JOURNAL OF TIME SERIES ANALYSIS
J. Time Ser. Anal. (2024)
Published online in Wiley Online Library
(wileyonlinelibrary.com) DOI: 10.1111/jtsa.12784

SPECIAL ISSUE ARTICLE

A NEW HETEROSKEDASTICITY-ROBUST TEST FOR EXPLOSIVE BUBBLES

DAVID I. HARVEY,[a] STEPHEN J. LEYBOURNE,[a] A. M. ROBERT TAYLOR[b*] AND YANG ZU[c]

[a]*School of Economics, University of Nottingham, Nottingham, UK*
[b]*Essex Business School, University of Essex, Colchester, UK*
[c]*Department of Economics, University of Macau, Macau, China*

We propose a new class of modified regression-based tests for detecting asset price bubbles designed to be robust to the presence of general forms of both conditional and unconditional heteroskedasticity in the price series. This modification, based on the approach developed in Beare (2018) in the context of conventional unit root testing, is achieved by purging the impact of unconditional heteroskedasticity from the data using a kernel estimate of volatility before the application of the bubble detection methods proposed in Phillips, Shi and Yu (2015) (PSY). The modified statistic is shown to achieve the same limiting null distribution as the corresponding (heteroskedasticity-uncorrected) statistic from PSY would obtain under homoskedasticity, such that the usual critical values provided in PSY may still be used. Versions of the test based on regressions including either no intercept or a (redundant) intercept are considered. Representations for asymptotic local power against a single bubble model are also derived. Monte Carlo simulation results highlight that neither one of these tests dominates the other across different bubble locations and magnitudes, and across different models of time-varying volatility. Accordingly, we also propose a test based on a union of rejections between the with- and without-intercept variants of the modified PSY test. The union procedure is shown to perform almost as well as the better of the constituent tests for a given DGP, and also performs very well compared to existing heteroskedasticity-robust tests across a large range of simulation DGPs.

Received 14 December 2023; Accepted 10 September 2024

Keywords: Rational bubble; explosive autoregression; time-varying volatility; kernel smoothing; right-tailed unit root testing; union of rejections.

JEL. C12; C32.

MOS subject classification: 62M10.

1. INTRODUCTION

Asset price bubbles tend to be characterised by a sudden and explosive increase in the price of an asset without a corresponding increase in the fundamental value of the asset (thereby representing a misallocation of resources), usually followed by a subsequent destruction of value through a sharp and catastrophic price collapse. As such, bubbles often presage economic recessions; indeed, the 2007/08 Global Financial Crisis was preceded by suspected price bubbles in the US housing, commodity, and stock markets. In the aftermath of the crisis, policymakers reacted by considering new rules for macroprudential regulation and intervention.

As a result, the development of econometric methods to empirically identify asset price bubbles has been the focus of much recent research. Explosive behaviour in financial asset price series is closely related to the theory of rational bubbles, with a rational bubble deemed to have occurred if explosive characteristics are manifest in the time path of prices, but not dividends. Accordingly, Phillips *et al.* (2015) (PSY) model potential bubble behaviour using a time-varying autoregression which allows for explosive autoregressive regimes within distinct subsets of

*Correspondence to: A. M. Robert Taylor, Essex Business School, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK.
E-mail: robert.taylor@essex.ac.uk

the data. PSY propose bubble detection tests based on a double supremum of forward and backward recursive right-tailed Dickey–Fuller (DF) statistics, a generalisation of the original and widely used Phillips *et al.* (2011) (PWY) test that employed a single supremum of forward-only recursive DF statistics.

A key assumption underlying both the PWY and PSY bubble detection procedures is that the innovations driving the asset price series are conditionally and unconditionally homoskedastic. While the presence of conditional heteroskedasticity, such as GARCH, in a wide range of financial variables, including asset returns, is a well-established stylised fact, a number of recent empirical studies have also questioned the reasonableness of the unconditional homoskedasticity assumption. In particular, strong evidence of structural breaks in the unconditional variance of asset returns is reported in, *inter alia*, Rapach *et al.* (2008), McMillan and Wohar (2011), Calvo-Gonzalez *et al.* (2010) and Vivian and Wohar (2012). Harvey *et al.* (2016) demonstrate that the asymptotic null distribution of the PWY test depends on the time-path of the underlying unconditional volatility. As a result if the test is compared to critical values derived under a homoskedastic error assumption, its size is not in general controlled if volatility is time varying, with a higher than expected probability of spurious bubble identification resulting. Harvey *et al.* (2016) propose a wild bootstrap implementation of the PWY test, which delivers correct asymptotic size in the presence of time-varying volatility. A similar wild bootstrap approach can be applied to the tests proposed in PSY to allow for unconditionally heteroskedastic errors and is considered in Harvey *et al.* (2020) (HLZ) as a comparator for their sign-based version of the PSY test. HLZ find that, under heteroskedasticity, their sign-based approach generally outperforms these bootstrap PSY tests.

