWY$ can offer substantial gains. In Table 3(b) we give the results for DGP 2. For $\tau_1 = 0.1$, PSY is less powerful than $sPSY$ and it is noticeable that PSY can have much lower power for $\sigma_1 < 1$ - something which was not observed under DGP 1. The same is also true when we compare PWY with $sPWY$. For $\tau_1 = 0.5$, PSY remains inferior to $sPSY$ when $\sigma_1 < 1$ while PWY is now better performing than $sPWY$ unless $\sigma_1 < 1$, where the ranking is reversed. Under DGP 4 in Table 3(c), the results are throughout very similar to those found for DGP 2 in Table 3(b), suggesting that the addition of the collapse period, in itself, has very little effect on the power of these suprema-based tests. As such, similar comments apply here as made under DGP 2. Under heteroskedasticity, then, our results clearly demonstrate that $sPSY$ and $sPWY$ can, in terms of asymptotic power, be considered as very worthy competitors to their standard counterparts. Particularly, but by no means exclusively, they have better power properties for downward volatility shifts; a case which proves to be a distinct weakness for PWY throughout.

Tables 3(a)-3(c) also report (infeasibly size-adjusted) local asymptotic power results for variants of PSY and PWY that exclude an intercept in the Dickey-Fuller regressions (cf. Remark 1), which we denote by PSY_0 and PWY_0. These results are obtained under the assumption that $\mu = o(T^{1/2})$, in which case the limit distributions for PSY_0

and PWY_0 are invariant to μ and take the same form as those for PSY and PWY, respectively, as given in section 2.1, but with $\tilde{U}(r)$ replaced by $U(r)$. Comparing PSY_0 and PWY_0 with PSY and PWY, it is clear that, as would be expected, exclusion of the intercept term results in superior local asymptotic power when $\mu = o(T^{1/2})$ is satisfied (apart from a few minor exceptions for PWY). What is noteworthy is that in cases where $sPSY$ and $sPWY$ display power gains over their PSY and PWY counterparts, $sPSY$ and $sPWY$ can often also achieve power gains over PSY_0 and PWY_0. In what follows, we retain our main emphasis on the original PSY and PWY tests rather than the PSY_0 and PWY_0 variants, due to the potential for the latter to have finite sample behaviour influenced by the unknown nuisance parameter μ, an issue we revisit in the finite sample results of section 5.

One interesting observation from the local asymptotic power results is that under homoskedasticity, PSY and PWY are in general more powerful than $sPSY$ and $sPWY$, respectively, for large c_1, while the sign-based tests are in general more powerful for smaller values of c_1. The latter finding may appear surprising since the sign-based tests are motivated by heteroskedasticity considerations, but there is no theoretical reason why these procedures cannot perform better than the original PSY and PWY tests, since there are no optimality claims associated with PSY and PWY in terms of power under homoskedasticity. What is also interesting is that (on the basis of these limit simulations at least), $sPSY$ is *always* found to be more powerful than $sPWY$, under homoskedasticity and heteroskedasticity, while there is no such unambiguity present between PSY and PWY.

For $sPSY$ and $sPWY$, as a robustness check we also evaluated powers of these tests under an (implicit) assumption that $z_t \sim t(5)$ (a fat tailed distribution) setting $f(0) = \Gamma(3)/(\Gamma(5/2)\sqrt{5\pi}) = 0.380$. Since this value of $f(0)$ is little different to 0.399, the powers change very little, but are always slightly smaller than for $z_t \sim NIID(0,1)$ because the offset terms in (6) are smaller in absolute value (the power differences have a mean (standard deviation) [maximum] of 0.009 (0.006) [0.021] for $sPSY$ and 0.008 (0.006) [0.023] for $sPWY$).

Of the four statistics, arguably then, $sPSY$ seems to emerge as the one with the best overall performance, followed by PSY. However, since there is no unique ranking between these two tests, we can consider a simple method which attempts at harnessing the better power of each for a given DGP setting, via a *union of rejections* strategy, which we detail in the next section.

4 A union of rejections strategy

Our approach is fundamentally based around that of Harvey et al. (2009), who consider the problem of testing for a unit root in the presence of uncertainty surrounding whether or not a linear trend is present in the deterministic component by combining tests which do and do not allow for trends, rejecting the unit root null if either test rejects. In the current context we consider a combination of $sPSY$ and PSY, although the same method is directly applicable to a combination of $sPWY$ and PWY. Specifically, denoting the asymptotic ξ level null critical value of $sPSY$ by cv_ξ^s (from the $\sigma(s)$-

invariant $MM^s_{0,0}$ distribution) and that of PSY by cv_ξ (from the $\sigma(s)$-dependent $MM_{0,0}$ distribution) a union of rejections strategy can be written as the decision rule

$$\text{Reject } H_0 \text{ if } \{sPSY > \psi_\xi cv^s_\xi \text{ or } PSY > \psi_\xi cv_\xi\}$$

where ψ_ξ is a scaling constant that ensures the decision rule yields asymptotic size of ξ under H_0. Defining a single statistic $uPSY$ as

$$uPSY := \max\left(sPSY, \frac{cv^s_\xi}{cv_\xi}PSY\right)$$

the decision rule is then equivalent to

$$\text{Reject } H_0 \text{ if } uPSY > \psi_\xi cv^s_\xi =: cv^u_\xi.$$

An application of the continuous mapping theorem (CMT) along with the results in Theorems 1 and 2 yields the asymptotic distribution of $uPSY$ as

$$uPSY \Rightarrow \max\left(MM^s_{c_1,c_2}, \frac{cv^s_\xi}{cv_\xi}MM_{c_1,c_2}\right).$$

Note that this union of rejections strategy as it stands is doubly infeasible as the $uPSY$ statistic itself uses the $\sigma(s)$-dependent cv_ξ, and also the critical value cv^u_ξ is $\sigma(s)$-dependent via ψ_ξ. The scaling constant ψ_ξ can be determined from the limit distribution of $uPSY$ with $c_1 = c_2 = 0$, but there is actually no need to calculate it explicitly since, for a given value of cv^s_ξ/cv_ξ, all we actually require is the critical value cv^u_ξ which is obtained directly from the null limit distribution of $uPSY$.

Infeasibly size-corrected limit powers for $uPSY$ and its $sPWY/PWY$-based counterpart, denoted $uPWY$, are also shown in Tables 3(a)-(c). The immediate feature we observe for the union procedures is that, throughout, their power levels are always really quite close to the higher of the two constituent tests. This is never something that can be guaranteed in general with such union-based procedures, due to the implicit scaling constant ψ_ξ essentially having the effect of always inflating the critical values applied to each constituent test. Here, however, the impact of this scaling appears to be really rather modest, thereby rendering the union a rather effective tool in this particular instance.

Thus far only $sPSY$ and $sPWY$ represent properly feasible test procedures as they are asymptotically size controlled without requiring knowledge of σ_t. For PSY and PWY, and $uPSY$ and $uPWY$, asymptotic size control can be obtained by employing a wild bootstrap scheme to construct the relevant critical values. This is shown to be valid in the context of the PWY test in HLST, and we now outline how this applies to PSY and $uPSY$.

The wild bootstrap algorithm is:

1. Generate a wild bootstrap sample $\{y^b_t\}^T_{t=1}$ by setting

$$y^b_1 = 0, \quad y^b_t = y^b_{t-1} + \Delta y_t w_t, \quad t = 2, .., T$$

where the w_t are $NIID(0, 1)$ variates.

2. Use the wild bootstrap sample to compute the pair of statistics PSY and $sPSY$.

3. Repeat step 1 and step 2 M times, denoting the resulting pairs of statistics by $\{PSY_1^b, sPSY_1^b\}, ..., \{PSY_M^b, sPSY_M^b\}$.

Note that under H_0, since $sign(\Delta y_t^b) = sign(z_t w_t)$, $sPSY_m^b$ is exact invariant to σ_t.

The next Theorem details the joint asymptotic distribution of PSY_m^b and $sPSY_m^b$ under a locally explosive H_1.

Theorem 3 *Under H_1 with $\delta_{i.T} = c_i T^{-1}$, $c_i > 0$, $i = 1, 2$ and Assumptions A1-A4*

$$\begin{pmatrix} PSY_m^b \\ sPSY_m^b \end{pmatrix} \overset{p}{\Rightarrow} \begin{pmatrix} MM_{0,0} \\ MM_{0,0}^s \end{pmatrix}$$

jointly, for any $1 \leq m \leq M$.

The marginal convergence result regarding PSY_m^b follows directly from HLST. Noting that $sign(\Delta y_t^b) = sign(z_t w_t) + o_p(1)$, the proof of the marginal convergence result for $sPSY_m^b$ follows the same strategy as HLST and the proof of Theorem 2 of this paper. The joint convergence occurs because both statistics are calculated from the *same* bootstrap sample (this result is needed below for the asymptotic validity of the union of rejections strategy). The proof of Theorem 3 is therefore straightforward and omitted for the sake of brevity. The Theorem demonstrates that the wild bootstrap procedure is first order valid in approximating the asymptotic joint *null* distribution of the PSY and $sPSY$ statistics under a locally explosive H_1 (which includes H_0 as a special case).

The ξ level bootstrap critical values are obtained from the empirical distribution functions of PSY_m^b and $sPSY_m^b$ calculated from M bootstrap replications. Denoting these critical values as cv_ξ^b and $cv_\xi^{b,s}$, a rejection of H_0 for PSY is obtained if $PSY > cv_\xi^b$ and a rejection of H_0 for $sPSY$ is obtained if $sPSY > cv_\xi^{b,s}$. As $T, N \to \infty$, it follows that cv_ξ^b and $cv_\xi^{b,s}$ converge in probability to cv_ξ and cv_ξ^s, so these individual bootstrap procedures are correctly sized in the limit. Consequently, PSY inherits exactly the same asymptotic local power properties under H_1 as its infeasibly size-corrected counterpart of section 3.2 (this is trivially true of $sPSY$ as cv_ξ^s does not depend on σ_t).

The wild bootstrap counterpart of the union statistic $uPSY$ is given by

$$uPSY_m^b := \max\left(sPSY_m^b, \frac{cv_\xi^{b,s}}{cv_\xi^b} PSY_m^b\right)$$

for $m = 1, ..., M$. It follows from Theorem 3 that

$$uPSY_m^b \overset{p}{\Rightarrow} \max\left(MM_{0,0}^s, \frac{cv_\xi^s}{cv_\xi} MM_{0,0}\right).$$

The ξ level bootstrap critical value for the union is obtained from the empirical distribution function of $uPSY_m^b$, and denoting this critical value as $cv_\xi^{b,u}$ we reject H_0 when $uPSY^b > cv_\xi^{b,u}$, where

$$uPSY^b := \max\left(sPSY, \frac{cv_\xi^{b,s}}{cv_\xi^b} PSY\right).$$

Here $uPSY^b$ is a feasible variant of $uPSY$ that replaces cv_ξ^s/cv_ξ with $cv_\xi^{b,s}/cv_\xi^b$. Note that this approach does not require knowledge of the scaling constant ψ_ξ, as the size control is obtained implicitly using the bootstrap critical values. As $T, N \to \infty$, $uPSY^b$ is correctly sized in the limit under H_0 because $cv_\xi^{b,u}$ converges in probability to cv_ξ^u, and it has the same limiting local power function as $uPSY$ under H_1. An entirely analogous wild bootstrap approach can be implemented for PWY (as in HLST), $sPWY$ and $uPWY$.

5 Finite sample size and power of the tests

We now turn to an examination of the finite sample properties of the various wild bootstrap procedures. Our simulations are based on the model (1) with $T = 100$. Here we set $\mu = 0$ and $u_1 = \varepsilon_1$, where $\varepsilon_t = \sigma_t z_t$ with the z_t generated as $NIID(0, 1)$ random variates. Table 4 shows 0.05-level finite sample sizes; Tables 5(a)-(b) report powers for the constellations of parameter settings used in our previous asymptotic size and power simulations given in Table 2 and Tables 3(a)-(b). For brevity we omit results pertaining to DGP 4. Here the limit volatility functions $\sigma(s)$ are discretised to $\sigma_t(t/T)$ in an obvious way. Once again 2000 Monte Carlo replications are used and we employ $M = 499$ bootstrap replications.

As regards finite sample size accuracy, there is a definite tendency for (bootstrap) PSY to be undersized, with size often dropping below 0.03 (and occasionally below 0.01, including in the homoskedastic case). In comparison, the size of (bootstrap) $sPSY$ (whose size is invariant to σ_1) is reasonably accurate at 0.059. Interestingly, the undersize of PSY does not translate into substantial undersize of the (bootstrap) union $uPSY$; its size is never below 0.04, so $sPSY$ is clearly having an offsetting effect within the combination. PWY also has some tendency to undersize (unless $\sigma_1 = 6$ when it can be modestly oversized), but to a lesser degree than PSY. The size of $sPWY$ is very accurate at 0.054 and $uPWY$ offers better size control than PWY.

Considering finite sample power, Table 5(a) gives results for DGP 1. When $\tau_1 = 0.4$, PSY is generally less powerful than $sPSY$. In fact, the latter's superiority in this regard is more readily apparent here than in the asymptotic context (Table 3(a)); this is probably a manifestation of the undersizing of PSY noted above. There is no clear winner when comparing PWY and $sPWY$, unlike in the asymptotic case where $sPWY$ was generally the better performing test. However, it is still the case that PWY can have very low power compared to $sPWY$ when $\sigma_1 < 1$. When $\tau_1 = 0.8$, PSY is dominated by $sPSY$ unless $\sigma_1 > 1$ and PWY outperforms $sPWY$ unless $\sigma_1 < 1$ which is similar to the asymptotic case although the magnitudes involved can differ between the finite sample and asymptotic cases. The finite sample results for DGP 1 would reasonably suggest that $sPSY$ is a better performing test than PSY overall; while there is no clear ranking between PWY and $sPWY$, PWY can have very low power for $\sigma_1 < 1$. Results for DGP 2 are given in Table 5(b). When $\tau_1 = 0.1$, we see that PSY and PWY are, respectively, less powerful than $sPSY$ and $sPWY$, and the differences are often particular marked. For $\tau_1 = 0.5$, PSY is again generally inferior to $sPSY$, while PWY performs better than $sPWY$ unless $\sigma_1 < 1$, in which case $sPWY$ can offer substantial gains. These findings are again largely in accordance with the asymptotic results (Table 3(b)).

Throughout Table 5, we also see that the union procedures have power levels close to whichever constituent test is displaying the higher power under given circumstances. Overall then, it would appear reasonable to conclude that our asymptotic power simulation results provide a decent indicator of how the various bootstrap procedures will perform in practice, even when only data series of modest length are available. Lastly, we note that while we have implemented the $sPSY$ and $sPWY$ tests using bootstrap critical values here, we could of course simply use their finite sample critical values from Table 1 ($T = 100$). Upon doing this, we found the difference in powers to have a mean (standard deviation) [maximum] of 0.017 (0.011) [0.041] for $sPSY$ and 0.005 (0.004) [0.014] for $sPWY$, implying the two critical value methods yield insignificantly different results.

Finally, in Table 6 we report finite sample size and power results for wild bootstrap versions of the PSY_0 and PWY_0 tests that exclude an intercept from the Dickey-Fuller regressions. Here we consider the illustrative case of DGP 1 using the same settings as in the results of Tables 4 and 5, only now we consider a range of values for μ ($\mu \in \{0, 10, 20, 50\}$) in order to investigate the sensitivity of these tests to μ in finite samples (note that the bootstrap critical values do not depend on μ). When $\mu = 0$, the bootstrap PSY_0 test displays some oversize, while bootstrap PWY_0 has approximately correct size unless $\sigma_1 > 1$ where it can be modestly oversized. As μ increases, we observe an increase in the upward size distortion for PSY_0 when $\sigma_1 \leq 1$ and a decrease in size when $\sigma_1 > 1$, to the point that undersize can be displayed for large σ_1 and μ. On the other hand, PWY_0 remains approximately correctly sized when $\sigma_1 \leq 1$ but can be very undersized for $\sigma_1 > 1$ and large μ. Turning to the test powers, even greater sensitivity to μ can be observed, and we observe that an increase in μ corresponds to a decrease in finite sample power for PSY_0 and PWY_0. Even for $\mu = 10$, the power of PSY_0 and PWY_0 can be reduced by up to 0.175 and 0.137, respectively, compared to the $\mu = 0$ case. These reductions in power become more marked as μ increases, with PSY_0 and PWY_0 powers for $\mu = 50$ reduced by up to 0.470 and 0.414, respectively, compared to when $\mu = 0$. Overall, we observe that the finite sample properties of the PSY_0 and PWY_0 tests are highly sensitive to the unknown DGP parameter μ, reinforcing our recommendation for use of the PSY/PWY and $sPSY/sPWY$ procedures which offer robustness (exact invariance) to μ.