Our contribution in this article is to develop a bootstrap-free approach to obtaining heteroskedasticity-robust versions of the PSY bubble detection tests. To that end, we follow the volatility re-scaling approach developed by Beare (2018) in the context of conventional full sample unit root tests directed against a stationary alternative. This entails calculating the PSY test statistic not from the original data, denoted y_t, but instead from the series of cumulated first differences, Δy_t, of the data standardised by a kernel estimate $\hat{\sigma}_t$ of σ_t, the volatility of Δy_t. That is, we cumulate $\Delta y_t/\hat{\sigma}_t$ and treat this volatility-purged cumulated series as the data that we input into the calculation of the PSY statistic. This approach has parallels with the approach taken in HLZ who, instead of cumulating $\Delta y_t/\hat{\sigma}_t$, base the PSY statistic on the cumulation of $\Delta y_t/\left|\Delta y_t\right| = \text{sign}(\Delta y_t)$, a quantity which is by construction exact invariant to the pattern of time-varying volatility in Δy_t under the null. HLZ show that this approach leads to power gains when compared to a wild bootstrap implementation of the usual PSY test. The main aim of this article is a comparison of the performance of the two non-bootstrap approaches to gauge whether using $\Delta y_t/\hat{\sigma}_t$ in place of the binary quantity $\Delta y_t/\left|\Delta y_t\right|$ might lead to further improvements in bubble detection efficacy. We would anticipate a gain in power, given that $\hat{\sigma}_t$ uses more sample information to estimate σ_t than in HLZ, where σ_t is essentially proxied by $\left|\Delta y_t\right|$ alone.

Under the unit root null and alternative of a single locally explosive bubble regime, we derive the asymptotic distribution of the two variants of PSY tests based on $\Delta y_t/\hat{\sigma}_t$: one which most closely follows Beare (2018) and fits an intercept term in the underlying DF regression, and another which omits this term which is in fact redundant under our model specification. Using a number of different specifications for a bubble/collapse DGP and pattern of time-varying volatility, we find that the local asymptotic power of our new tests compares very well with that of the HLZ tests, with the better of our two new tests outperforming the better of the HLZ tests for a given DGP. Between our two new tests, we find that which offers the better power performance depends on the bubble/collapse/volatility specification under consideration. This prompts us to consider a union of rejections testing strategy that combines the with-intercept and without-intercept test variants. We find that this strategy performs very well, capturing almost all of the power available from the better performing of the two individual tests across the full range of bubble/collapse/volatility specifications that we examine. As such, it also outperforms the better performing of the two HLZ tests.

The remainder of the article is organised as follows. Section 2 outlines the heteroskedastic bubble DGP we work with and the assumptions under which we will operate. Section 3 introduces our heteroskedasticity-modified version of the PSY test. Here we also establish the limit distributions of the re-scaled PSY statistics under local alternatives for the case of a single bubble episode. Asymptotic local powers of our new tests are numerically compared with the HLZ tests in Section 4. A union of rejections procedure is outlined in Section 5. The results from a Monte Carlo study exploring the finite sample properties of the tests are discussed in Section 6. Section 7

DOI: 10.1111/jtsa.12784

concludes. Proofs of our asymptotic results are provided in Appendix A. An accompanying Appendix S1 contains the finite sample simulation results discussed in Section 6.

In what follows '$\lfloor . \rfloor$' denotes the integer part of its argument, '$\mathbf{1}(\cdot)$' denotes the indicator function, '$\xrightarrow{p}$' and '$\Rightarrow$' respectively, denote convergence in probability and weak convergence, in each case as the sample size diverges, and '$x := y$' ('$x =: y$') denotes that x (y) is defined by y (x).

2. THE HETEROSKEDASTIC STOCHASTIC BUBBLE MODEL

To keep the contents of the article as tractable as possible, we focus attention on the case where a single bubble episode is present under the alternative. However, it is important to stress that the modified heteroskedasticity-robust versions of PSY's GSADF test that we develop in this article are, like the original GSADF test, also valid tests for models where multiple bubbles are present under the alternative. We will also only discuss volatility re-scaled modifications of the leading doubly recursive GSADF test from PSY. The same re-scaling principle can be applied to the (singly) forward and backward recursive tests, SADF and BSADF respectively, discussed in PSY (the former coinciding with the test developed in PWY).