6 Dating bubble start and end points using sign-based statistics

We now consider how to consistently estimate the bubble start and end points, τ_1 and τ_2 when $\delta_{1,T} > 0$. For simplicity, we examine this issue within the context of DGP 2. Consistent estimation is not possible using the current Pitman drift bubble magnitude $\delta_{1,T} = c_1 T^{-1}$, $c_1 > 0$, so in what follows we replace this with a stronger, mildly explosive bubble magnitude of the form $\delta_{1,T} = c_1 T^{-\alpha}$ where $\alpha \in (0,1)$.[3] PWY and PSY propose dating strategies for the start and end points of the bubble based on repeated implementation of their recursive tests over expanding samples, using critical values that diverge to infinity but at a rate slower than the derived divergence rate of the statistics over the mildly explosive regime. We considered implementing the PWY/PSY dating approach using our sign-based statistics, and it is not difficult to show that the sign-based statistics diverge at the rate $T^{1/2}$ in the mildly explosive regime, allowing τ_1 to be estimated consistently.[4] However, the sign-based statistics are unable to consistently estimate τ_2 with the PWY/PSY dating strategy.[5] In view of this, we pursue a dating strategy based on maximising sub-sample statistics.

By way of motivation, in view of the $sPSY$ statistic of (4), fairly intuitive estimators of (τ_1, τ_2) are provided by the maximisers

$$(\hat{\tau}_1, \hat{\tau}_2) := \underset{\lambda_1 \in [0,1-\pi]\ \lambda_2 \in [\lambda_1+\pi,1]}{\arg\max} sDF(\lambda_1, \lambda_2).$$

However, while it can be shown (under Assumptions A1-A4) that $\hat{\tau}_2 \xrightarrow{p} \tau_2$ we find that $\hat{\tau}_1$ does not consistently estimate τ_1. A solution to this inconsistency is obtained on replacing $\hat{s}^2(\lambda_1, \lambda_2)$ with the quantity $(\tilde{s}^2(\lambda_1, \lambda_2))^{\varepsilon}$, for some $0 < \varepsilon \leq 1$, where

$$\tilde{s}^2(\lambda_1, \lambda_2) := \frac{\lfloor \lambda_2 T \rfloor \hat{s}^2(0, \lambda_2) - \lfloor \lambda_1 T \rfloor \hat{s}^2(0, \lambda_1)}{\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 1}$$

with the definitions of $\hat{s}^2(0, \lambda_1)$ and $\hat{s}^2(0, \lambda_2)$ being implied from Section 2.2. We can then state a result which pertains to the modified statistic

$$sDF^*(\lambda_1, \lambda_2) := \frac{\hat{\rho}(\lambda_1, \lambda_2)}{\sqrt{(\tilde{s}^2(\lambda_1, \lambda_2))^{\varepsilon} / \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} C_{t-1}^2}}$$

by showing that (after suitable normalisation) the probability limit of $sDF^*(\lambda_1, \lambda_2)$ is a well-defined (piecewise) *deterministic* function, with (τ_1, τ_2) being its unique maximiser. In the next theorem, we show that the maximiser of the dating statistic

[3] The consistency results in this section also hold when $\alpha = 0$, which represents a fixed magnitude (rather than mildly explosive) bubble.

[4] The PWY and PSY dating statistics diverge at a rate dependent on the mildly explosive parameter α. By construction, our sign-based statistics are independent of the actual magnitude of the mildly explosive bubble, yielding a fixed rate of divergence $T^{1/2}$ for all $\alpha \in (0,1)$.

[5] Based on our later results in Lemma 1, it is not difficult to see that the order of divergence of our sign-based statistics is unchanged before and after τ_2, hence this change point cannot be identified by directly applying the PWY/PSY dating strategy.

$sDF^*(\lambda_1, \lambda_2)$, not unsurprisingly, can estimate the maximiser (τ_1, τ_2) of the limiting function consistently.

Theorem 4 *For model (1), under DGP 2 with $\delta_{1,T} = c_1 T^{-\alpha}$, $c_1 > 0$, $\alpha \in (0, 1)$ and Assumptions A1-A4, with*

$$(\tilde{\tau}_1, \tilde{\tau}_2) := \underset{\lambda_1 \in [0,1-\pi] \ \lambda_2 \in [\lambda_1+\pi,1]}{\arg\max} sDF^*(\lambda_1, \lambda_2)$$

then, provided $\tau_2 - \tau_1 \geq \pi$,

$$(\tilde{\tau}_1, \tilde{\tau}_2) \xrightarrow{p} (\tau_1, \tau_2)$$

for $0 < \varepsilon \leq 1$.

In essence, replacing $\hat{s}^2(\lambda_1, \lambda_2)$ with $\tilde{s}^2(\lambda_1, \lambda_2)$, which is a weighted average of $\hat{s}^2(0, \lambda_2)$ and $\hat{s}^2(0, \lambda_1)$, creates a kink in the limiting function, making (τ_1, τ_2) a unique maximiser of it, while raising $\tilde{s}^2(\lambda_1, \lambda_2)$ to the power ε relates to finite sample considerations. Since Theorem 1 holds for all $0 < \varepsilon \leq 1$, it might be tempting simply to set $\varepsilon = 1$ for any practical application. However, unreported simulation evidence suggests the finite sample properties of this choice can be very poor, with the estimator $\tilde{\tau}_1$ ($\tilde{\tau}_2$) being badly biased upwards (downwards). Setting ε to a much smaller value, such as $\varepsilon = 0.01$, was found to yield much less biased estimators. The results of Theorem 4 can be shown to continue to hold in the presence of any form of bubble collapse (as in DGP 3 or DGP 4, or indeed an instantaneous collapse). They also hold under DGP 1, i.e. for $\tau_2 = 1$.

Note that the consistency result in Theorem 4 requires $\tau_2 - \tau_1 \geq \pi$, so that the length of the bubble regime is at least as long as the minimum window width for which $sDF^*(\lambda_1, \lambda_2)$ is computed. To allow for bubbles of short length, for example a bubble that emerges late in the sample period (under DGP 1), a relatively small value of π may be appropriate for accurate dating. Note that the setting for π here does not need to coincide with the setting for π in the testing procedure.

In principle, we could also use $sDF^*(\lambda_1, \lambda_2)$ in place of $sDF(\lambda_1, \lambda_2)$ in our testing setup. This would unify the testing and dating aspects of our procedure rather conveniently. Unfortunately, the finite sample size and power properties of $sDF^*(\lambda_1, \lambda_2)$ we found to be somewhat inferior in comparison to those of $sDF(\lambda_1, \lambda_2)$ and we therefore cannot recommend such a strategy.

7 Asymmetric errors

Assumption A4 implies that the mean and median of z_t are the same (zero). It is possible that this assumption could fail to hold, for example Campbell et al. (1997, Table 1.1) find (very) mild asymmetries in daily financial returns which is likely to imply violation of Assumption A4. Suppose, then, that $E(z_t) = 0$ but $F(0) \neq 1/2$. In this case, under the null hypothesis, $E\{sign(z_t)\} \neq 0$ and the invariance principle (2) clearly breaks down. It is then obvious that some form of de-meaning of $sign(z_t)$ is required to make progress. This could be carried out in a number of different ways, but a convenient method is to employ recursive de-meaning of $sign(\Delta y_t)$ before cumulating to form C_t. Specifically, we replace $sign(\Delta y_t)$ in the construction of $sDF(\lambda_1, \lambda_2)$ with

$$sign(\Delta y_t) - (t-1)^{-1} \sum_{j=2}^{t} sign(\Delta y_j) \tag{7}$$

which is invariant to $E[sign(z_t)]$. The advantage of recursive de-meaning is that (7) only involves data up to time t, which is, of course, relevant for any kind of real-time bubble monitoring exercise (full-sample de-meaning, for example, would not have this property).[6]

Under a locally explosive H_1 and Assumptions A1-A3, we can show (along the lines of the proof of Theorem 2) that

$$\begin{aligned} T^{-1/2} \sum_{t=2}^{\lfloor rT \rfloor} \left\{ sign(\Delta y_t) - (t-1)^{-1} \sum_{j=2}^{t} sign(\Delta y_j) \right\} \quad &\Rightarrow \quad \sigma_{sz} \left\{ U^s(r) - \int_0^r x^{-1} U^s(x) dx \right\} \\ &=: \quad \sigma_{sz} \tilde{U}^s(r) \end{aligned}$$

where $\sigma_{sz}^2 = Var[sign(z_t)]$ and $U^s(r)$ is a copy of the distribution of that given in (6) (we duplicate the notation only to avoid repeating each expression). Then, denoting the new statistic by $\bar{s}PSY$, we find that

$$\bar{s}PSY \Rightarrow \sup_{\lambda_1 \in [0, 1-\pi]} \sup_{\lambda_2 \in [\lambda_1 + \pi, 1]} \tilde{L}^s(\lambda_1, \lambda_2)$$

where

$$\tilde{L}^s(\lambda_1, \lambda_2) := \frac{\tilde{U}^s(\lambda_2)^2 - \tilde{U}^s(\lambda_1)^2 - (\lambda_2 - \lambda_1)}{2\sqrt{\int_{\lambda_1}^{\lambda_2} \tilde{U}^s(r)^2 dr}}.$$

Asymptotic and finite sample critical values for this test, and its $sPWY$ counterpart, denoted $\bar{s}PWY$, are given in Table 6. The corresponding bootstrap statistics are based on recursive de-meaning of $sign(\Delta y_t^b)$. The consistency results of Theorem 4 can also be shown to hold in the case of recursive de-meaning.

With traditional left-tail unit root testing, it is well known that any form of de-trending of the data to account for deterministic terms in the observed series reduces the power of the test relative to the case where no de-trending is required. We would have little reason to suggest the same will not happen in the current context of right-tail testing. Since recursive de-meaning of $sign(\Delta y_t)$ is *de facto* equivalent to de-trending of C_t, we should therefore expect to find $\bar{s}PSY$ and $\bar{s}PWY$ to have lower power than $sPSY$ and $sPWY$, respectively. We examine the extent to which this occurs both asymptotically and in finite samples. In addition to the results for $sPSY$ and $sPWY$,

[6]Note that if a bubble is present from the beginning of the sample period, it is theoretically possible (although unlikely in practice) that all the values of $sign(\Delta y_t)$ are equal to one, in which case the recursively de-meaned series would be zero for all time periods. While this causes problems for calculation of the test statistic, it is clear that such an occurrence should be taken as evidence of a bubble.

Table 3 also reports asymptotic local powers for $\bar{s}PSY$ and $\bar{s}PWY$, along with results for the corresponding union of rejections procedures $\bar{u}PSY$ (union of $\bar{s}PSY$ and PSY) and $\bar{u}PWY$ (union of $\bar{s}PWY$ and PWY). In Tables 3(a) and 3(c) we observe a loss in power through using $\bar{s}PSY$ or $\bar{s}PWY$ compared to $sPSY$ or $sPWY$, as anticipated, with these being most apparent for smaller values of c_1. Although some of the power gains that $\bar{s}PSY$ and $\bar{s}PWY$ offered over PSY and PWY are removed, there are still many cases where the sign-based approach outperforms the standard tests, sometimes by a substantial margin. In Table 3(b) we observe the unexpected feature that, for some DGP settings, $\bar{s}PSY$ and $\bar{s}PWY$ can have higher local asymptotic power than the non-recursively demeaned variants $sPSY$ and $sPWY$, respectively. Given that the latter tests were already seen to outperform PSY and PWY in a majority of cases for DGP 2, it follows that $\bar{s}PSY$ and $\bar{s}PWY$ offer valuable power gains relative to the original PSY and PWY tests also. Throughout Table 3, the union of rejections procedures $\bar{u}PSY$ and $\bar{u}PWY$ behave in a similar way to the union procedures of section 4, with power levels displayed that are close to the better of the two constituent tests that comprise each union.

Tables 4 and 5 report finite sample size and power results for the recursively de-meaned variants of the tests (and the corresponding unions), using bootstrap critical values throughout as in section 5. Table 4 shows the sizes of $\bar{s}PSY$ and $\bar{s}PWY$ to be close to the nominal level in finite samples, and the corresponding $\bar{u}PSY$ and $\bar{u}PWY$ union sizes are similar to those for $uPSY$ and $uPWY$. In Table 5, the finite sample power results for $\bar{s}PSY$ and $\bar{s}PWY$ follow broadly similar patterns to the asymptotic results, although for many settings the finite sample powers can be considerably lower than their asymptotic counterparts. This is particularly noticeable in Table 5(b) where the unexpected result that $\bar{s}PSY$ and $\bar{s}PWY$ had higher local asymptotic power than $sPSY$ and $sPWY$ is now reversed in finite samples. While our finite sample results are limited to $T = 100$, unreported simulations using larger finite sample sizes confirm that the finite sample powers of the tests converge to the local asymptotic results in Table 3. Table 5 also shows that, once again, each union of rejections procedure has power close to the better of its constituent tests.

8 Extensions and discussions

8.1 Higher order dynamics

We have assumed thus far that ε_t is serially uncorrelated. More generally, we may consider it to have an autoregressive representation of the form

$$\varepsilon_t = \sum_{i=1}^{p} \rho_i \varepsilon_{t-i} + \sigma_t z_t$$

with ρ_i such that ε_t is stationary under homoskedasticity. In this case, in the spirit of the recursive de-meaning described above, we fit the recursive OLS regressions

$$\Delta y_j = \hat{\alpha}(t) + \hat{\phi}(t) y_{j-1} + \sum_{i=1}^{p} \hat{\rho}_i(t) \Delta y_{j-i} + e_j, \qquad j = p+5, ..., t. \tag{8}$$

We then construct $sDF(\lambda_1, \lambda_2)$ using $sign(\Delta y_t - \sum_{i=1}^{p} \hat{\rho}_i(t) \Delta y_{t-i})$. The null limit distribution of $sPSY$ can be shown to remain the same as given in section 2.2. There is no need to alter the bootstrap data generation scheme nor the form of $sPSY_m^b$ because the wild bootstrap removes any weak dependence present in Δy_t. In practice, p is unknown but could be determined using standard information criteria, for example BIC.

8.2 More general deterministic terms

If the constant deterministic term μ in (1) is replaced by a process undergoing a finite number, n say, of deterministic level shifts, the limit distributions of the sign-based statistics are unchanged. This occurs because only n of the $sign(\Delta y_t)$ are affected, so the effect is asymptotically negligible. Moreover, there is no restriction on the *magnitudes* of the level shifts due to the sign transformation. The limit distributions of PSY and PWY are also unchanged, but only provided the level shift magnitudes are $o(T^{1/2})$. The practical consequence of this is that, in finite samples, the sizes and powers of PSY and PWY will be rather more sensitive to large level shifts than those of the sign-based tests.

Phillips et al. (2014) consider the possibility of a local-to-zero drift term. In the context of our model, this translates to replacing μ with $\mu + \beta T^{-d} t$, where d is a positive constant. It can be shown that the null limit distributions for the sign-based statistics continue to hold provided $d > 1/2$, coinciding with the restriction that PSY and PWY require for their tests to be asymptotically invariant to the local drift. To see this, define

$$\Lambda_T := T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} sign(\Delta y_t) - T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} sign(z_t).$$

It is straightforward to calculate that

$$\begin{aligned} E(\Lambda_T) &= 2T^{-1/2} \sum_{t=1}^{\lfloor rT \rfloor} \left(F(0) - F\left(-\frac{\beta T^{-d}}{\sigma_t} \right) \right) \\ &= O(T^{\frac{1}{2}-d}) \end{aligned}$$

as $F(0) - F\left(-\frac{\beta T^{-d}}{\sigma_t}\right)$ is $O(T^{-d})$ uniformly for all t. Also, $Var(\Lambda_T) = O(T^{-d})$ using a similar argument. A simple application of Markov's inequality then shows that $\Lambda_T = o_p(1)$ when $d > 1/2$.

9 An empirical illustration

By way of a practical illustration of the use of our sign-based tests and dating methods, we apply them to Bitcoin price data (measured in pounds sterling) to study the possibility of explosive behaviour being present in Bitcoin prices from late 2017. Bitcoin is a digital asset designed to work as a medium of exchange that uses cryptography (a so-called "cryptocurrency") and is considered a speculative asset among economists. The data range we choose is for the period 1/9/2017 to 28/1/2018. Bitcoin is traded 24/7 globally so price observations are available on all days, giving 149 observations. The data, which is plotted in Figure 1, is the daily closing price and was obtained from the website https://finance.yahoo.com/quote/BTC-GBP. In what follows, testing and dating are based on setting $\pi = 0.1$.

Table 7 shows the value of the statistics PSY, $sPSY$ and $\bar{s}PSY$, with $p = 1$ in (8) (selected by BIC assuming a maximum value of $p = 4$) and one lagged difference included in the OLS regressions underlying PSY (a small number of observations are lost through accounting for serial correlation; for consistency we compute all tests over the same effective sample size). The entries in round brackets are bootstrap p-values for the tests based on $M = 9999$ bootstrap replications. The PSY test clearly fails to reject the null hypothesis (measured at any conventional significance level), while both $sPSY$ and $\bar{s}PSY$ show strong rejections. The strength of rejection obtained from $sPSY$ is slightly higher than for $\bar{s}PSY$, which might be expected in view of the simulation results of section 7 above. Table 7 also reports results for the feasible union of rejections procedures $uPSY^b$ and $\bar{u}PSY^b$. The p-values associated with these procedures imply rejection of the null in both cases, albeit at a slightly higher significance level than was found for the $sPSY$ and $\bar{s}PSY$ tests, as would be expected. It can also be seen that the values of the $uPSY^b$ and $\bar{u}PSY^b$ statistics coincide with the $sPSY$ and $\bar{s}PSY$ statistics, consistent with the rejections coming from the sign-based tests rather than the original test.