To that end, we follow HLZ and focus attention on the time series process $\{y_t\}$ generated according to the following DGP,

$$y_t = \mu + u_t, \quad t = 1, \dots, T, \tag{1}$$

$$u_t = \rho_t u_{t-1} + \varepsilon_t, \quad \varepsilon_t = \sigma_t z_t, \quad t = 2, \dots, T \tag{2}$$

where $\rho_t := \rho(t/T)$ with

$$\rho(t/T) = \begin{cases} 1 & t = 2, \dots, \lfloor \tau_1 T \rfloor, \\ 1 + \delta_1 & t = \lfloor \tau_1 T \rfloor + 1, \dots, \lfloor \tau_2 T \rfloor, \\ 1 - \delta_2 & t = \lfloor \tau_2 T \rfloor + 1, \dots, \lfloor \tau_3 T \rfloor, \\ 1 & t = \lfloor \tau_3 T \rfloor + 1, \dots, T, \end{cases} \tag{3}$$

where $\delta_1 \geq 0$ and $\delta_2 \geq 0$, and where $0 \leq \tau_1 < \tau_2 < \tau_3 \leq 1$. We assume that the initial condition u_1 is such that $u_1 = o_p(T^{1/2})$. In the context of (2), ε_t is a zero-mean innovation process with (possibly) time-varying volatility function, σ_t, precise conditions on which are given in Assumptions 1 and 2 below.

The DGP given by (1)–(3) imposes a unit root on y_t up to time $\lfloor \tau_1 T \rfloor$, after which y_t is an explosive bubble process when $\delta_1 > 0$ until time $\lfloor \tau_2 T \rfloor$. Notice that this explosive regime would originate at the beginning of the sample if $\tau_1 = 0$. If $\tau_2 < 1$, the explosive regime then ends in-sample, at which point the model permits a possible collapse, with $\delta_2 > 0$ creating a collapse regime modelled by stationary mean-reverting behaviour. 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. $H_0 : \delta_i = 0, i = 1, 2$. The alternative hypothesis is $H_1 : \delta_1 > 0$ and $\delta_2 \geq 0$. In its most general form, where $\delta_2 > 0$ with $\tau_3 < 1$, the alternative is that y_t is unit root, followed by bubble, then collapse, before returning to a unit root regime. Special cases of this alternative are clearly permitted within the framework of (3). For example, if $\delta_2 = 0$ the bubble regime does not collapse but terminates in a unit root regime; if $\delta_2 > 0$ with $\tau_3 = 1$, the collapse period runs to the end of the sample. Under H_1 we will consider locally explosive alternatives (and collapses) of the form $\delta_i = c_i T^{-1}$, $c_i \geq 0$, $i = 1, 2$, the scaling by T^{-1} providing the appropriate Pitman drift for asymptotic power comparisons of the tests when $c_1 > 0$ and $c_2 \geq 0$. Throughout our analysis, we assume that a collapse does not occur without the presence of a prior bubble, i.e. when $c_1 = 0$, we assume $c_2 = 0$.

With respect to the error process, ε_t, in (1), we make Assumptions 1 and 2:

Assumption 1. (a) Let z_t be a martingale difference sequence (MDS) with respect to the natural filtration, $\mathcal{F}_t$, generated by $\{z_s, s \leq t\}$, with unit unconditional variance, $E(z_t^2) = 1$, and where $T^{-1} \sum_{t=2}^{T} z_t^2 \xrightarrow{p} 1$. (b) For all integers

q such that $2 \leqslant q \leqslant 8$ and for all integers $r_1, \ldots, r_{q-1} \geqslant 0$, the qth order cumulants $\kappa_q(t, t-r_1, \ldots, t-r_{q-1})$ of $(z_t, z_{t-r_1}, \ldots, z_{t-r_{q-1}})$ satisfy the condition that $\sup_t \sum_{r_1,\ldots,r_{q-1}=-\infty}^{\infty} |\kappa_q(t, t-r_1, \ldots, t-r_{q-1})| < \infty$.

Assumption 2. The volatility function is non-stochastic and satisfies $\sigma_t := \sigma(t/T)$, where $\sigma(\cdot)$ is a strictly positive càdlàg function (which therefore allows for a countable number of jumps in volatility) with support [0, 1] and is uniformly bounded by a constant M.

Remark 2.1. Part (a) of Assumption 1 ensures that a MDS functional central limit theorem (FCLT) holds on the innovations $\{z_t\}$; cf. Assumption 1 of Chang and Park (2002, p. 433). Part (b) of Assumption 1 coincides with Assumption 1 (iii) of Goncalves and Kilian (2007). These conditions allow for conditional heteroskedasticity of an unknown and quite general form. An implication of the restrictions placed on the cumulants by Assumption 1(b) is that $\sup_t E(\epsilon_t^8) < \infty$. This moment assumption appears standard in the related literature where a kernel smoothed estimate of the volatility function is required and is also imposed by, *inter alia*, Xu and Phillips (2008), Harvey *et al.* (2019), Cavaliere *et al.* (2022) (CNT), Boswijk and Zu (2018, 2022). An exception is Beare (2018) whose method of proof only requires a finite fourth moment assumption. The trade-off for this weaker moment condition is that Beare (2018) needs to impose a continuous differentiability condition on $\sigma(\cdot)$ which is stronger than our Assumption 2 which allows for a discontinuous unconditional volatility function. The large sample results given in this article should also hold under the mixing and moment conditions on ϵ_t and the continuously differentiable condition on $\sigma(\cdot)$ adopted in Beare (2018). ⋄

Remark 2.2. Assumption 2 implies that σ_t is the unconditional volatility of ϵ_t. Under Assumption 2 the volatility process can display (possibly) multiple instantaneous 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 with $\sigma(t/T) = \sigma$ for all t. Consequently both the conditional and unconditional variance of ϵ_t are allowed to display time-varying behaviour under Assumptions 1 and 2. ⋄

Remark 2.3. Broadly similar conditions to those placed on $\sigma(\cdot)$ by Assumption 2 are adopted in Assumption 1(b) of CNT who also allow $\sigma(\cdot)$ to be a càdlàg function. Our conditions on $\sigma(\cdot)$ are rather weaker than those imposed by, for example, assumption (i) of Xu and Phillips (2008) which requires $\sigma(\cdot)$ to satisfy a uniform first-order Lipschitz condition with at most a finite number of discontinuities. ⋄

Remark 2.4. It is also instructive to compare our assumptions with those made in the extant bubble testing literature. PWY and PSY impose that $\epsilon_t \sim i.i.d(0, \sigma^2)$. Harvey *et al.* (2016) develop wild bootstrap implementations of the PWY tests which allow for unconditional heteroskedasticity in ϵ_t of a similar form to Assumption 2, but impose conditional homoskedasticity on z_t. Similarly, although HLZ allow for unconditional heteroskedasticity in developing their sign-based tests for explosive bubbles, they also impose conditional homoskedasticity on z_t. Harvey *et al.* (2019) also impose conditional homoskedasticity on z_t in their weighted least squares implementation of the PWY test. To the best of our knowledge then, this article develops the only regression-based bubble detection tests currently available in the literature that allow for both conditional and unconditional heteroskedasticity in the errors. Given the empirical findings on the nature of volatility in financial price series discussed in Section 1, this should render these methods attractive to practitioners. ⋄

3. VOLATILITY RE-SCALED PSY TESTS

To obtain a volatility-robust version of the GSADF test developed in PSY, rather than calculating the statistic directly from the observed data series, y_t, we instead propose calculating it from the series of cumulated first differences of the data, Δy_t, re-scaled by a kernel estimate of the volatility of Δy_t. That is, we cumulate $\Delta y_t / \hat{\sigma}_t$, where $\hat{\sigma}_t$ denotes the kernel estimate of σ_t, and treat this volatility-adjusted cumulated series as the data for the GSADF statistic. This approach follows in the spirit of Beare (2018), who used cumulated standardised differences in the context of full sample unit root testing against a stationary alternative in the presence of heteroskedasticity.

We first define a non-parametric variance estimator of the form

$$\hat{\sigma}_t^2 := \sum_{j=2}^{T} w_{t,j}(y_j - y_{j-1})^2, \tag{4}$$

where the kernel weights, $w_{t,j}$, are given by

$$w_{t,j} := \frac{K\left(\frac{j-t}{Th}\right)}{\sum_{j=2}^{T} K\left(\frac{j-t}{Th}\right)},$$

where $K(\cdot)$ is a kernel function and h the associated bandwidth, precise conditions on both of which will be given later.