The additional entry for PSY in square brackets is the bootstrap p-value obtained when we do *not* account for any heteroskedasticity, which we carry out by constructing the increments of y_t^b using w_t instead of $\Delta y_t w_t$. This approach is then essentially the same as using standard finite sample critical values obtained using $NIID(0, 1)$ errors. Interestingly, this leads to a complete overturn of the previous non-rejection by PSY. That this occurs, however, we take as a potential indication of substantial levels of heteroskedasticity being present in the data, and therefore an indication of the need to correct for it before we are in a position to make size-controlled inference. These contrasting findings are perhaps particularly pertinent given that changing volatility is widely considered to be a trademark characteristic of Bitcoin price data. The plot of absolute price changes shown in Figure 1 would seem to support such a view.

Having provided significant evidence for the presence of explosivity in the data on the basis of $sPSY$ and $\bar{s}PSY$, we can proceed to date it (we do not attempt to date using the PSY approach given that PSY failed to reject). Using the dating statistic sDF^* (with $\varepsilon = 0.01$), the start date for the explosive regime is identified as 13/11/2017. The Bitcoin price suffered a small crash from 8/11/2017, and was undergoing a continuous 5-day decrease until 12/11/2017, after which it started a rapid increase period until mid December 2018. As such, the sDF^* statistic seems to be reasonably accurate in identifying 13/11/2017 as the start of the period of rapid increase. The sDF^* statistic finds that 7/12/2017 is the end date of the explosive regime. On 7/12/2017, the Bitcoin price reached a local maximum after continuous increasing of about 3 weeks from £4379, closing at £12882. Then it suffered a short one-day crash, before it was pushed to its historical high on 22/12/2017. Our sDF^* therefore seems to be rather accurate in identifying this crash, placing the end time on the crash day (so no dating delay is inherent). We also apply the same dating strategy to the recursively de-meaned data, denoting this statistic as $\bar{s}DF^*$. We find that the $\bar{s}DF^*$ statistic places the start date rather earlier than sDF^*, at 4/10/2017. This essentially treats the one-week crash starting from 8/11/2017 as a random shock and also picks up the relatively gradual increase in the Bitcoin price from 10/2017 as an explosive regime, which also seems reasonable to us. Notice that the end date identified by $\bar{s}DF^*$ is identical to that from sDF^*. As such, sDF^* suggests an explosive regime that is concentrated around the period of most intensive upward movement in the prices, while $\bar{s}DF^*$ is suggestive of a more gradually emerging explosive regime. Both these scenarios seem plausible and we would not wish to take a stance in favour of one or other without conducting deeper analysis that lies outside of the remit of this paper.

10 Conclusions

In this paper we have proposed a sign-based variant of the PSY test for explosive autoregressive behaviour in financial time series. In contrast to the original test, this test does not require bootstrap-type methods to control size in the presence of heteroskedastic innovations, thereby offering computational efficiency gains when applied in practice. Under a locally explosive bubble alternative, we also find that the sign-based test has appealing asymptotic power properties, with the potential to deliver substantially greater power than the original test for many volatility and bubble model specifications. However, because the original test may still outperform the sign-based test for some specifications, we also suggested a union of rejections procedure that combines the sign-based and original tests and employs a joint wild bootstrap to control size. This union is seen to succeed in capturing most of the power available from the better performing of the two tests for a given alternative. Our finite sample simulations indicate that our new procedures should work well in practice. We have also shown how a slight variant of the sign-based test can be used to consistently date the start and end points of a mildly explosive bubble, and how a recursively de-meaned variant of the test can allow for asymmetry in the innovations. We applied our sign-based testing and dating procedures to recent Bitcoin price data and uncovered robust evidence for the existence of an explosive regime in this data, subsequently identifying what we consider to be plausible start and end dates for this regime.

Finally, we consider some areas for future research. First, our assumptions consider only deterministically time-varying volatility, and it would be interesting to extend our work to ARCH-type and stochastic volatility dynamics, investigating conditions under which our tests remain asymptotically valid. Secondly, we recognise that the DGP we have chosen to analyse in this paper is not the only specification capable of generating bubble-type behaviour. Random coefficient autoregressive models (e.g. Blanchard and Watson, 1982, Granger and Swanson, 1997) and certain types of noncausal model (e.g. Gourieroux and Jasiak, 2018) represent plausible alternative specifications. An investigation of the performance of our sign-based tests in the context of such alternative models would be interesting. Finally, in principle the procedures developed in this paper could be extended to the context of real-time detection and dating of possibly multiple bubbles. Additional issues arise when considering a real-time analysis, for example the need to control size for tests applied at multiple sequential points in time (see, e.g., Homm and Breitung, 2012), and a full development of such real-time monitoring procedures would also be of value.

References

Beare, B.K. (2018). Unit root testing with unstable volatility. *Journal of Time Series Analysis*, forthcoming.

Blanchard, O.J. and Watson M.W. (1982). Bubbles, rational expectations and financial markets. In Wachtel, P. (ed.), *Crises in the Economic and Financial Structure*, pp.295–316. Lexington Books.

Boldin, M.V. (2013). On qualitatively robust sign test in random walk model. Proceedings of the 10th International Conference 'Computer Data Analysis and Modeling: Theoretical and Applied Stochastics', Minsk, Volume 1, 194–197.

Campbell, B. and Dufour, J.-M. (1995). Exact nonparametric orthogonality and random walk tests. *Review of Economics and Statistics* 77, 1–16.

Campbell, J.Y., Lo, A.W.C. and MacKinlay, A.C. (1997). *The Econometrics of Financial Markets.* Princeton University Press.

Christoffersen, P.F. (2012). *Elements of Financial Risk Management (2nd edn.).* Academic Press.

Gourieroux, C. and Jasiak J. (2018). Robust analysis of the martingale hypothesis. Discussion Paper, York University, Canada.

Granger, C.W.J. and Swanson N.R. (1997). An introduction to stochastic unit-root processes. *Journal of Econometrics* 80, 35-62.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2009). Unit root testing in practice: dealing with uncertainty over the trend and initial condition (with commentaries and rejoinder). *Econometric Theory* 25, 587–667.

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

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

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 Shi, S.-P. (2018a). Financial bubble implosion and reverse regression. *Econometric Theory* 34, 705–753.

Phillips, P.C.B. and Shi, S.-P. (2018b). Real time monitoring of asset markets: bubbles and crises. Discussion Paper.

Phillips, P.C.B., Shi, S.-P. and Yu, J. (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics* 76, 315–333.

Phillips, P.C.B., Shi, S.-P. and Yu, J. (2015). Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. *International Economic Review* 56, 1043–1077.

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

Rapach, D.E., Strauss, J.K. and Wohar, M.E. (2008). Forecasting stock return volatility in the presence of structural breaks. In: Rapach, D.E. and Wohar, M.E. (eds.), *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, pp. 381–416, Vol. 3 of Emerald Series Frontiers of Economics and Globalization, Emerald.

Shi, S.-P., Hurn, S. and Phillips, P.C.B. (2018a). Causal change detection in possibly integrated systems: revisiting the money-income relationship. Discussion Paper.

Shi, S.-P., Phillips, P.C.B. and Hurn, S. (2018b). Change detection and the causal impact of the yield curve. *Journal of Time Series Analysis* 39, 966–987.

So, B.S. and Shin, D.W. (2001). An invariant sign test for random walks based on recursive median adjustment. *Journal of Econometrics* 102, 197–229.

Tsay, R.S. (2010). *Analysis of Financial Time Series (3rd edn.).* Wiley Series in Probability and Statistics, Wiley.

Van der Vaart, A.W. (2000). *Asymptotic Statistics.* Cambridge University Press.

Appendix A: Proofs of Theorems

Without loss of generality we can set $\mu = 0$ in what follows.

Proof of Theorem 1

It follows from HLST that

$$\begin{aligned} T^{-1/2} y_{\lfloor rT \rfloor} &\Rightarrow \begin{cases} W_\sigma(r) & r \leq \tau_1 \\ V_1(r) & \tau_1 < r \leq \tau_2 \\ V_2(r) & \tau_2 < r \leq \tau_3 \\ V_2(\tau_3) + W_\sigma(r) - W_\sigma(\tau_3) & r > \tau_3. \end{cases} \\ &= U(r) \end{aligned} \tag{A.1}$$

(where $W_\sigma(r)$ corresponds to the variance-transformed Brownian motion process $\bar{\omega} W^\eta(r)$ in the notation of that paper). Also, since $\Delta y_t = \varepsilon_t + O_p(T^{-1/2})$ for all t, we find that $\hat{\sigma}^2(\lambda_1, \lambda_2) = (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} \Delta y_t^2 + o_p(1) \xrightarrow{p} (\lambda_2 - \lambda_1)^{-1} \int_{\lambda_1}^{\lambda_2} \sigma(r)^2 dr$.

A little straightforward manipulation allows us to write $DF(\lambda_1, \lambda_2)$ in the form

$$DF(\lambda_1, \lambda_2) = \frac{T^{-1}(y_{\lfloor \lambda_2 T \rfloor} - \bar{y})^2 - T^{-1}(y_{\lfloor \lambda_1 T \rfloor} - \bar{y})^2 - T^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} \Delta y_t^2}{2\sqrt{\hat{\sigma}^2(\lambda_1, \lambda_2) T^{-2} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} (y_{t-1} - \bar{y})^2}}.$$

Then, following the arguments of the proof of Theorem 1 in PSY, $DF(\lambda_1, \lambda_2)$ can be interpreted as a continuous functional of the partial sum process $T^{-1/2} y_{\lfloor rT \rfloor}$ and $\hat{\sigma}^2(\lambda_1, \lambda_2)$, allowing application of the continuous mapping theorem to give

$$\begin{aligned} DF(\lambda_1, \lambda_2) &\Rightarrow \frac{\tilde{U}(\lambda_2)^2 - \tilde{U}(\lambda_1)^2 - \int_{\lambda_1}^{\lambda_2} \sigma(r)^2 dr}{2\sqrt{(\lambda_2 - \lambda_1)^{-1} \int_{\lambda_1}^{\lambda_2} \sigma(r)^2 dr \int_{\lambda_1}^{\lambda_2} \tilde{U}(r)^2 dr}} \\ &= L_{c_1, c_2}(\lambda_1, \lambda_2). \end{aligned}$$

The stated limit for the PSY statistic is then obtained from a further continuous mapping argument following the proof of Theorem 1 in PSY, since the double sup operator can be written as a continuous functional over the space of functions $\{DF(\lambda_1, \lambda_2), (\lambda_1, \lambda_2) \in ([0, 1 - \pi], [\lambda_1 + \pi, 1])\}$ with respect to the uniform norm. See also Shi et al. (2018a) and Shi et al. (2018b) for similar arguments in deriving sup-type limit results.

Proof of Theorem 2

We will first show that

$$\begin{aligned} T^{-1/2} C_{\lfloor rT \rfloor} &\Rightarrow \begin{cases} W^s(r) & r \leq \tau_1 \\ W^s(r) + 2f(0) X_1(r) & \tau_1 < r \leq \tau_2 \\ W^s(r) + 2f(0) X_2(r) & \tau_2 < r \leq \tau_3 \\ W^s(r) + 2f(0) X_2(\tau_3) & r > \tau_3. \end{cases} \\ &= U^s(r) \end{aligned} \tag{A.2}$$

from which the main result follows easily. The result in (A.2) extends Theorem 1 of Boldin (2013) to allow for time-varying volatility and a time-varying coefficient mean level model. In what follows, we only demonstrate the result for the last regime where $r > \tau_3$; the results in the other regimes can be obtained in the same way.

First note that under H_1,

$$T^{-1/2}C_{\lfloor rT\rfloor} = T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathrm{sign}(\varepsilon_t) + \left(T^{-1/2}C_{\lfloor rT\rfloor} - T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathrm{sign}(\varepsilon_t)\right).$$

We first examine the difference $T^{-1/2}C_{\lfloor rT\rfloor} - T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathrm{sign}(\varepsilon_t)$. Using the definition $\mathrm{sign}(x) = -2\mathbf{1}(x \leqslant 0) + 1$, we have

$$\begin{aligned}
T^{-1/2}C_{\lfloor rT\rfloor} - T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathrm{sign}(\varepsilon_t) &= 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}(\mathbf{1}(\varepsilon_t \leqslant 0) - \mathbf{1}(y_t - y_{t-1} \leqslant 0)) \\
&= 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}(\mathbf{1}(\varepsilon_t \leqslant 0) - \mathbf{1}(\delta_t y_{t-1} + \varepsilon_t \leqslant 0)) \\
&= 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathbf{1}(\varepsilon_t \leqslant 0) - 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathbf{1}(\varepsilon_t \leqslant -\delta_t y_{t-1})
\end{aligned}$$

where $\delta_t = \delta_{1,T}\mathbf{1}(\lfloor\tau_1 T\rfloor < t \leqslant \lfloor\tau_2 T\rfloor) - \delta_{2,T}\mathbf{1}(\lfloor\tau_2 T\rfloor < t \leqslant \lfloor\tau_3 T\rfloor)$. Next, we make decompositions of the above two sums of indicator functions around the corresponding distribution function $F(.)$ of z_t:

$$\begin{aligned}
2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathbf{1}(\varepsilon_t &\leqslant 0) - 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\mathbf{1}(\varepsilon_t \leqslant -\delta_t y_{t-1}) \\
&= 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}(\mathbf{1}(\varepsilon_t \leqslant 0) - F(0)) \\
&\quad -2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\left(\mathbf{1}(\varepsilon_t \leqslant -\delta_t y_{t-1}) - F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)\right) \\
&\quad +2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\left(F(0) - F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)\right) \\
&= A + B + C
\end{aligned}$$

where A-C are defined implicitly.

Looking at terms A and B terms together and denoting

$$H_t := (\mathbf{1}(\varepsilon_t \leqslant 0) - F(0)) - \left(\mathbf{1}(\varepsilon_t \leqslant -\delta_t y_{t-1}) - F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)\right)$$

for $t = 2, \ldots, T$, we have

$$A + B = 2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}H_t.$$

Our aim is to evaluate the mean and variance of $A+B$. First notice that $\{H_t\}_{t=2}^T$ is a martingale difference sequence with respect to the natural filtration:

$$\begin{aligned} E(H_t|\mathcal{F}_{t-1}) &= E\left((\mathbf{1}(\varepsilon_t \leqslant 0) - F(0)) - \left(\mathbf{1}(\varepsilon_t \leqslant -\delta_t y_{t-1}) - F\left(-\delta_t \frac{y_{t-1}}{\sigma_t}\right)\right)|\mathcal{F}_{t-1}\right) \\ &= E\left((\mathbf{1}(z_t \leqslant 0) - F(0)) - \left(\mathbf{1}\left(z_t \leqslant -\delta_t \frac{y_{t-1}}{\sigma_t}\right) - F\left(-\delta_t \frac{y_{t-1}}{\sigma_t}\right)\right)|\mathcal{F}_{t-1}\right) \\ &= 0. \end{aligned}$$

This implies $E(A+B)=0$. Next,

$$\begin{aligned} \operatorname{Var}(H_t|\mathcal{F}_{t-1}) &= \operatorname{Var}\left(\mathbf{1}(z_t \leqslant 0) - \mathbf{1}\left(z_t \leqslant -\delta_t \frac{y_{t-1}}{\sigma_t}\right)|\mathcal{F}_{t-1}\right) \\ &\leqslant E\left(\left(\mathbf{1}(z_t \leqslant 0) - \mathbf{1}\left(z_t \leqslant -\delta_t \frac{y_{t-1}}{\sigma_t}\right)\right)^2 |\mathcal{F}_{t-1}\right) \\ &= E\left(\left(\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} < z_t \leqslant 0\right)\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} \leqslant 0\right)\right.\right. \\ &\qquad \left.\left. -\mathbf{1}\left(0 < z_t \leqslant -\delta_t \frac{y_{t-1}}{\sigma_t}\right)\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} > 0\right)\right)^2 |\mathcal{F}_{t-1}\right) \\ &= E\left(\left(\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} < z_t \leqslant 0\right)\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} \leqslant 0\right)\right.\right. \\ &\qquad \left.\left. +\mathbf{1}\left(0 < z_t \leqslant -\delta_t \frac{y_{t-1}}{\sigma_t}\right)\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} > 0\right)\right) |\mathcal{F}_{t-1}\right) \end{aligned} \tag{A.3}$$

where we have used the inequality $\operatorname{Var}(X|\mathcal{F}_{t-1}) \leqslant E(X^2|\mathcal{F}_{t-1})$ in the second step; in the last step, the cross product term of the quadratic expansion

$$E\left(\mathbf{1}\left(-\delta_t \frac{y_{t-1}}{\sigma_t} < z_t \leqslant 0\right)\mathbf{1}\left(0 < z_t \leqslant -\delta_t \frac{y_{t-1}}{\sigma_t}\right)|\mathcal{F}_{t-1}\right) = 0$$

as the two sets considered in the indicator functions are mutually exclusive. In the previous derivation for $\operatorname{Var}(H_t|\mathcal{F}_{t-1})$, we have also used the result that $E(\mathbf{1}(A)^2) = E(\mathbf{1}(A))$ for any set A. We have

$$\operatorname{Var}(A+B) = \operatorname{Var}\left(2T^{-1/2}\sum_{t=2}^{\lfloor rT \rfloor} H_t\right) = 4T^{-1}E\left(\left(\sum_{t=2}^{\lfloor rT \rfloor} H_t\right)^2\right).$$