Next we construct the cumulated first differences of Δy_t standardised by $\hat{\sigma}_t$, i.e.

$$x_t := \sum_{j=2}^{t} \frac{\Delta y_j}{\hat{\sigma}_j}, \quad t = 2, \ldots, T. \tag{5}$$

Notice that the volatility-robust test of HLZ is based on $\Delta y_t / |\Delta y_t|$, which arises as a special case of $\Delta y_t / \hat{\sigma}_t$ if we set $w_{t,j} = \mathbf{1}(j = t)$. This is therefore essentially equivalent to imposing a bandwidth of zero in the kernel function.

Our proposed statistic is then a volatility re-scaled version of PSY's GSADF statistic, constructed from x_t rather than y_t; that is,

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

where using generic notation, DF_σ is the t-ratio for $\hat{\phi}(\lambda_1, \lambda_2)$ in the (with-intercept) fitted Dickey–Fuller OLS regression

$$\Delta x_t = \hat{\alpha}(\lambda_1, \lambda_2) + \hat{\phi}(\lambda_1, \lambda_2) x_{t-1} + e_t, \tag{6}$$

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

$$\text{DF}_\sigma(\lambda_1, \lambda_2) := \frac{\hat{\phi}(\lambda_1, \lambda_2)}{\sqrt{\hat{s}^2(\lambda_1, \lambda_2) / \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} \left(x_{t-1} - \bar{x}\right)^2}}, \tag{7}$$

where $\bar{x} := (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} x_{t-1}$ and $\hat{s}^2(\lambda_1, \lambda_2) := (\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 2)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor + 1}^{\lfloor \lambda_2 T \rfloor} e_t^2$, with e_t being the residuals in (6). The PSY_σ statistic is 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.[1]

In the fitted regression (6), we have followed Beare (2018) and included an intercept term. However, the x_t are, by construction, numerically invariant to the nuisance parameter μ in the DGP (1) and so there is no requirement to include this intercept term. As such, we also consider the corresponding statistic that excludes the intercept term in (6). Denoting the corresponding without-intercept version of (7) by $\text{DF}_\sigma^*(\lambda_1, \lambda_2)$, the corresponding PSY statistic is given by

$$\text{PSY}_\sigma^* := \sup_{\lambda_1 \in [0, 1-\pi]} \sup_{\lambda_2 \in [\lambda_1 + \pi, 1]} \text{DF}_\sigma^*(\lambda_1, \lambda_2).$$

[1] In practice, in view of (5), we need to impose a minimum value of $2/T$ on λ_1, such that the earliest possible starting time index in the regression in (6) is $t = \lfloor T(2/T) \rfloor + 1 = 3$.

To obtain the asymptotic distributions of the two statistics, we first need to establish the large sample behaviour of the partial sum process, x_t. This requires us to place certain conditions on the kernel and bandwidth used in constructing the non-parametric variance estimator, $\hat{\sigma}_t^2$ in (4). Specifically, we make the following assumptions, both of which are typical in the literature; see, for example, Xu and Phillips (2008), Harvey *et al.* (2019), CNT, and Boswijk and Zu (2018, 2022). Notice, however, that Assumption 4 is less restrictive than the corresponding bandwidth rate condition imposed in Beare (2018), which requires that $Th^4 \to \infty$.

Assumption 3. The kernel function $K(\cdot)$ is a bounded, non-negative and continuous function defined on the real number line and $\int_{-\infty}^{\infty} K(x)\mathrm{d}x = 1$.

Assumption 4. The bandwidth h is such that as $T \to \infty$, $h \to 0$ and $Th^2 \to \infty$.

We are now in a position to detail the large sample behaviour of the partial sum process x_t under the most general form of the model in (1)–(3):

Theorem 1. Let y_t satisfy (1)–(3) and let Assumptions 1–4 hold. Then, under $H_1 : \delta_i = c_i T^{-1}$, $c_i \geq 0$, $i = 1, 2$,

$$T^{-1/2} x_{\lfloor Tr \rfloor} \Rightarrow \mathcal{X}(r) =: \begin{cases} W(r) & r \leqslant \tau_1 \\ W(r) + c_1 \int_{\tau_1}^{r} \frac{V_1(s)}{\sigma(s)} \mathrm{d}s & \tau_1 < r \leqslant \tau_2 \\ W(r) + c_1 \int_{\tau_1}^{\tau_2} \frac{V_1(s)}{\sigma(s)} \mathrm{d}s - c_2 \int_{\tau_2}^{r} \frac{V_2(s)}{\sigma(s)} \mathrm{d}s & \tau_2 < r \leqslant \tau_3 \\ W(r) + c_1 \int_{\tau_1}^{\tau_2} \frac{V_1(s)}{\sigma(s)} \mathrm{d}s - c_2 \int_{\tau_2}^{\tau_3} \frac{V_2(s)}{\sigma(s)} \mathrm{d}s & r > \tau_3 \end{cases}, \tag{8}$$

where $V_1(r) := \mathrm{e}^{(r-\tau_1)c_1} \int_0^{\tau_1} \sigma(v)\mathrm{d}W(v) + \int_{\tau_1}^{r} \mathrm{e}^{(r-v)c_1} \sigma(v)\mathrm{d}W(v)$, $V_2(r) := \mathrm{e}^{-(r-\tau_2)c_2} V_1(\tau_2) + \int_{\tau_2}^{r} \mathrm{e}^{-(r-v)c_2} \sigma(v)\mathrm{d}W(v)$, and where $W(r)$ is a standard Brownian motion on $[0, 1]$.

Remark 3.1. Theorem 1 establishes the limiting distribution of x_t under both the null hypothesis and local alternatives. Under H_0, $c_1 = c_2 = 0$ and $\mathcal{X}(\cdot)$ reduces to the standard Brownian motion, $W(\cdot)$, and hence does not depend on the underlying volatility process, $\sigma(\cdot)$. Under H_1 the asymptotic distribution of $\mathcal{X}(r)$ depends on the constants c_1 and c_2, and also on the volatility process $\sigma(\cdot)$. ⋄

Next, in Theorem 2, we detail the large sample behaviour of PSY_σ and PSY_σ^* under both the null and local alternative hypotheses.

Theorem 2. Let the conditions of Theorem 1 hold. Then,

$$\mathrm{PSY}_\sigma \Rightarrow \sup_{\lambda_1 \in [0, 1-\pi]} \sup_{\lambda_2 \in [\lambda_1 + \pi, 1]} L_{c_1, c_2, \sigma}(\lambda_1, \lambda_2) =: G_{c_1, c_2, \sigma}$$
$$\mathrm{PSY}_\sigma^* \Rightarrow \sup_{\lambda_1 \in [0, 1-\pi]} \sup_{\lambda_2 \in [\lambda_1 + \pi, 1]} L^*_{c_1, c_2, \sigma}(\lambda_1, \lambda_2) =: G^*_{c_1, c_2, \sigma},$$

where

$$L_{c_1, c_2, \sigma}(\lambda_1, \lambda_2) := \frac{\tilde{\mathcal{X}}(\lambda_2)^2 - \tilde{\mathcal{X}}(\lambda_1)^2 - (\lambda_2 - \lambda_1)}{2\sqrt{\int_{\lambda_1}^{\lambda_2} \tilde{\mathcal{X}}(r)^2 \mathrm{d}r}}$$
$$L^*_{c_1, c_2, \sigma}(\lambda_1, \lambda_2) := \frac{\mathcal{X}(\lambda_2)^2 - \mathcal{X}(\lambda_1)^2 - (\lambda_2 - \lambda_1)}{2\sqrt{\int_{\lambda_1}^{\lambda_2} \mathcal{X}(r)^2 \mathrm{d}r}},$$

with $\tilde{\mathcal{X}}(r) := \mathcal{X}(r) - (\lambda_2 - \lambda_1)^{-1} \int_{\lambda_1}^{\lambda_2} \mathcal{X}(v)\mathrm{d}v$, where $\mathcal{X}(r)$ is defined in Theorem 1.

Table I. Asymptotic and finite sample critical values for ξ-level tests

	PSY_σ			PSY^*_σ			UPSY_σ		
	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$	$\xi = 0.10$	$\xi = 0.05$	$\xi = 0.01$
$T = 100$	1.629	1.828	2.392	3.637	4.158	5.553	3.950	4.527	6.129
$T = 200$	1.608	1.789	2.140	3.226	3.595	4.330	3.468	3.804	4.589
$T = 400$	1.712	1.935	2.296	3.167	3.446	4.007	3.361	3.598	4.145
$T = \infty$	1.875	2.094	2.486	2.978	3.296	3.859	3.186	3.486	3.951

Remark 3.2. The limit distributions of PSY_σ and PSY^*_σ under the null hypothesis H_0 are given by $G_{0,0,\sigma}$ and $G^*_{0,0,\sigma}$ respectively, which do not depend on $\sigma(s)$. For PSY_σ, $G_{0,0,\sigma}$ coincides with the limiting null distribution given for the GSADF statistic in Theorem 1 of PSY (p. 1049) for the case where $\varepsilon_t \sim i.i.d(0, \sigma^2)$. Consequently, for any volatility process satisfying Assumption 2, the limiting null distribution of PSY_σ coincides with the limit null distribution of the standard GSADF statistic proposed in PSY (based on y_t) that would obtain were the volatility constant, with the same applying for PSY^*_σ. Under H_1, the limiting distributions of PSY_σ and PSY^*_σ depend on c_1 and c_2 and on the volatility process $\sigma(\cdot)$. ⋄