Since $\{H_t\}$ is a martingale difference sequence, by Burkholder's inequality (e.g. Hall and Heyde (1980) Theorem 2.10), for a generic constant $\mathbb{C} > 0$, it is satisfied that

$$\begin{aligned}
4T^{-1}E\left(\left(\sum_{t=2}^{\lfloor rT \rfloor} H_t\right)^2\right) &\leqslant \mathbb{C}T^{-1}E\left(\sum_{t=2}^{\lfloor rT \rfloor} H_t^2\right) \\
&= \mathbb{C}T^{-1}E\left(\sum_{t=2}^{\lfloor rT \rfloor} E(H_t^2|\mathcal{F}_{t-1})\right) \\
&= \mathbb{C}T^{-1}E\left(\sum_{t=2}^{\lfloor rT \rfloor} \mathrm{Var}(H_t|\mathcal{F}_{t-1})\right) \\
&= \mathbb{C}T^{-1}E\left(\sum_{t=2}^{\lfloor rT \rfloor} E\left(\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} < z_t \leqslant 0\right)\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} \leqslant 0\right)\right.\right. \\
&\quad \left.\left. +\mathbf{1}\left(0 < z_t \leqslant -\delta_t\frac{y_{t-1}}{\sigma_t}\right)\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} > 0\right)|\mathcal{F}_{t-1}\right)\right) \\
&= \mathbb{C}T^{-1}\sum_{t=2}^{\lfloor rT \rfloor} E\left(\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} < z_t \leqslant 0\right)\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} \leqslant 0\right)\right. \\
&\quad \left. +\mathbf{1}\left(0 < z_t \leqslant -\delta_t\frac{y_{t-1}}{\sigma_t}\right)\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} > 0\right)\right)
\end{aligned}$$

where we have substituted in the expression derived in (A.3). Since $\delta_t y_{t-1} \xrightarrow{p} 0$ in the locally explosive regime, $\delta_t = 0$ identically in the unit root regimes, and $\delta_t y_{t-1} \xrightarrow{p} 0$ in the stationary regime, the set $\{-\delta_t y_{t-1}/\sigma_t < z_t \leqslant 0\}$ converges to a null set and we have

$$E\left(\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} < z_t \leqslant 0\right)\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} \leqslant 0\right)\right) \to 0.$$

A similar argument also shows that

$$E\left(\mathbf{1}\left(0 < z_t \leqslant -\delta_t\frac{y_{t-1}}{\sigma_t}\right)\mathbf{1}\left(-\delta_t\frac{y_{t-1}}{\sigma_t} > 0\right)\right) \to 0.$$

Thus we find that $\mathrm{Var}(A + B) \to 0$. This, together with the previous result that $E(A + B) = 0$, implies $A + B = o_p(1)$, using the Markov inequality.

Now we look at term C. Expanding C according to 4 regimes, we have

$$\begin{aligned}
C &= -2T^{-1/2}\sum_{t=2}^{\lfloor rT\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&= -2T^{-1/2}\sum_{t=2}^{\lfloor \tau_1 T\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&\quad -2T^{-1/2}\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor \tau_2 T\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&\quad -2T^{-1/2}\sum_{t=\lfloor \tau_2 T\rfloor+1}^{\lfloor \tau_3 T\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&\quad -2T^{-1/2}\sum_{t=\lfloor \tau_3 T\rfloor+1}^{\lfloor rT\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&= C1+C2+C3+C4
\end{aligned}$$

where $C1$-$C4$ are defined implicitly. First notice that $\delta_t = 0$ in the two unit root regimes, such that $C1 = C4 = 0$. Next we look at $C2$ and $C3$. For $C2$,

$$\begin{aligned}
C2 &= -2T^{-1/2}\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor \tau_2 T\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&= 2T^{-1/2}\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor \tau_2 T\rfloor} f(0)\delta_t\frac{y_{t-1}}{\sigma_t}\left(1+o_p\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)\right) \\
&= 2c_1T^{-1}f(0)\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor \tau_2 T\rfloor}\frac{T^{-1/2}y_{t-1}}{\sigma_t}(1+o_p(1)).
\end{aligned}$$

From (A.1), when $\tau_1 \leqslant r < \tau_2$, $T^{-1/2}y_{\lfloor rT\rfloor} \Rightarrow V_1(r)$, so

$$C2 \Rightarrow 2f(0)c_1\int_{\tau_1}^{\tau_2}\frac{V_1(h)}{\sigma(h)}dh$$

by the CMT. For $C3$,

$$\begin{aligned}
C3 &= -2T^{-1/2}\sum_{t=\lfloor \tau_2 T\rfloor+1}^{\lfloor \tau_3 T\rfloor}\left(F\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)-F(0)\right) \\
&= 2T^{-1/2}\sum_{t=\lfloor \tau_2 T\rfloor+1}^{\lfloor \tau_3 T\rfloor} f(0)\delta_t\frac{y_{t-1}}{\sigma_t}\left(1+o_p\left(-\delta_t\frac{y_{t-1}}{\sigma_t}\right)\right) \\
&= -2c_2T^{-1}f(0)\sum_{t=\lfloor \tau_2 T\rfloor+1}^{\lfloor \tau_3 T\rfloor}\frac{T^{-1/2}y_{t-1}}{\sigma_t}(1+o_p(1)).
\end{aligned}$$

Also from (A.1), when $\tau_1 \leqslant r < \tau_2$, $T^{-1/2} y_{\lfloor rT \rfloor} \Rightarrow V_2(r)$, so we have

$$C3 \Rightarrow -2f(0)c_2 \int_{\tau_2}^{\tau_3} \frac{V_2(h)}{\sigma(h)} dh$$

also by the CMT. Together then,

$$C \Rightarrow 2f(0)c_1 \int_{\tau_1}^{\tau_2} \frac{V_1(h)}{\sigma(h)} dh - 2f(0)c_2 \int_{\tau_2}^{\tau_3} \frac{V_2(h)}{\sigma(h)} dh.$$

Lastly we look at the convergence in distribution of the sum of the signs. We have

$$T^{-1/2} \sum_{t=2}^{\lfloor rT \rfloor} \text{sign}(\varepsilon_t) \Rightarrow W^s(r)$$

as a consequence of the imposed assumptions.

We have thus derived the weak limit of $T^{-1/2} C_{\lfloor rT \rfloor}$ when $r > \tau_3$. Its weak limit in other regimes can be derived in the same manner.

Now we show that $\hat{s}^2(\lambda_1, \lambda_2) \xrightarrow{p} 1$. By definition of the variance estimator $\hat{s}^2(\lambda_1, \lambda_2)$ and the least squares estimator $\hat{\rho}(\lambda_1, \lambda_2)$, we make the following expansion

$$\begin{aligned} \hat{s}^2(\lambda_1, \lambda_2) &= (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 2)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} (\Delta C_t - \hat{\rho}(\lambda_1, \lambda_2) C_{t-1})^2 \\ &= (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 2)^{-1} \left(\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} (\Delta C_t)^2 - \frac{\left(\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} \Delta C_t C_{t-1} \right)^2}{\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} C_{t-1}^2} \right). \end{aligned}$$

Using (A.2) we have $\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} \Delta C_t C_{t-1} = O_p(T)$ and $\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} C_{t-1}^2 = O_p(T^2)$; we thus have

$$(\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 2)^{-1} \frac{\left(\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} \Delta C_t C_{t-1} \right)^2}{\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} C_{t-1}^2} = o_p(1).$$

Also notice that

$$\sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} (\Delta C_t)^2 = \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} (\text{sign}(\Delta u_t))^2$$

where by definition that $\text{sign}(x) = -2\mathbf{1}(x \leqslant 0) + 1$, $(\Delta C_t)^2 = (\text{sign}(\Delta u_t))^2 = 1$ identically. Hence

$$\hat{s}^2(\lambda_1, \lambda_2) = (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 2)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} (\Delta C_t)^2 + o_p(1) \xrightarrow{p} 1.$$

We are now in a position to derive the main result of the theorem, using similar arguments to those in the proof of Theorem 1 (see also PSY, Shi et al., 2018a, and Shi et al., 2018b). We can write $sDF(\lambda_1, \lambda_2)$ in the form

$$\begin{aligned} sDF(\lambda_1,\lambda_2) &= \frac{T^{-1}C^2_{\lfloor\lambda_2 T\rfloor} - T^{-1}C^2_{\lfloor\lambda_1 T\rfloor} - T^{-1}\sum_{t=\lfloor\lambda_1 T\rfloor+1}^{\lfloor\lambda_2 T\rfloor}(\Delta C_t)^2}{2\sqrt{\hat{s}^2(\lambda_1,\lambda_2)T^{-2}\sum_{t=\lfloor\lambda_1 T\rfloor+1}^{\lfloor\lambda_2 T\rfloor}C^2_{t-1}}} \\ &\Rightarrow \frac{U^s(\lambda_2)^2 - U^s(\lambda_1)^2 - (\lambda_2-\lambda_1)}{2\sqrt{\int_{\lambda_1}^{\lambda_2}U^s(r)^2dr}} \\ &= L^s_{c_1,c_2}(\lambda_1,\lambda_2) \end{aligned}$$

where the weak convergence follows from (A.2) and the CMT. The limit for $sPSY$ then follows from a further application of the CMT and the proof of the theorem is complete.

Proof of Theorem 4

Define $D(\lambda_1, \lambda_2) = T^{-1/2}sDF^*(\lambda_1, \lambda_2)$. Clearly $\arg\max_{\lambda_1\in[0,1-\pi]\ \lambda_2\in[\lambda_1+\pi,1]} D(\lambda_1, \lambda_2) = \arg\max_{\lambda_1\in[0,1-\pi]\ \lambda_2\in[\lambda_1+\pi,1]} sDF^*(\lambda_1, \lambda_2)$ for any T. It will be shown that $D(\lambda_1, \lambda_2)$ has a non-explosive limit and its maximiser is (τ_1, τ_2).

Lemma 5 shows that

$$D(\lambda_1,\lambda_2) \xrightarrow{p} \mathbb{D}(\lambda_1,\lambda_2)$$

uniformly in $0 < \lambda_1 < \lambda_2 < 1$, where $\mathbb{D}(\lambda_1, \lambda_2)$ is defined in the lemma. It is also straightforward to verify that

$$(\tau_1,\tau_2) = \underset{\lambda_1\in[0,1-\pi]\ \lambda_2\in[\lambda_1+\pi,1]}{\arg\max} \mathbb{D}(\lambda_1,\lambda_2)$$

which is the unique maximiser of the $\mathbb{D}(\lambda_1, \lambda_2)$ function in the considered domain. That is, the coordinate (τ_1, τ_2) defined by the true bubble start and end time consists of a unique maximiser of the function $\mathbb{D}(\lambda_1, \lambda_2)$ in the domain $[0, 1-\pi] \times [\lambda_1 + \pi, 1]$. Applying the Argmax Theorem (e.g. Theorem 5.7 of Van der Vaart (1998)), we thus have the result that the maximiser $(\hat{\tau}_1, \hat{\tau}_2)$ of the dating statistic $D(\lambda_1, \lambda_2)$ converges in probability to the maximiser (τ_1, τ_2) of its probability limit $\mathbb{D}(\lambda_1, \lambda_2)$, i.e.

$$(\hat{\tau}_1,\hat{\tau}_2) \xrightarrow{p} (\tau_1,\tau_2).$$

Appendix B: Lemmas

Lemmas in this section are for the proof of Theorem 4 relating to consistency of the sign-based dating strategy, hence the model we consider here is DGP 2 with the bubble magnitude set to be $\delta_{1,T} = c_1T^{-\alpha}$, $\alpha \in (0, 1)$. Notice that in DGP 2, there is no mean-reverting regime so essentially $\delta_{2,T} = 0$. In this section, we also assume the bubble is upwards, implying that in large samples the signs of the increments in the explosive regime are predominantly $+1$. However, this assumption is not necessary; in the case of a downwards bubble, the signs of the increments in the explosive regime will be predominantly -1, and the deterministic functions derived in Lemma 2 will be the negative of the results presented, while the results of the other lemmas and the main results of Theorem 4 are unchanged.

Lemma 1

(a) When $r \leqslant \tau_1$,

$$T^{-1/2} u_{\lfloor rT \rfloor} \Rightarrow \int_0^r \sigma(h) dW(h).$$

(b) When $\tau_1 < r \leqslant \tau_2$,

$$T^{-1/2}(1+\delta_{1,T})^{-(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor)} u_{\lfloor rT \rfloor} \Rightarrow \int_0^{\tau_1} \sigma(h) dW(h).$$

(c) When $r > \tau_2$,

$$T^{-1/2}(1+\delta_{1,T})^{-(\lfloor \tau_2 T \rfloor - \lfloor \tau_1 T \rfloor)} u_{\lfloor rT \rfloor} \Rightarrow \int_0^{\tau_1} \sigma(h) dW(h).$$

Proof of Lemma 1

Part (a) follows directly from the functional central limit theorem.

For (b), by repeated backward substitution for $u_{\lfloor rT \rfloor}$ up to $u_{\lfloor \tau_1 T \rfloor}$, we have

$$\begin{aligned} & T^{-1/2}(1+\delta_{1,T})^{-(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor)} u_{\lfloor rT \rfloor} \\ = \; & T^{-1/2} u_{\lfloor \tau_1 T \rfloor} + T^{-1/2}(1+\delta_{1,T})^{-(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor)}(\varepsilon_{\lfloor rT \rfloor} + \ldots + (1+\delta_{1,T})^{\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor - 1} \varepsilon_{\lfloor \tau_1 T \rfloor + 1}). \end{aligned}$$

Denote $A := \varepsilon_{\lfloor rT \rfloor} + \ldots + (1+\delta_{1,T})^{\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor - 1} \varepsilon_{\lfloor \tau_1 T \rfloor + 1} = \sum_{i=\lfloor \tau_1 T \rfloor + 1}^{\lfloor rT \rfloor} (1+\delta_{1,T})^{\lfloor rT \rfloor - i} \varepsilon_i$. Notice that $E(A) = 0$ and

$$\begin{aligned} \mathrm{Var}(A) &= \sum_{i=\lfloor \tau_1 T \rfloor + 1}^{\lfloor rT \rfloor} (1+\delta_{1,T})^{2(\lfloor rT \rfloor - i)} \sigma_i^2 \\ &\leqslant C \sum_{i=\lfloor \tau_1 T \rfloor + 1}^{\lfloor rT \rfloor} (1+\delta_{1,T})^{2(\lfloor rT \rfloor - i)} \\ &= C \frac{(1+\delta_{1,T})^{2(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor - 1)} - 1}{(1+\delta_{1,T})^2 - 1} \\ &= C \frac{(1+\delta_{1,T})^{2(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor - 1)}}{\delta_{1,T}} (1 + o(1)) \\ &= C T^{\alpha} (1+\delta_{1,T})^{2(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor - 1)} (1 + o(1)) \end{aligned}$$

where we have used the uniform boundedness of the variance function. Applying Markov's inequality we have $A = O_p(T^{\alpha/2}(1+\delta_{1,T})^{(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor - 1)})$. Thus

$$T^{-1/2}(1+\delta_{1,T})^{-(\lfloor rT \rfloor - \lfloor \tau_1 T \rfloor)} A = O_p(T^{(\alpha-1)/2}) = o_p(1).$$

Also, notice that $T^{-1/2}u_{\lfloor \tau_1 T\rfloor} \Rightarrow \int_0^{\tau_1} \sigma(h)dW(h)$ as τ_1 is in the unit root regime. In total, we have when $\tau_1 < r \leqslant \tau_2$, $T^{-1/2}(1+\delta_{1,T})^{-(\lfloor rT\rfloor - \lfloor \tau_1 T\rfloor)}u_{\lfloor rT\rfloor} \Rightarrow \int_0^{\tau_1} \sigma(h)dW(h)$.

For (c), by repeated backward substitution for $u_{\lfloor rT\rfloor}$ up to $u_{\lfloor \tau_2 T\rfloor}$, we have

$$u_{\lfloor rT\rfloor} = u_{\lfloor \tau_2 T\rfloor} + (\varepsilon_{\lfloor rT\rfloor} + \varepsilon_{\lfloor rT\rfloor -1} + \ldots + \varepsilon_{\lfloor \tau_2 T\rfloor +1}).$$

Note that for the first term, τ_2 is in the mildly explosive regime so

$$T^{-1/2}(1+\delta_{1,T})^{-(\lfloor \tau_2 T\rfloor - \lfloor \tau_1 T\rfloor)}u_{\lfloor \tau_2 T\rfloor} \Rightarrow \int_0^{\tau_1} \sigma(h)dW(h).$$

Applying the functional central limit theorem for the second term, we have

$$T^{-1/2}(\varepsilon_{\lfloor rT\rfloor} + \varepsilon_{\lfloor rT\rfloor -1} + \ldots + \varepsilon_{\lfloor \tau_2 T\rfloor +1}) \Rightarrow \int_{\tau_2}^{1} \sigma(h)dW(h).$$

Clearly, the first term dominates and we obtain

$$T^{-1/2}(1+\delta_{1,T})^{-(\lfloor rT\rfloor - \lfloor \tau_1 T\rfloor)}u_{\lfloor rT\rfloor} \Rightarrow \int_0^{\tau_1} \sigma(h)dW(h).$$

Lemma 2

(a) When $r \leqslant \tau_1$,

$$T^{-1/2}C_{\lfloor \tau T\rfloor} \Rightarrow W^s(r).$$

(b) When $\tau_1 < r \leqslant \tau_2$, uniformly in r,

$$T^{-1}C_{\lfloor rT\rfloor} - (r - \tau_1) = o_p(1).$$

(c) When $r > \tau_2$, uniformly in r,

$$T^{-1}C_{\lfloor rT\rfloor} - (\tau_2 - \tau_1) = o_p(1).$$

Proof of Lemma 2

Part (a) relates to a unit root regime, and the claimed weak convergence is known from the proof of Theorem 2.

For (b), by definition

$$C_{\lfloor rT\rfloor} = \sum_{i=2}^{\lfloor \tau_1 T\rfloor} \operatorname{sign}(\varepsilon_i) + \sum_{i=\lfloor \tau_1 T\rfloor +1}^{\lfloor rT\rfloor} \operatorname{sign}\left(c_1 T^{-\alpha} u_{i-1} + \varepsilon_i\right).$$

The first term satisfies $T^{-1/2}\sum_{i=2}^{\lfloor \tau_1 T\rfloor}\operatorname{sign}(\varepsilon_i) \Rightarrow W^s(\tau_1)$, while for the second term, notice that

$$\begin{aligned}
&\frac{1}{T}\sum_{i=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\operatorname{sign}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\right)\\
&= \frac{1}{T}\sum_{i=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\left(1-2\times\mathbf{1}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\leqslant 0\right)\right)\\
&= (r-\tau_1)-2\frac{1}{T}\sum_{i=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\mathbf{1}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\leqslant 0\right)\\
&= (r-\tau_1)-2\frac{1}{T}\left(\sum_{i=\lfloor \tau_1 T\rfloor+1}^{\lfloor r^*T\rfloor}\mathbf{1}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\leqslant 0\right)+\sum_{i=\lfloor r^* T\rfloor+1}^{\lfloor rT\rfloor}\mathbf{1}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\leqslant 0\right)\right)
\end{aligned}$$

where $r^*=\tau_1+a^*T^{-\kappa}$, and $a^*,\kappa>0$ are constants. Here we decompose the sum of indicators into two parts. The first part is a sum in a shrinking neighbourhood after time $\lfloor \tau_1 T\rfloor$ (with length of order $T^{-\kappa}$). The first part will have at most $\lfloor r^*T\rfloor-\lfloor \tau_1 T\rfloor = a^*T^{1-\kappa}$ terms of 1s, thus will be at most $a^*T^{1-\kappa}$, which is clearly of a smaller order than T, uniformly in r. The remainder of the decomposition is the sum of all the terms after time $\lfloor r^*T\rfloor$. The same strategy used in proving Lemma 1(b) implies that for any $i\geqslant \lfloor r^*T\rfloor+1$, $T^{-\alpha}u_{i-1}=T^{1/2-\alpha}(1+\delta_{1,T})^{a^*T^{1-\kappa}}$. Choosing $0<\kappa<1$, it follows easily that $T^{-\alpha}u_{i-1}\to+\infty$ as $T\to\infty$ (given an upwards bubble without loss of generality, as discussed in the beginning of Appendix B). Thus the second sum will be identically zero in the limit, uniformly in r. Hence, uniformly in r,

$$\frac{1}{T}\sum_{i=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\operatorname{sign}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\right)-(r-\tau_1)=o_p(1)$$

and the result follows trivially.

The result for (c) follows in the same way as for (b), since here

$$T^{-1}C_{\lfloor rT\rfloor}=T^{-1}\sum_{i=2}^{\lfloor \tau_1 T\rfloor}\operatorname{sign}(\varepsilon_i)+T^{-1}\sum_{i=\lfloor \tau_1 T\rfloor+1}^{\lfloor \tau_2 T\rfloor}\operatorname{sign}\left(c_1T^{-\alpha}u_{i-1}+\varepsilon_i\right)+T^{-1}\sum_{i=\lfloor \tau_2 T\rfloor+1}^{\lfloor rT\rfloor}\operatorname{sign}(\varepsilon_i)$$

and notice that $T^{-1/2}\sum_{i=\lfloor \tau_2 T\rfloor+1}^{\lfloor rT\rfloor}\operatorname{sign}(\varepsilon_i)\Rightarrow W^s(r)-W^s(\tau_2)$, thus the final term is asymptotically negligible, uniformly in r.

Lemma 3

(a) When $r\leqslant\tau_1$,

$$\begin{aligned}
\frac{1}{T^2}\sum_{t=2}^{\lfloor rT\rfloor}C_{t-1}^2 &\Rightarrow \int_0^r (W^s(r))^2dr,\\
\frac{1}{T}\sum_{t=2}^{\lfloor rT\rfloor}\Delta C_tC_{t-1} &\Rightarrow \int_0^r W^s(r)dW^s(r).
\end{aligned}$$

(b) When $\tau_1 < r \leqslant \tau_2$, uniformly in r,

$$\begin{aligned} T^{-3}\sum_{t=2}^{\lfloor rT\rfloor} C_{t-1}^2 &\xrightarrow{p} \quad (r-\tau_1)^3/3, \\ T^{-2}\sum_{t=2}^{\lfloor rT\rfloor} \Delta C_t C_{t-1} &\xrightarrow{p} \quad (r-\tau_1)^2/2. \end{aligned}$$

(c) When $r > \tau_2$,uniformly in r,

$$\begin{aligned} T^{-3}\sum_{t=2}^{\lfloor rT\rfloor} C_{t-1}^2 &\xrightarrow{p} (\tau_2-\tau_1)^3/3 + (\tau_2-\tau_1)^2(r-\tau_2), \\ T^{-2}\sum_{t=2}^{\lfloor rT\rfloor} \Delta C_t C_{t-1} &\xrightarrow{p} (\tau_2-\tau_1)^2/2. \end{aligned}$$

Proof of Lemma 3

Part (a) relates to a unit root regime, and the results are known from the proof of Theorem 2.

For the first claimed result in (b), first notice that

$$T^{-3}\sum_{t=2}^{\lfloor rT\rfloor} C_{t-1}^2 = T^{-3}\sum_{t=2}^{\lfloor \tau_1 T\rfloor} C_{t-1}^2 + T^{-3}\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor} C_{t-1}^2.$$

From the result of part (a), we have $T^{-2}\sum_{t=2}^{\lfloor rT\rfloor} C_{t-1}^2 \Rightarrow \int_0^r (W^s(r))^2 dr$ when $r \leqslant \tau_1$. It thus follows easily $T^{-3}\sum_{t=2}^{\lfloor \tau_1 T\rfloor} C_{t-1}^2 = o_p(1)$. The order also holds uniformly in r, since this term clearly does not depend on r. For the second term, notice that

$$\begin{aligned} &\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-3}\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor} C_{t-1}^2 - (r-\tau_1)^3/3\right| \\ \leqslant\ &\sup_{\tau_1<r\leqslant\tau_2} T^{-3}\left|\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor} C_{t-1}^2 - \sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor}(t-1-\lfloor \tau_1 T\rfloor)^2\right| \\ &+\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-3}\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor}(t-1-\lfloor \tau_1 T\rfloor)^2 - (r-\tau_1)^3/3\right|. \end{aligned} \tag{A.4}$$

Using the result of Lemma 2(b), the first term of (A.4) can be bounded as follows

$$\begin{aligned}
&\sup_{\tau_1<r\leqslant\tau_2} T^{-3}\left|\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor} C_{t-1}^2 - \sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}(t-\lfloor\tau_1 T\rfloor)^2\right| \\
\leqslant\ & \sup_{\tau_1<r\leqslant\tau_2}\left(\sup_{\lfloor\tau_1 T\rfloor+1\leqslant t\leqslant\lfloor rT\rfloor}\left|T^{-2}C_{t-1}^2-\left(\frac{t-1-\lfloor\tau_1 T\rfloor}{T}\right)^2\right|\right)\left(T^{-1}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}1\right) \\
=\ & o_p(1).
\end{aligned}$$

In the second term of (A.4), notice that $T^{-3}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}(t-1-\lfloor\tau_1 T\rfloor)^2$ converges to $(r-\tau_1)^3/3$ as a deterministic sequence, for any $\tau_1<r\leqslant\tau_2$. This pointwise convergence is also uniform in $\tau_1<r\leqslant\tau_2$, by noticing that the limiting function is uniformly continuous. Thus in total we have

$$\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-3}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}C_{t-1}^2-(r-\tau_1)^3/3\right|=o_p(1)$$

which further implies that $\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-3}\sum_{t=2}^{\lfloor rT\rfloor}C_{t-1}^2-(r-\tau_1)^3/3\right|=o_p(1)$.

Next we show the second claimed result of part (b). First,

$$T^{-2}\sum_{t=2}^{\lfloor rT\rfloor}\Delta C_t C_{t-1}=T^{-2}\sum_{t=2}^{\lfloor\tau_1 T\rfloor}\Delta C_t C_{t-1}+T^{-2}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\Delta C_t C_{t-1}.$$

Again using the result of part (a), $T^{-2}\sum_{t=2}^{\lfloor\tau_1 T\rfloor}\Delta C_t C_{t-1}=o_p(1)$ uniformly in r. For the second term,

$$\begin{aligned}
&\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-2}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\Delta C_t C_{t-1}-(\tau_2-\tau_1)^2/2\right| \\
\leqslant\ & \sup_{\tau_1<r\leqslant\tau_2}T^{-2}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}\left|\Delta C_t C_{t-1}-(t-1-\lfloor\tau_1 T\rfloor)\right| \\
&+\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-2}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}(t-1-\lfloor\tau_1 T\rfloor)-(\tau_2-\tau_1)^2/2\right| \\
=\ & \sup_{\tau_1<r\leqslant\tau_2}T^{-2}\left(\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor r^* T\rfloor}+\sum_{t=\lfloor r^* T\rfloor+1}^{\lfloor rT\rfloor}\right)\left|\Delta C_t C_{t-1}-(t-1-\lfloor\tau_1 T\rfloor)\right| \\
&+\sup_{\tau_1<r\leqslant\tau_2}\left|T^{-2}\sum_{t=\lfloor\tau_1 T\rfloor+1}^{\lfloor rT\rfloor}(t-1-\lfloor\tau_1 T\rfloor)-(\tau_2-\tau_1)^2/2\right|
\end{aligned}$$

where $r^* = \tau_1 + a^* T^{-\kappa}$ is chosen as in the proof of Lemma 2. In the decomposed summation in the first term above, first consider the $\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor r^* T\rfloor}$ part. As argued before, when $t \leqslant \lfloor r^* T\rfloor$, ΔC_t is not always 1, even when T is large. But since the sum of the terms has an order of at most $T^{1-\kappa}$, the contribution of the $\sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor r^* T\rfloor}$ part (when normalised by T^{-2}) is $o_p(1)$, uniformly in r. For the $\sum_{i=\lfloor r^* T\rfloor+1}^{\lfloor rT\rfloor}$ part, notice that when $t > \lfloor r^* T\rfloor$, ΔC_t is identically 1 when T is large, and this part (when normalised) can be bounded uniformly in r using Lemma 2(b), as follows

$$
\begin{aligned}
& \sup_{\tau_1 < r \leqslant \tau_2} T^{-2} \sum_{t=\lfloor r^* T\rfloor+1}^{\lfloor rT\rfloor} \left|\Delta C_t C_{t-1} - (t-1-\lfloor \tau_1 T\rfloor)\right| \\
\leqslant\ & \sup_{\tau_1 < r \leqslant \tau_2} \left(\sup_{\lfloor \tau_1 T\rfloor+1 \leqslant t \leqslant \lfloor rT\rfloor} \left| T^{-1} C_{t-1} - \frac{(t-1-\lfloor \tau_1 T\rfloor)}{T} \right| \right) \left(T^{-1} \sum_{t=\lfloor r^* T\rfloor+1}^{\lfloor rT\rfloor} 1 \right) \\
=\ & o_p(1).
\end{aligned}
$$

Then using the same argument as for the first claim of part (b),

$$
\left| T^{-2} \sum_{t=\lfloor \tau_1 T\rfloor+1}^{\lfloor rT\rfloor} (t-1-\lfloor \tau_1 T\rfloor) - (\tau_2-\tau_1)^2/2 \right| = o(1)
$$

uniformly in r as a deterministic sequence. We have thus shown the second claimed result of part (b).

For (c),

$$
T^{-3} \sum_{t=2}^{\lfloor rT\rfloor} C_{t-1}^2 = T^{-3} \sum_{t=2}^{\lfloor \tau_2 T\rfloor} C_{t-1}^2 + T^{-3} \sum_{t=\lfloor \tau_2 T\rfloor+1}^{\lfloor rT\rfloor} C_{t-1}^2.
$$

The first term above has a limit $(\tau_2-\tau_1)^3/3$, uniformly in r using the result of part (b). Using the same argument in proving (b), and the result of Lemma 2(c), the second term above can be shown to have a limit $(\tau_2-\tau_1)^2(r-\tau_2)$, uniformly in r. In total, we thus have

$$
\sup_{r > \tau_2} \left| T^{-3} \sum_{t=2}^{\lfloor rT\rfloor} C_{t-1}^2 - (\tau_2-\tau_1)^3/3 + (\tau_2-\tau_1)^2(r-\tau_2) \right| = o_p(1).
$$

Finally,

$$
\begin{aligned}
& T^{-2} \sum_{t=2}^{\lfloor rT\rfloor} \Delta C_t C_{t-1} \\
=\ & T^{-2} \sum_{t=2}^{\lfloor \tau_2 T\rfloor} \Delta C_t C_{t-1} + T^{-2} \sum_{t=\lfloor \tau_2 T\rfloor+1}^{\lfloor rT\rfloor} \Delta C_t C_{t-1}
\end{aligned}
$$

where the first term has a limit $(\tau_2-\tau_1)^2/2$ uniformly in r by the result of part (b). For the second term, notice that

$$\begin{aligned}
& \sup_{r>\tau_2}\left|T^{-2}\sum_{t=\lfloor\tau_2 T\rfloor+1}^{\lfloor rT\rfloor}\Delta C_t C_{t-1}-T^{-1}(\tau_2-\tau_1)\sum_{t=\lfloor\tau_2 T\rfloor+1}^{\lfloor rT\rfloor}\Delta C_t\right| \\
\leqslant\ & \sup_{r>\tau_2} T^{-1}\sum_{t=\lfloor\tau_2 T\rfloor+1}^{\lfloor rT\rfloor}|\Delta C_t|\left|T^{-1}C_{t-1}-(\tau_2-\tau_1)\right| \\
\leqslant\ & \sup_{r>\tau_2}\left(\sup_{\lfloor\tau_2 T\rfloor+1\leqslant t\leqslant\lfloor rT\rfloor}\left|T^{-1}C_{t-1}-(\tau_2-\tau_1)\right|\right)\left(T^{-1}\sum_{t=\lfloor\tau_2 T\rfloor+1}^{\lfloor rT\rfloor}|\Delta C_t|\right) \\
=\ & o_p(1)
\end{aligned}$$

where we have used the results of Lemma 2(c) and the fact that $T^{-1}\sum_{t=\lfloor\tau_2 T\rfloor+1}^{\lfloor rT\rfloor}|\Delta C_t| = O_p(1)$, uniformly in r. In total, we have thus shown that

$$\sup_{r>\tau_2}\left|T^{-2}\sum_{t=\lfloor\tau_2 T\rfloor+1}^{\lfloor rT\rfloor}\Delta C_t C_{t-1}-(\tau_2-\tau_1)^2/2\right| = o_p(1).$$

Lemma 4

For the sub-sample variance estimator

$$\hat{s}^2(0,r)=(\lfloor rT\rfloor-1)^{-1}\sum_{t=2}^{\lfloor rT\rfloor}(\Delta C_t-\hat{\rho}(0,r)C_{t-1})^2,$$

(a) When $r\leqslant\tau_1$, uniformly in r, $\hat{s}^2(0,r)\xrightarrow{p}1$.
(b) When $\tau_1<r\leqslant\tau_2$, uniformly in r,

$$\hat{s}^2(0,r)\xrightarrow{p}1-\frac{3(r-\tau_1)}{4r}.$$

(c) When $r>\tau_2$,uniformly in r,

$$\hat{s}^2(0,r)\xrightarrow{p}1-\frac{(\tau_2-\tau_1)^2}{4(\tau_2-\tau_1)r/3+4(r-\tau_2)r}.$$

Proof of Lemma 4

As in the proof of Theorem 2, we have the following expansion:

$$\hat{s}^2(0,r)=(\lfloor rT\rfloor-1)^{-1}\sum_{t=2}^{\lfloor rT\rfloor}(\Delta C_t)^2-(\lfloor rT\rfloor-1)^{-1}\frac{\left(\sum_{t=2}^{\lfloor rT\rfloor}\Delta C_t C_{t-1}\right)^2}{\sum_{t=2}^{\lfloor rT\rfloor}C_{t-1}^2}. \qquad (A.5)$$

First notice that in the first term of (A.5), uniformly for all $0 < r < 1$,

$$(\Delta C_t)^2 = (\text{sign}(\Delta u_t))^2.$$

By our definition of the sign function, $(\Delta C_t)^2 = 1$ identically and

$$(\lfloor rT \rfloor - 1)^{-1} \sum_{t=2}^{\lfloor rT \rfloor} (\Delta C_t)^2 \xrightarrow{p} 1.$$

So the difference in parts (a), (b) and (c) only lies in the second term of (A.5).