Remark 3.3. We have assumed thus far that ε_t is serially uncorrelated. More generally, we might wish to allow it to admit a finite autoregressive representation of the form[2]

$$\varepsilon_t = \sum_{i=1}^{p} \theta_i \varepsilon_{t-i} + \sigma_t z_t, \tag{9}$$

with the autoregressive coefficients θ_i, $i = 1, \ldots, p$, satisfying standard stability conditions, such that ε_t would be covariance stationary in the unconditionally homoskedastic case where $\sigma_t^2 = \sigma^2$, for all t. In this situation, in constructing PSY_σ and PSY^*_σ the subsample regression (6), and its without-intercept equivalent respectively, need to be augmented with the p lagged difference terms $\Delta x_{t-1}, ..., \Delta x_{t-p}$. The non-parametric variance estimator defined in (4), based on first-differences, can still be used to construct the sample data $\{x_t\}$, as in (5). Under the null, in this context $\hat{\sigma}_t^2$ provides an estimate of $\text{Var}(\Delta y_t)$, and so the resulting $\Delta x_t = \Delta y_t / \hat{\sigma}_t$ sequence will follow an *approximate* AR(p) model with homoskedastic innovations. Under the conditions of Theorem 1, our proof can in principle be extended, to show that the limiting null distributions of the resulting augmented PSY_σ and PSY^*_σ statistics are still given by $G_{0,0,\sigma}$ and $G^*_{0,0,\sigma}$ respectively. Notice that $\text{Var}(\Delta y_t)$ under specification (9) is approximately equal under the null to σ_t multiplied by a (time-invariant) constant determined by the AR coefficients $\theta_j, j = 1, ..., p$, from (9). In view of this, one could alternatively estimate σ_t non-parametrically directly using the residuals from estimating a full-sample ADF regression (that is, the regression of Δy_t on a constant, y_{t-1}, and p lags of Δy_t) and use this residual-based estimate to construct the sample $\{x_t\}$. The resulting $\{x_t\}$ sample will also follow an approximate homoskedastic $AR(p)$ model under the null. ⋄

For the setting $\pi = 0.1$, asymptotic upper-tail critical values of the null distributions $G_{0,0,\sigma}$ and $G^*_{0,0,\sigma}$ are given in Table I for the usual significance levels. Here, and throughout our asymptotic analysis, we approximated $W(r)$ using NIID(0, 1) random variates with discretised normalised sums of 1000 steps. Table I also provides finite sample null critical values of PSY_σ and PSY^*_σ based on generating ε_t as NIID(0, 1) and using a Gaussian kernel for $K(\cdot)$ with bandwidth setting $h = 0.1T^{-0.25}$ (in line with our setting for the finite sample simulations below). Monte Carlo results throughout this article are based on 2000 replications.

[2] At a practical level, the doubly recursive nature of the PSY-type approach, which includes relatively short sub-sample sizes, means that only a small fixed value of p should be used, in line with the recommendation of PSY. In any case, substantial serial correlation would not be expected in ε_t, given that under the null this process represents asset returns which should, at least in theory, be serially uncorrelated.

In the next section, we examine the asymptotic local powers of the PSY_σ and PSY^*_σ tests, comparing these with the sign-based tests of HLZ, and explore the extent to which any gain is obtained by excluding the intercept term in the underlying DF regressions.

4. ASYMPTOTIC LOCAL POWER COMPARISONS

We examine the asymptotic power of the PSY_σ and PSY^*_σ tests under the locally explosive alternative H_1. By way of comparison, we also simulate the asymptotic power of the two sign-based tests of HLZ: sPSY and $\bar{s}$PSY, using HLZ's notation.[3] The sPSY test relies on an assumption of a zero median in the distribution of z_t, while $\bar{s}$PSY, which is based on recursively demeaned $sign(\Delta y_t)$ rather than simply $sign(\Delta y_t)$, controls size irrespective of whether the median is zero or not. We consider both a benchmark case of homoskedasticity, $\sigma(r) = 1$, and also heteroskedastic settings with volatility functions $\sigma(r) = S(r, \sigma_1, \sigma_2, \tau_\sigma, \gamma)$ where $S(r, \sigma_1, \sigma_2, \tau_\sigma, \gamma)$ is a logistic smooth transition function of the form

$$S(r, \sigma_1, \sigma_2, \tau_\sigma, \gamma) := \sigma_1 + \frac{\sigma_2 - \sigma_1}{1 + \exp\{-\gamma(r - \tau_\sigma)\}}.$$