For part (a), from the proof of Theorem 2 it is clear that

$$(\lfloor rT \rfloor - 1)^{-1} \frac{\left(\sum_{t=2}^{\lfloor rT \rfloor} \Delta C_t C_{t-1} \right)^2}{\sum_{t=2}^{\lfloor rT \rfloor} C_{t-1}^2} = o_p(1)$$

uniformly in r, using the weak convergence results.

For (b), first consider generic functions $f_T(x)$, $f(x)$, $g_T(x)$, $g(x)$ satisfying $\sup_x |f_T(x) - f(x)| = o_p(1)$ and $\sup_x |g_T(x) - g(x)| = o_p(1)$. For functions defined as $h_T(x) := f_T(x)g_T(x)$ and $h(x) := f(x)g(x)$,

$$\begin{aligned}
\sup_x |h_T(x) - h(x)| &= \sup_x |f_T(x)g_T(x) - f(x)g(x)| \\
&\leqslant \sup_x |f_T(x)g_T(x) - f_T(x)g(x) + f_T(x)g(x) - f(x)g(x)| \\
&\leqslant \sup_x |f_T(x)|\,|g_T(x) - g(x)| + \sup_x |g(x)|\,|f_T(x) - f(x)| \\
&\leqslant \sup_x |f_T(x)| \sup_x |g_T(x) - g(x)| + \sup_x |g(x)| \sup_x |f_T(x) - f(x)|.
\end{aligned}$$

As long as $\sup_x |f_T(x)|$ and $\sup_x |g(x)|$ are bounded (in probability), we have $\sup_x |h_T(x) - h(x)| = o_p(1)$. A similar argument could be applied to functions defined as the division of uniformly converging functions (provided that the involved operations are well defined). In our context, since all the convergence results we consider are over a bounded interval of r, the boundedness requirement is easily satisfied. Now, applying the uniform convergence result of Lemma 3(b), and the obvious uniform convergence of $T(\lfloor rT \rfloor - 1)^{-1}$ to $1/r$ over the considered domain of r, we have

$$\begin{aligned}
(\lfloor rT \rfloor - 1)^{-1} \frac{\left(\sum_{t=2}^{\lfloor rT \rfloor} \Delta C_t C_{t-1} \right)^2}{\sum_{t=2}^{\lfloor rT \rfloor} C_{t-1}^2} &= T(\lfloor rT \rfloor - 1)^{-1} \frac{\left(T^{-2} \sum_{t=2}^{\lfloor rT \rfloor} \Delta C_t C_{t-1} \right)^2}{T^{-3} \sum_{t=2}^{\lfloor rT \rfloor} C_{t-1}^2} \\
&\xrightarrow{p} \frac{3(r - \tau_1)}{4r}
\end{aligned}$$

uniformly in r.

For (c), using the results of Lemma 3(c) and the argument in proving part (b), we have

$$\begin{aligned}
(\lfloor rT \rfloor - 1)^{-1} \frac{\left(\sum_{t=2}^{\lfloor rT \rfloor} \Delta C_t C_{t-1} \right)^2}{\sum_{t=2}^{\lfloor rT \rfloor} C_{t-1}^2} &= T(\lfloor rT \rfloor - 1)^{-1} \frac{\left(T^{-2} \sum_{t=2}^{\lfloor rT \rfloor} \Delta C_t C_{t-1} \right)^2}{T^{-3} \sum_{t=2}^{\lfloor rT \rfloor} C_{t-1}^2} \\
&\xrightarrow{p} \frac{(\tau_2 - \tau_1)^2}{4(\tau_2 - \tau_1)r/3 + 4(r - \tau_2)r}
\end{aligned}$$

uniformly in r.

The claimed results of the lemma then follow easily.

Lemma 5

Under the mildly explosive alternative, uniformly for $0 < r < s < 1$,

$$D(r,s) \xrightarrow{p} \mathbb{D}(r,s)$$

where

$$\mathbb{D}(r,s) = \frac{\mathbb{D}_1(r,s)}{\sqrt{(\mathbb{D}_3(r,s))^{\varepsilon}\mathbb{D}_2(r,s)}}$$

with

$$\mathbb{D}_1(r,s) := \begin{cases} 0 & \pi < r < s \leqslant \tau_1 \\ (s-\tau_1)^2/2 & \pi < r \leqslant \tau_1 < s \leqslant \tau_2 \\ (\tau_2-\tau_1)^2/2 & \pi < r \leqslant \tau_1 < \tau_2 < s < 1 \\ ((s-\tau_1)^2-(r-\tau_1)^2)/2 & \tau_1 < r < s \leqslant \tau_2 \\ ((\tau_2-\tau_1)^2-(r-\tau_1)^2)/2 & \tau_1 < r \leqslant \tau_2 < s < 1 \\ 0 & \tau_2 < r < s < 1 \end{cases},$$

$$\mathbb{D}_2(r,s) := \begin{cases} 0 & \pi < r < s \leqslant \tau_1 \\ (s-\tau_1)^3/3 & \pi < r \leqslant \tau_1 < s \leqslant \tau_2 \\ (\tau_2-\tau_1)^3/3+(\tau_2-\tau_1)^2(s-\tau_2) & \pi < r \leqslant \tau_1 < \tau_2 < s < 1 \\ ((s-\tau_1)^3-(r-\tau_1)^3)/3 & \tau_1 < r < s \leqslant \tau_2 \\ (\tau_2-\tau_1)^3/3+(\tau_2-\tau_1)^2(s-\tau_2)-(r-\tau_1)^3/3 & \tau_1 < r \leqslant \tau_2 < s < 1 \\ (\tau_2-\tau_1)^2(s-r) & \tau_2 < r < s < 1 \end{cases},$$

$$\mathbb{D}_3(r,s) := \begin{cases} 1 & \pi < r < s \leqslant \tau_1 \\ 1-\frac{3(s-\tau_1)}{4(s-r)} & \pi < r \leqslant \tau_1 < s \leqslant \tau_2 \\ 1-\frac{(\tau_2-\tau_1)^2}{(4(\tau_2-\tau_1)/3+4(s-\tau_2))(s-r)} & \pi < r \leqslant \tau_1 < \tau_2 < s < 1 \\ 1/4 & \tau_1 < r < s \leqslant \tau_2 \\ 1-\frac{(\tau_2-\tau_1)^2}{(4(\tau_2-\tau_1)/3+4(s-\tau_2))(s-r)}+\frac{3(r-\tau_1)}{4(s-r)} & \tau_1 < r \leqslant \tau_2 < s < 1 \\ 1-\frac{(\tau_2-\tau_1)^2}{s-r}\left(\frac{1}{4(\tau_2-\tau_1)/3+4(s-\tau_2)}-\frac{1}{4(\tau_2-\tau_1)/3+4(r-\tau_2)}\right) & \tau_2 < r < s < 1 \end{cases}.$$

Proof of Lemma 5

The proof is straightforward using the results of Lemmas 3 and 4. To avoid repetition, we derive the limit of $D(r,s)$ when $\tau_1 < r \leqslant \tau_2 < s < 1$ as an example; the other cases can be derived following the same strategy.

First consider generic functions $f_T(x)$, $f(x)$ satisfying $\sup_x |f_T(x) - f(x)| = o_p(1)$. Functions defined as $g_T(x,y) := f_T(x) - h_T(y)$ and $g(x,y) := f(x) - h(y)$ also satisfy

$\sup_{x,y} |g_T(x,y) - g(x,y)| = o_p(1)$ by simply noting the inequality

$$\begin{aligned}
\sup_{x,y} |g_T(x,y) - g(x,y)| &= \sup_{x,y} |f_T(x) - h_T(y) - f(x) + h(y)| \\
&\leqslant \sup_{x,y} |f_T(x) - f(x)| + \sup_{x,y} |h_T(y) - h(y)| \\
&= \sup_{x} |f_T(x) - f(x)| + \sup_{y} |h_T(y) - h(y)|.
\end{aligned}$$

That is, the uniform convergence property is preserved when multivariate functions are defined using the difference of uniformly convergent univariate functions evaluated at two points, as above.

Then

$$\begin{aligned}
T^{-2} \sum_{t=\lfloor rT \rfloor+1}^{\lfloor sT \rfloor} \Delta C_t C_{t-1} &= T^{-2} \sum_{t=2}^{\lfloor sT \rfloor} \Delta C_t C_{t-1} - T^{-2} \sum_{t=2}^{\lfloor rT \rfloor} \Delta C_t C_{t-1} \\
&\xrightarrow{p} (\tau_2 - \tau_1)^2/2 - (r - \tau_1)^2/2
\end{aligned}$$

uniformly in $\tau_1 < r \leqslant \tau_2 < s < 1$, where we have used the fact that $\tau_1 < r \leqslant \tau_2$ and $\tau_2 < s < 1$, together with the results of Lemma 3(b) and 3(c). Similarly

$$\begin{aligned}
T^{-3} \sum_{t=\lfloor rT \rfloor+1}^{\lfloor sT \rfloor} C_{t-1}^2 &= T^{-3} \sum_{t=2}^{\lfloor sT \rfloor} C_{t-1}^2 - T^{-3} \sum_{t=2}^{\lfloor rT \rfloor} C_{t-1}^2 \\
&\xrightarrow{p} (\tau_2 - \tau_1)^3/3 + (\tau_2 - \tau_1)^2 (s - \tau_2) - (r - \tau_1)^3/3
\end{aligned}$$

uniformly in $\tau_1 < r \leqslant \tau_2 < s < 1$, where we again use the results of Lemma 3(b) and 3(c).

For the adjusted variance estimator $\tilde{s}^2(r,s)$, by definition

$$\begin{aligned}
\tilde{s}^2(r,s) &= \frac{\lfloor sT \rfloor \hat{s}^2(0,s) - \lfloor rT \rfloor \hat{s}^2(0,r)}{\lfloor sT \rfloor - \lfloor rT \rfloor - 1} \\
&\xrightarrow{p} \frac{s\left(1 - \frac{(\tau_2-\tau_1)^2}{4(\tau_2-\tau_1)s/3+4(s-\tau_2)s}\right) - r\left(1 - \frac{3(r-\tau_1)}{4r}\right)}{s - r} \\
&= 1 - \frac{(\tau_2 - \tau_1)^2}{(4(\tau_2 - \tau_1)/3 + 4(s - \tau_2))(s - r)} + \frac{3(r - \tau_1)}{4(s - r)}
\end{aligned}$$

uniformly in $\tau_1 < r \leqslant \tau_2 < s < 1$, where we have used the results of Lemma 4(b) and 4(c), together with the fact that uniform convergence is preserved under defined function multiplication/division, as discussed in the proof of Lemma 4. Thus when

$\tau_1 < r \leqslant \tau_2 < s < 1$,

$$
\begin{aligned}
D(r,s) &= T^{-1/2} \frac{\sum_{t=\lfloor rT \rfloor+1}^{\lfloor sT \rfloor} \Delta C_t C_{t-1}}{\sqrt{(\tilde{s}^2(r,s))^{\varepsilon} \sum_{t=\lfloor rT \rfloor+1}^{\lfloor sT \rfloor} C_{t-1}^2}} \\
&= \frac{T^{-2} \sum_{t=\lfloor rT \rfloor+1}^{\lfloor sT \rfloor} \Delta C_t C_{t-1}}{\sqrt{(\tilde{s}^2(r,s))^{\varepsilon} T^{-3} \sum_{t=\lfloor rT \rfloor+1}^{\lfloor sT \rfloor} C_{t-1}^2}} \\
&\xrightarrow{p} \frac{\mathbb{D}_1(r,s)}{\sqrt{(\mathbb{D}_3(r,s))^{\varepsilon} \mathbb{D}_2(r,s)}}
\end{aligned}
$$

uniformly in $\tau_1 < r \leqslant \tau_2 < s < 1$, as in the claimed result.

Table 1. Asymptotic and finite sample ξ-level critical values of sign-based tests.

	$sPSY$			$sPWY$		
	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
$T = 100$	4.381	5.578	13.056	2.470	2.859	3.656
$T = 200$	3.469	3.901	4.957	2.405	2.735	3.434
$T = 400$	3.213	3.547	4.231	2.430	2.776	3.408
$T = \infty$	2.933	3.180	3.655	2.410	2.734	3.248
	$\bar{s}PSY$			$\bar{s}PWY$		
	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
$T = 100$	3.898	4.558	9.566	2.478	2.804	3.461
$T = 200$	3.377	3.787	4.880	2.467	2.818	3.508
$T = 400$	3.201	3.551	4.091	2.407	2.734	3.466
$T = \infty$	2.969	3.211	3.624	2.349	2.664	3.112

Table 2. Asymptotic size of nominal 0.05-level tests: $\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_{\sigma 1}) + \sigma_1 \mathbf{1}(\tau_{\sigma 1} < s \leq \tau_{\sigma 2}) + \mathbf{1}(\tau_{\sigma 2} < s \leq 1)$.

$\tau_{\sigma 1}$	σ_1	PSY	PWY
Panel (a). $\tau_{\sigma 2} = 1$			
0.4	1	0.050	0.050
	1/6	0.049	0.030
	1/3	0.043	0.030
	3	0.380	0.447
	6	0.628	0.717
0.8	1	0.050	0.050
	1/6	0.045	0.047
	1/3	0.042	0.047
	3	0.360	0.292
	6	0.643	0.591
Panel (b). $\tau_{\sigma 2} = 0.7$			
0.1	1	0.050	0.050
	1/6	0.658	0.203
	1/3	0.367	0.133
	3	0.245	0.421
	6	0.438	0.621
0.5	1	0.050	0.050
	1/6	0.538	0.045
	1/3	0.302	0.044
	3	0.339	0.344
	6	0.609	0.618

Table 3(a). Asymptotic local powers of nominal 0.05-level tests: DGP 1, $\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq 1)$.

τ_1	σ_1	c_1	PSY	$sPSY$	$uPSY$	$\bar{s}PSY$	$\bar{u}PSY$	PSY_0	PWY	$sPWY$	$uPWY$	$\bar{s}PWY$	$\bar{u}PWY$	PWY_0
0.4	1	2	0.252	0.300	0.293	0.082	0.228	0.351	0.299	0.264	0.310	0.069	0.270	0.350
		4	0.782	0.761	0.777	0.649	0.777	0.804	0.802	0.723	0.794	0.590	0.796	0.789
		6	0.945	0.935	0.945	0.919	0.945	0.949	0.949	0.919	0.947	0.891	0.948	0.939
		8	0.984	0.983	0.982	0.975	0.982	0.984	0.983	0.975	0.983	0.969	0.983	0.981
	1/6	2	0.703	0.769	0.749	0.628	0.690	0.810	0.189	0.737	0.710	0.609	0.565	0.389
		4	0.947	0.948	0.949	0.933	0.947	0.955	0.827	0.939	0.933	0.921	0.910	0.821
		6	0.983	0.983	0.984	0.981	0.984	0.984	0.952	0.980	0.979	0.975	0.972	0.947
		8	0.995	0.994	0.995	0.993	0.995	0.995	0.986	0.993	0.992	0.990	0.989	0.984
	1/3	2	0.513	0.592	0.571	0.364	0.489	0.650	0.212	0.551	0.507	0.322	0.294	0.388
		4	0.900	0.904	0.904	0.863	0.898	0.921	0.822	0.884	0.875	0.843	0.838	0.813
		6	0.973	0.968	0.971	0.964	0.973	0.974	0.957	0.965	0.964	0.964	0.963	0.951
		8	0.990	0.988	0.989	0.986	0.990	0.991	0.983	0.986	0.985	0.984	0.984	0.981
	3	2	0.098	0.181	0.158	0.067	0.089	0.197	0.130	0.154	0.155	0.070	0.109	0.200
		4	0.504	0.597	0.559	0.453	0.498	0.584	0.544	0.536	0.549	0.380	0.521	0.590
		6	0.851	0.870	0.865	0.841	0.855	0.880	0.868	0.839	0.857	0.800	0.860	0.880
		8	0.958	0.964	0.961	0.958	0.959	0.962	0.962	0.952	0.961	0.944	0.962	0.962
	6	2	0.060	0.165	0.134	0.075	0.068	0.093	0.069	0.139	0.126	0.070	0.075	0.091
		4	0.304	0.563	0.520	0.412	0.387	0.446	0.329	0.508	0.478	0.349	0.366	0.449
		6	0.751	0.853	0.842	0.811	0.798	0.800	0.769	0.814	0.806	0.759	0.776	0.800
		8	0.921	0.944	0.940	0.940	0.939	0.932	0.925	0.928	0.928	0.922	0.931	0.934
0.8	1	2	0.088	0.106	0.104	0.053	0.083	0.125	0.094	0.076	0.088	0.050	0.079	0.105
		4	0.299	0.281	0.294	0.153	0.277	0.353	0.283	0.202	0.278	0.087	0.272	0.283
		6	0.534	0.495	0.522	0.367	0.519	0.561	0.518	0.384	0.501	0.241	0.506	0.477
		8	0.713	0.671	0.708	0.581	0.706	0.720	0.683	0.557	0.671	0.455	0.675	0.640
	1/6	2	0.169	0.555	0.520	0.418	0.399	0.650	0.052	0.438	0.407	0.301	0.264	0.077
		4	0.643	0.802	0.791	0.739	0.749	0.845	0.240	0.731	0.722	0.654	0.625	0.265
		6	0.857	0.893	0.894	0.873	0.881	0.915	0.513	0.857	0.849	0.818	0.804	0.490
		8	0.928	0.937	0.940	0.925	0.934	0.958	0.715	0.913	0.907	0.897	0.888	0.673
	1/3	2	0.109	0.323	0.287	0.168	0.161	0.402	0.052	0.236	0.209	0.102	0.086	0.083
		4	0.468	0.641	0.615	0.520	0.545	0.710	0.242	0.519	0.493	0.401	0.373	0.270
		6	0.738	0.791	0.791	0.732	0.769	0.834	0.513	0.723	0.709	0.647	0.628	0.493
		8	0.870	0.879	0.881	0.845	0.872	0.902	0.708	0.827	0.819	0.788	0.777	0.667
	3	2	0.091	0.070	0.080	0.051	0.076	0.112	0.118	0.056	0.099	0.051	0.101	0.124
		4	0.186	0.117	0.163	0.063	0.160	0.228	0.242	0.079	0.202	0.055	0.206	0.251
		6	0.335	0.233	0.312	0.131	0.306	0.372	0.406	0.149	0.355	0.073	0.361	0.397
		8	0.499	0.388	0.474	0.266	0.477	0.523	0.547	0.274	0.522	0.154	0.527	0.538
	6	2	0.067	0.061	0.067	0.051	0.062	0.074	0.078	0.054	0.068	0.051	0.069	0.077
		4	0.105	0.095	0.096	0.065	0.087	0.146	0.146	0.068	0.110	0.056	0.109	0.160
		6	0.194	0.172	0.175	0.102	0.165	0.247	0.242	0.099	0.206	0.067	0.206	0.256
		8	0.319	0.285	0.299	0.188	0.287	0.374	0.367	0.177	0.329	0.111	0.330	0.383

Table 3(b). Asymptotic local powers of nominal 0.05-level tests: DGP 2, $\tau_2 = 0.7$,
$\sigma(s) = \mathbf{1}(0 \le s \le \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \le \tau_2) + \mathbf{1}(\tau_2 < s \le 1)$.

τ_1	σ_1	c_1	PSY	$sPSY$	$uPSY$	$\bar{s}PSY$	$\bar{u}PSY$	PSY_0	PWY	$sPWY$	$uPWY$	$\bar{s}PWY$	$\bar{u}PWY$	PWY_0
0.1	1	2	0.149	0.236	0.215	0.340	0.373	0.264	0.187	0.204	0.231	0.192	0.270	0.313
		4	0.651	0.665	0.674	0.638	0.740	0.710	0.695	0.627	0.685	0.550	0.717	0.719
		6	0.920	0.913	0.924	0.890	0.929	0.928	0.928	0.894	0.925	0.863	0.930	0.923
		8	0.980	0.979	0.982	0.981	0.986	0.980	0.979	0.976	0.980	0.972	0.982	0.978
	1/6	2	0.024	0.635	0.590	0.641	0.609	0.525	0.035	0.603	0.578	0.548	0.520	0.369
		4	0.699	0.921	0.910	0.895	0.886	0.891	0.694	0.896	0.896	0.871	0.863	0.826
		6	0.945	0.981	0.977	0.974	0.973	0.981	0.938	0.975	0.974	0.970	0.969	0.958
		8	0.988	0.994	0.993	0.993	0.992	0.993	0.986	0.991	0.991	0.990	0.990	0.989
	1/3	2	0.045	0.435	0.375	0.504	0.486	0.413	0.054	0.394	0.362	0.359	0.348	0.373
		4	0.673	0.847	0.826	0.810	0.800	0.843	0.718	0.816	0.808	0.773	0.781	0.819
		6	0.922	0.955	0.949	0.950	0.943	0.956	0.939	0.940	0.944	0.935	0.943	0.946
		8	0.985	0.990	0.990	0.987	0.989	0.991	0.986	0.984	0.986	0.982	0.986	0.986
	3	2	0.065	0.197	0.149	0.262	0.251	0.138	0.065	0.169	0.141	0.133	0.135	0.132
		4	0.398	0.584	0.538	0.542	0.555	0.548	0.403	0.539	0.508	0.464	0.491	0.534
		6	0.812	0.869	0.850	0.855	0.855	0.849	0.812	0.833	0.835	0.814	0.833	0.839
		8	0.955	0.967	0.962	0.965	0.967	0.960	0.956	0.955	0.958	0.950	0.960	0.957
	6	2	0.051	0.191	0.131	0.247	0.221	0.073	0.052	0.161	0.131	0.127	0.115	0.073
		4	0.222	0.590	0.527	0.533	0.480	0.438	0.228	0.530	0.496	0.450	0.404	0.437
		6	0.745	0.866	0.847	0.859	0.834	0.806	0.745	0.831	0.818	0.814	0.799	0.803
		8	0.922	0.960	0.955	0.955	0.948	0.940	0.923	0.940	0.937	0.942	0.940	0.939
0.5	1	2	0.086	0.091	0.091	0.098	0.114	0.112	0.098	0.069	0.092	0.060	0.088	0.117
		4	0.247	0.235	0.250	0.205	0.300	0.304	0.272	0.179	0.262	0.098	0.260	0.287
		6	0.492	0.431	0.484	0.408	0.534	0.523	0.509	0.357	0.488	0.225	0.498	0.486
		8	0.674	0.625	0.677	0.596	0.706	0.703	0.693	0.536	0.678	0.443	0.688	0.653
	1/6	2	0.032	0.498	0.445	0.453	0.397	0.535	0.056	0.414	0.379	0.283	0.242	0.102
		4	0.079	0.779	0.749	0.756	0.728	0.820	0.219	0.705	0.688	0.647	0.616	0.313
		6	0.486	0.890	0.875	0.883	0.864	0.915	0.542	0.849	0.841	0.817	0.793	0.528
		8	0.770	0.942	0.930	0.931	0.921	0.954	0.731	0.913	0.906	0.898	0.889	0.705
	1/3	2	0.043	0.255	0.196	0.229	0.198	0.316	0.060	0.196	0.176	0.105	0.095	0.102
		4	0.104	0.573	0.531	0.544	0.498	0.649	0.231	0.493	0.472	0.377	0.356	0.311
		6	0.438	0.768	0.738	0.748	0.723	0.825	0.537	0.692	0.683	0.636	0.624	0.532
		8	0.724	0.876	0.862	0.862	0.842	0.899	0.725	0.827	0.822	0.783	0.784	0.701
	3	2	0.076	0.061	0.080	0.064	0.082	0.089	0.086	0.053	0.081	0.052	0.081	0.093
		4	0.132	0.112	0.131	0.107	0.149	0.171	0.154	0.080	0.135	0.062	0.138	0.173
		6	0.238	0.207	0.234	0.185	0.267	0.294	0.288	0.147	0.252	0.091	0.263	0.306
		8	0.396	0.343	0.396	0.290	0.430	0.446	0.448	0.268	0.407	0.169	0.418	0.458
	6	2	0.059	0.058	0.060	0.060	0.064	0.073	0.069	0.052	0.060	0.051	0.059	0.076
		4	0.084	0.092	0.085	0.086	0.090	0.110	0.103	0.068	0.081	0.058	0.081	0.121
		6	0.131	0.164	0.143	0.141	0.147	0.185	0.162	0.111	0.134	0.078	0.127	0.195
		8	0.224	0.264	0.235	0.236	0.242	0.302	0.267	0.190	0.231	0.125	0.225	0.311

Table 3(c). Asymptotic local powers of nominal 0.05-level tests: DGP 4, $\tau_2 = 0.7$, $\tau_3 = 0.8$, $\sigma(s) = \mathbf{1}(0 \le s \le \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \le \tau_3) + \mathbf{1}(\tau_3 < s \le 1)$.

τ_1	σ_1	c_1	PSY	$sPSY$	$uPSY$	$\bar{s}PSY$	$\bar{u}PSY$	PSY_0	PWY	$sPWY$	$uPWY$	$\bar{s}PWY$	$\bar{u}PWY$	PWY_0
0.1	1	2	0.160	$\underline{0.225}$	0.208	0.090	0.153	0.249	0.176	0.210	0.227	0.055	0.160	0.294
		4	0.659	0.645	0.658	0.461	0.642	0.695	$\underline{0.681}$	0.628	0.675	0.383	0.664	0.695
		6	0.919	0.906	0.915	0.856	0.912	0.923	0.924	0.890	0.921	0.829	0.920	0.916
		8	0.980	0.975	0.980	0.969	0.980	0.980	0.979	0.974	0.979	0.962	0.979	0.978
	1/6	2	0.035	$\underline{0.604}$	0.552	0.309	0.241	0.501	0.021	$\underline{0.598}$	0.563	0.219	0.171	0.302
		4	0.698	$\underline{0.906}$	0.890	0.848	0.829	0.884	0.654	$\underline{0.895}$	0.889	0.830	0.817	0.791
		6	0.940	0.978	0.972	0.966	0.963	0.973	0.927	$\underline{0.972}$	0.971	0.962	0.959	0.944
		8	0.986	0.993	0.991	0.990	0.989	0.992	0.980	0.990	0.990	0.988	0.988	0.984
	1/3	2	0.060	$\underline{0.405}$	0.344	0.165	0.121	0.382	0.028	$\underline{0.395}$	0.346	0.072	0.058	0.323
		4	0.675	$\underline{0.824}$	0.806	0.715	0.699	0.822	0.697	$\underline{0.811}$	0.795	0.683	0.691	0.783
		6	0.926	0.947	0.943	0.934	0.927	0.953	0.932	0.938	0.939	0.923	0.929	0.936
		8	0.983	0.989	0.988	0.984	0.984	0.990	0.983	0.985	0.983	0.978	0.981	0.982
	3	2	0.065	$\underline{0.180}$	0.140	0.074	0.075	0.140	0.066	$\underline{0.171}$	0.143	0.059	0.076	0.133
		4	0.400	$\underline{0.564}$	0.518	0.378	0.414	0.548	0.403	$\underline{0.537}$	0.508	0.311	0.394	0.534
		6	0.814	$\underline{0.856}$	0.839	0.802	0.819	0.849	0.813	0.833	0.835	0.769	0.813	0.839
		8	0.956	0.964	0.960	0.951	0.957	0.960	0.956	0.954	0.957	0.942	0.955	0.957
	6	2	0.051	$\underline{0.175}$	0.117	0.070	0.064	0.073	0.052	$\underline{0.163}$	0.134	0.060	0.064	0.073
		4	0.222	$\underline{0.563}$	0.501	0.373	0.316	0.439	0.229	$\underline{0.531}$	0.498	0.309	0.283	0.437
		6	0.745	$\underline{0.850}$	0.832	0.809	0.781	0.806	0.745	$\underline{0.828}$	0.816	0.768	0.765	0.803
		8	0.922	0.955	0.950	0.943	0.937	0.940	0.923	0.939	0.936	0.930	0.934	0.939
0.5	1	2	0.085	0.089	0.090	0.055	0.082	0.101	0.085	0.072	0.087	0.051	0.075	0.099
		4	0.250	0.229	0.249	0.115	0.236	0.280	$\underline{0.241}$	0.184	0.235	0.074	0.221	0.258
		6	$\underline{0.492}$	0.420	0.476	0.307	0.467	0.503	$\underline{0.482}$	0.358	0.471	0.190	0.469	0.460
		8	$\underline{0.675}$	0.618	0.665	0.526	0.663	0.685	$\underline{0.669}$	0.537	0.657	0.401	0.657	0.625
	1/6	2	0.058	$\underline{0.485}$	0.444	0.342	0.290	0.512	0.050	$\underline{0.417}$	0.380	0.231	0.200	0.069
		4	0.290	$\underline{0.766}$	0.745	0.703	0.677	0.795	0.169	$\underline{0.706}$	0.686	0.617	0.584	0.245
		6	0.574	$\underline{0.882}$	0.872	0.859	0.843	0.898	0.478	$\underline{0.848}$	0.840	0.801	0.780	0.468
		8	0.765	$\underline{0.937}$	0.926	0.916	0.908	0.942	0.683	$\underline{0.912}$	0.906	0.890	0.883	0.665
	1/3	2	0.060	$\underline{0.250}$	0.211	0.122	0.104	0.289	0.051	$\underline{0.200}$	0.175	0.075	0.067	0.070
		4	0.235	$\underline{0.559}$	0.531	0.452	0.416	0.616	0.184	$\underline{0.494}$	0.466	0.328	0.299	0.247
		6	0.518	$\underline{0.755}$	0.735	0.700	0.677	0.803	0.475	$\underline{0.694}$	0.675	0.606	0.586	0.476
		8	0.729	$\underline{0.865}$	0.852	0.833	0.818	0.883	0.682	$\underline{0.826}$	0.814	0.767	0.754	0.662
	3	2	0.076	0.063	0.075	0.053	0.074	0.087	0.080	0.055	0.080	0.049	0.077	0.094
		4	0.132	0.110	0.129	0.070	0.121	0.166	$\underline{0.147}$	0.081	0.134	0.055	0.133	0.173
		6	0.238	0.201	0.228	0.125	0.225	0.290	$\underline{0.277}$	0.147	0.248	0.079	0.248	0.304
		8	$\underline{0.398}$	0.335	0.386	0.223	0.386	0.439	$\underline{0.434}$	0.268	0.402	0.149	0.404	0.454
	6	2	0.059	0.058	0.059	0.053	0.059	0.073	0.070	0.053	0.060	0.048	0.059	0.074
		4	0.085	0.091	0.082	0.066	0.078	0.110	0.103	0.069	0.084	0.053	0.081	0.117
		6	0.132	0.159	0.138	0.099	0.122	0.187	$\underline{0.162}$	0.113	0.136	0.069	0.125	0.191
		8	0.224	0.255	0.230	0.187	0.209	0.301	$\underline{0.268}$	0.193	0.232	0.113	0.222	0.307

Table 4. Finite sample empirical size of nominal 0.05-level tests: $T = 100$,
$\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq \tau_2) + \mathbf{1}(\tau_2 < s \leq 1)$.

τ_1	σ_1	PSY	$sPSY$	$uPSY$	$\bar{s}PSY$	$\bar{u}PSY$	PWY	$sPWY$	$uPWY$	$\bar{s}PWY$	$\bar{u}PWY$
					Panel (a). $\tau_2 = 1$						
0.4	1	0.007	0.059	0.042	0.047	0.030	0.028	0.054	0.039	0.047	0.036
	1/6	0.006	0.059	0.043	0.047	0.029	0.026	0.054	0.035	0.047	0.035
	1/3	0.006	0.059	0.043	0.047	0.030	0.026	0.054	0.035	0.047	0.035
	3	0.026	0.059	0.046	0.047	0.036	0.066	0.054	0.059	0.047	0.060
	6	0.045	0.059	0.054	0.047	0.043	0.073	0.054	0.066	0.047	0.068
0.8	1	0.007	0.059	0.042	0.047	0.030	0.028	0.054	0.039	0.047	0.036
	1/6	0.009	0.059	0.043	0.047	0.029	0.026	0.054	0.036	0.047	0.034
	1/3	0.009	0.059	0.043	0.047	0.029	0.026	0.054	0.036	0.047	0.034
	3	0.036	0.059	0.049	0.047	0.037	0.066	0.054	0.069	0.047	0.070
	6	0.060	0.059	0.057	0.047	0.049	0.078	0.054	0.076	0.047	0.074
					Panel (b). $\tau_2 = 0.7$						
0.1	1	0.007	0.059	0.042	0.047	0.030	0.028	0.054	0.039	0.047	0.036
	1/6	0.051	0.059	0.052	0.047	0.045	0.056	0.054	0.057	0.047	0.057
	1/3	0.030	0.059	0.045	0.047	0.036	0.058	0.054	0.062	0.047	0.059
	3	0.018	0.059	0.045	0.047	0.032	0.047	0.054	0.041	0.047	0.045
	6	0.032	0.059	0.044	0.047	0.035	0.046	0.054	0.042	0.047	0.042
0.5	1	0.007	0.059	0.042	0.047	0.030	0.028	0.054	0.039	0.047	0.036
	1/6	0.032	0.059	0.045	0.047	0.036	0.025	0.054	0.037	0.047	0.035
	1/3	0.021	0.059	0.042	0.047	0.030	0.026	0.054	0.039	0.047	0.035
	3	0.029	0.059	0.045	0.047	0.036	0.056	0.054	0.057	0.047	0.056
	6	0.043	0.059	0.051	0.047	0.038	0.077	0.054	0.067	0.047	0.068

Table 5(a). Finite sample powers of nominal 0.05-level tests: $T = 100$, DGP 1, $\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq 1)$.

τ_1	σ_1	c_1	PSY	$sPSY$	$uPSY$	$\bar{s}PSY$	$\bar{u}PSY$	PWY	$sPWY$	$uPWY$	$\bar{s}PWY$	$\bar{u}PWY$
0.4	1	2	0.055	0.233	0.172	0.137	0.124	0.207	0.233	0.220	0.052	0.178
		4	0.644	0.704	0.683	0.648	0.656	0.766	0.683	0.750	0.406	0.748
		6	0.922	0.918	0.920	0.903	0.916	0.940	0.899	0.936	0.719	0.937
		8	0.981	0.980	0.983	0.977	0.982	0.988	0.964	0.986	0.863	0.986
	1/6	2	0.501	0.724	0.693	0.675	0.665	0.157	0.720	0.672	0.464	0.387
		4	0.910	0.935	0.926	0.924	0.925	0.801	0.922	0.911	0.833	0.855
		6	0.974	0.981	0.979	0.978	0.978	0.943	0.974	0.968	0.909	0.952
		8	0.994	0.994	0.994	0.993	0.994	0.981	0.991	0.988	0.942	0.988
	1/3	2	0.228	0.539	0.476	0.433	0.410	0.170	0.538	0.465	0.183	0.183
		4	0.833	0.880	0.864	0.857	0.857	0.809	0.862	0.846	0.706	0.810
		6	0.964	0.968	0.966	0.964	0.963	0.950	0.960	0.959	0.870	0.950
		8	0.992	0.992	0.992	0.990	0.992	0.988	0.985	0.990	0.920	0.988
	3	2	0.072	0.148	0.111	0.102	0.093	0.171	0.138	0.151	0.059	0.134
		4	0.458	0.532	0.497	0.447	0.463	0.571	0.497	0.544	0.228	0.533
		6	0.831	0.837	0.836	0.804	0.828	0.867	0.791	0.857	0.571	0.857
		8	0.954	0.953	0.954	0.941	0.953	0.960	0.921	0.958	0.766	0.958
	6	2	0.075	0.139	0.110	0.100	0.096	0.132	0.124	0.131	0.063	0.114
		4	0.338	0.488	0.438	0.411	0.404	0.438	0.451	0.453	0.226	0.399
		6	0.754	0.799	0.788	0.758	0.773	0.794	0.755	0.785	0.545	0.781
		8	0.926	0.931	0.930	0.915	0.927	0.936	0.897	0.932	0.728	0.935
0.8	1	2	0.015	0.095	0.069	0.059	0.043	0.060	0.078	0.070	0.052	0.062
		4	0.097	0.227	0.181	0.137	0.132	0.209	0.146	0.196	0.064	0.191
		6	0.327	0.412	0.385	0.294	0.335	0.439	0.263	0.415	0.094	0.417
		8	0.555	0.574	0.566	0.440	0.542	0.623	0.362	0.606	0.124	0.609
	1/6	2	0.074	0.495	0.448	0.365	0.340	0.029	0.337	0.252	0.111	0.068
		4	0.463	0.756	0.728	0.618	0.646	0.183	0.547	0.449	0.200	0.225
		6	0.758	0.858	0.845	0.718	0.818	0.457	0.630	0.587	0.231	0.484
		8	0.875	0.909	0.900	0.759	0.892	0.661	0.675	0.716	0.258	0.673
	1/3	2	0.024	0.261	0.205	0.155	0.129	0.032	0.172	0.121	0.066	0.042
		4	0.232	0.566	0.528	0.436	0.429	0.185	0.389	0.313	0.129	0.184
		6	0.582	0.737	0.711	0.595	0.650	0.459	0.516	0.507	0.185	0.457
		8	0.760	0.829	0.821	0.688	0.797	0.659	0.595	0.680	0.220	0.658
	3	2	0.061	0.071	0.062	0.052	0.052	0.124	0.059	0.107	0.049	0.110
		4	0.121	0.106	0.109	0.070	0.097	0.217	0.079	0.193	0.051	0.195
		6	0.230	0.171	0.197	0.106	0.191	0.351	0.117	0.317	0.059	0.318
		8	0.393	0.270	0.339	0.189	0.336	0.493	0.168	0.467	0.076	0.469
	6	2	0.081	0.066	0.070	0.050	0.060	0.121	0.056	0.107	0.049	0.102
		4	0.122	0.086	0.091	0.061	0.092	0.178	0.067	0.152	0.050	0.158
		6	0.186	0.125	0.150	0.087	0.141	0.263	0.087	0.224	0.053	0.219
		8	0.292	0.189	0.231	0.135	0.222	0.389	0.121	0.333	0.064	0.335

Table 5(b). Finite sample powers of nominal 0.05-level tests: $T = 100$, DGP 2, $\tau_2 = 0.7$,
$\sigma(s) = \mathbf{1}(0 \leq s \leq \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \leq \tau_2) + \mathbf{1}(\tau_2 < s \leq 1)$.

τ_1	σ_1	c_1	PSY	$sPSY$	$uPSY$	$\bar{s}PSY$	$\bar{u}PSY$	PWY	$sPWY$	$uPWY$	$\bar{s}PWY$	$\bar{u}PWY$
0.1	1	2	0.015	$\underline{0.168}$	0.125	0.092	0.076	0.110	$\underline{0.200}$	0.160	0.045	0.084
		4	0.429	$\underline{0.583}$	0.532	0.489	0.495	0.610	0.596	0.607	0.154	0.579
		6	0.831	0.863	0.849	0.827	0.839	0.887	0.851	0.879	0.458	0.878
		8	0.945	0.951	0.951	0.943	0.948	0.967	0.945	0.962	0.657	0.963
	1/6	2	0.153	$\underline{0.540}$	0.483	0.413	0.398	0.051	$\underline{0.576}$	0.520	0.083	0.067
		4	0.788	$\underline{0.875}$	0.856	0.847	0.844	0.717	$\underline{0.869}$	0.857	0.487	0.713
		6	0.951	0.968	0.966	0.960	0.964	0.929	0.962	0.959	0.698	0.928
		8	0.986	0.988	0.987	0.987	0.987	0.978	0.984	0.984	0.767	0.978
	1/3	2	0.069	$\underline{0.322}$	0.268	0.191	0.174	0.070	$\underline{0.365}$	0.313	0.048	0.064
		4	0.678	$\underline{0.768}$	0.749	0.720	0.725	0.722	$\underline{0.780}$	0.765	0.298	0.689
		6	0.909	0.929	0.924	0.917	0.919	0.927	0.924	0.926	0.606	0.918
		8	0.978	0.978	0.979	0.973	0.975	0.980	0.974	0.978	0.735	0.978
	3	2	0.028	$\underline{0.144}$	0.104	0.096	0.083	0.081	$\underline{0.168}$	0.125	0.054	0.065
		4	0.282	$\underline{0.501}$	0.442	0.397	0.391	0.414	$\underline{0.519}$	0.477	0.157	0.367
		6	0.750	$\underline{0.819}$	0.802	0.777	0.786	0.803	0.799	0.805	0.436	0.781
		8	0.923	0.932	0.927	0.916	0.924	0.939	0.918	0.934	0.633	0.933
	6	2	0.037	$\underline{0.139}$	0.102	0.097	0.087	0.057	$\underline{0.162}$	0.114	0.056	0.054
		4	0.211	$\underline{0.489}$	0.425	0.392	0.378	0.283	$\underline{0.506}$	0.433	0.164	0.248
		6	0.710	$\underline{0.804}$	0.787	0.761	0.762	0.745	$\underline{0.786}$	0.778	0.446	0.726
		8	0.909	0.931	0.925	0.916	0.920	0.915	0.919	0.919	0.649	0.909
0.5	1	2	0.011	$\underline{0.089}$	0.060	0.057	0.042	0.059	0.083	0.065	0.045	0.050
		4	0.056	$\underline{0.193}$	0.155	0.115	0.107	0.183	0.168	0.178	0.051	0.155
		6	0.237	$\underline{0.363}$	0.322	0.250	0.273	$\underline{0.411}$	0.302	0.393	0.083	0.378
		8	0.493	$\underline{0.555}$	0.537	0.415	0.495	$\underline{0.626}$	0.441	0.603	0.139	0.603
	1/6	2	0.071	$\underline{0.439}$	0.396	0.316	0.296	0.032	$\underline{0.368}$	0.290	0.102	0.059
		4	0.385	$\underline{0.740}$	0.710	0.619	0.633	0.161	$\underline{0.648}$	0.564	0.230	0.204
		6	0.704	$\underline{0.841}$	0.830	0.743	0.800	0.463	$\underline{0.754}$	0.706	0.290	0.478
		8	0.844	$\underline{0.899}$	0.894	0.806	0.879	0.665	$\underline{0.810}$	0.793	0.320	0.681
	1/3	2	0.032	$\underline{0.218}$	0.178	0.127	0.113	0.033	$\underline{0.191}$	0.134	0.051	0.036
		4	0.178	$\underline{0.526}$	0.483	0.403	0.390	0.160	$\underline{0.439}$	0.358	0.134	0.158
		6	0.512	$\underline{0.718}$	0.697	0.601	0.632	0.462	$\underline{0.627}$	0.579	0.212	0.450
		8	0.724	$\underline{0.813}$	0.797	0.705	0.774	0.654	$\underline{0.710}$	0.710	0.266	0.651
	3	2	0.045	0.071	0.058	0.052	0.045	0.106	0.067	0.087	0.047	0.088
		4	0.089	0.100	0.089	0.071	0.079	$\underline{0.186}$	0.087	0.155	0.051	0.156
		6	0.175	0.163	0.159	0.111	0.144	$\underline{0.304}$	0.136	0.262	0.058	0.260
		8	0.316	0.278	0.288	0.194	0.270	$\underline{0.440}$	0.214	0.392	0.078	0.392
	6	2	0.068	0.067	0.060	0.053	0.051	$\underline{0.105}$	0.065	0.085	0.049	0.086
		4	0.100	0.090	0.081	0.068	0.072	$\underline{0.156}$	0.075	0.125	0.051	0.126
		6	0.149	0.130	0.129	0.101	0.120	$\underline{0.230}$	0.110	0.189	0.061	0.192
		8	0.242	0.206	0.214	0.160	0.208	$\underline{0.335}$	0.158	0.289	0.080	0.290

Table 6. Finite sample empirical size and powers of nominal 0.05-level tests: $T = 100$, DGP 1, $\sigma(s) = \mathbf{1}(0 \le s \le \tau_1) + \sigma_1 \mathbf{1}(\tau_1 < s \le 1)$.

			$\mu = 0$		$\mu = 10$		$\mu = 20$		$\mu = 50$	
τ_1	σ_1	c_1	PSY_0	PWY_0	PSY_0	PWY_0	PSY_0	PWY_0	PSY_0	PWY_0
0.4	1	0	0.101	0.053	0.109	0.044	0.115	0.047	0.128	0.049
		2	0.367	0.311	0.334	0.217	0.294	0.172	0.239	0.161
		4	0.800	0.773	0.781	0.701	0.739	0.581	0.577	0.373
		6	0.945	0.937	0.940	0.904	0.923	0.847	0.836	0.676
		8	0.990	0.985	0.984	0.974	0.977	0.954	0.942	0.858
	1/6	0	0.118	0.055	0.124	0.051	0.135	0.053	0.145	0.054
		2	0.823	0.382	0.685	0.276	0.532	0.227	0.416	0.220
		4	0.955	0.804	0.888	0.707	0.803	0.588	0.610	0.400
		6	0.988	0.936	0.958	0.894	0.923	0.827	0.824	0.639
		8	0.996	0.979	0.983	0.963	0.971	0.934	0.929	0.837
	1/3	0	0.113	0.055	0.123	0.053	0.134	0.055	0.147	0.056
		2	0.676	0.377	0.588	0.267	0.456	0.221	0.351	0.216
		4	0.913	0.810	0.864	0.710	0.786	0.593	0.591	0.396
		6	0.979	0.947	0.958	0.900	0.926	0.836	0.820	0.645
		8	0.996	0.986	0.987	0.967	0.974	0.937	0.929	0.837
	3	0	0.088	0.070	0.081	0.032	0.073	0.021	0.070	0.014
		2	0.231	0.227	0.191	0.148	0.172	0.108	0.141	0.072
		4	0.625	0.613	0.611	0.577	0.596	0.521	0.529	0.361
		6	0.881	0.874	0.877	0.861	0.866	0.835	0.833	0.723
		8	0.964	0.959	0.963	0.956	0.960	0.947	0.949	0.903
	6	0	0.086	0.077	0.044	0.024	0.033	0.009	0.033	0.004
		2	0.180	0.179	0.126	0.106	0.111	0.079	0.091	0.044
		4	0.534	0.521	0.509	0.490	0.494	0.469	0.464	0.383
		6	0.825	0.825	0.823	0.815	0.818	0.802	0.801	0.739
		8	0.947	0.946	0.946	0.942	0.944	0.935	0.933	0.906
0.8	1	0	0.101	0.053	0.109	0.044	0.115	0.047	0.128	0.049
		2	0.174	0.098	0.164	0.073	0.152	0.068	0.152	0.067
		4	0.352	0.244	0.319	0.162	0.267	0.126	0.232	0.115
		6	0.549	0.438	0.508	0.314	0.424	0.214	0.316	0.182
		8	0.708	0.601	0.653	0.471	0.555	0.320	0.384	0.254
	1/6	0	0.102	0.051	0.113	0.044	0.126	0.047	0.132	0.049
		2	0.635	0.083	0.506	0.064	0.367	0.064	0.341	0.069
		4	0.841	0.247	0.666	0.162	0.499	0.134	0.428	0.125
		6	0.916	0.467	0.761	0.336	0.608	0.223	0.469	0.201
		8	0.951	0.640	0.818	0.503	0.673	0.341	0.481	0.273
	1/3	0	0.102	0.051	0.114	0.045	0.125	0.047	0.134	0.049
		2	0.411	0.085	0.355	0.065	0.269	0.067	0.251	0.069
		4	0.695	0.248	0.581	0.163	0.436	0.131	0.367	0.122
		6	0.826	0.465	0.708	0.331	0.564	0.221	0.427	0.194
		8	0.895	0.637	0.792	0.502	0.648	0.335	0.461	0.273
	3	0	0.099	0.081	0.096	0.044	0.086	0.029	0.081	0.022
		2	0.139	0.132	0.125	0.079	0.100	0.047	0.090	0.032
		4	0.216	0.223	0.179	0.145	0.130	0.087	0.108	0.055
		6	0.325	0.356	0.278	0.246	0.208	0.163	0.149	0.104
		8	0.472	0.500	0.412	0.389	0.337	0.260	0.231	0.161
	6	0	0.091	0.086	0.070	0.037	0.042	0.012	0.034	0.005
		2	0.128	0.128	0.095	0.065	0.054	0.026	0.040	0.008
		4	0.184	0.183	0.141	0.117	0.086	0.059	0.050	0.018
		6	0.268	0.284	0.205	0.193	0.133	0.112	0.077	0.044
		8	0.383	0.402	0.320	0.319	0.231	0.209	0.123	0.091

Table 7. Application to Bitcoin data: test values, bootstrap p-values, bubble start and end dates.

PSY	$sPSY$	$\bar{s}PSY$	$uPSY^b$	$\bar{u}PSY^b$
3.203	5.515	4.758	5.515	4.758
(0.180)	(0.019)	(0.025)	(0.033)	(0.050)
[0.007]				
	start: 13/11/2017	start: 4/10/2017		
	end: 7/12/2017	end: 7/12/2017		

Figure 1. Bitcoin data: explosive regime start and end dates.