This function transitions from the value σ_1 to σ_2 over r, with midpoint fraction τ_σ and transition speed γ. Specifically, we set $\gamma = 30$, $\sigma_1 = 1$ and consider $\sigma_2 \in \{1/6, 1/3, 3, 6\}$ allowing downshifts and upshifts in volatility with $\tau_\sigma \in \{0.4, 0.8\}$ to represent earlier and later volatility midpoint timings. For the locally explosive alternatives (and possible collapses) we have $\{\tau_1, \tau_2, \tau_3\} = \{0.1, 0.4, 0.6\}$, $\{0.3, 0.6, 0.8\}$ and $\{0.5, 0.8, 1.0\}$ for early, middle and late explosive episodes respectively, with explosive magnitudes $c_1 \in \{2, 4, 6, 8, 10\}$ and collapse magnitudes $c_2 = kc_1$ with $k \in \{0, 0.5, 1\}$, such that an explosive episode is either unaccompanied by a collapse, or accompanied by a collapse with a parameter value of half or equal to the explosive magnitude. We simulate the asymptotic powers of upper-tail nominal 0.05-level tests using limit null critical values for PSY_σ and PSY^*_σ taken from Table I. For sPSY and $\bar{s}$PSY we use the limit null critical values in Table I of HLZ.[4] Here $\sigma(r)$, like $W(r)$, is discretised over 1000 steps.

Consider first the homoskedastic case of Table II. Comparing PSY_σ and PSY^*_σ, we see that PSY^*_σ is generally more powerful than PSY_σ, particularly when the explosive episode occurs early in the sample, where power gains of up to around 0.11 can be seen. In turn, PSY_σ is generally more powerful than sPSY (with some exceptions for the earlier explosive cases), while sPSY is always more powerful than $\bar{s}$PSY. Here then, PSY^*_σ emerges pretty unambiguously as the preferred test.

Turning to the first heteroskedastic specification in Table III, $S(r, 1, 1/6, 0.4, 30)$ (early large downward shift in volatility) a rather more involved picture emerges. While there are some cases where the power of PSY^*_σ exceeds that of PSY_σ, the reverse pattern can arise when a collapse is present, with the power of PSY_σ actually much higher than that of PSY^*_σ when the explosive episode occurs early in the sample, with power differences up to around 0.22. Notice also that $\bar{s}$PSY is more powerful than sPSY and PSY^*_σ in this early explosive episode with collapse environment. PSY_σ almost always has higher power than the better of sPSY and $\bar{s}$PSY, and therefore emerges as arguably the best test overall for these settings. For $S(r, 1, 1/3, 0.4, 30)$ (early small downward shift in volatility), much the same overall comments apply. For the heteroskedastic specifications in Table IV, $S(r, 1, 3, 0.4, 30)$ (early small upward shift in volatility) and $S(r, 1, 6, 0.4, 30)$ (early large upward shift in volatility), the pattern of results is broadly in line with the homoskedastic case, with PSY^*_σ representing the best-performing test, almost without exception.

In Table V, the timing of the volatility change is now later, with $S(r, 1, 1/6, 0.8, 30)$ (late large downward shift in volatility) and $S(r, 1, 1/3, 0.8, 30)$ (late small downward shift in volatility). Here, we see that PSY^*_σ is the best performing test when the explosive episode timing is early or central, but PSY_σ performs best when the explosive episode occurs late in the sample. Lastly, for the heteroskedastic specifications in Table VI $S(r, 1, 3, 0.8, 30)$

---

[3] Given that HLZ find that sPSY generally outperforms a wild bootstrap implementation of the PSY test under heteroskedasticity, we do not consider bootstrap-based tests in our present comparison.

[4] Local asymptotic powers of sPSY and $\bar{s}$PSY depend on the probability density function of z_t. For the purposes of this comparison exercise we assume, as in HLZ, that z_t is normally distributed.

DOI: 10.1111/jtsa.12784

Table II. Local asymptotic power of nominal 0.05-level tests

					$\sigma(r) = 1$				
τ_1	τ_2	τ_3	c_1	c_2	PSY$_\sigma$	PSY$^*_\sigma$	UPSY$_\sigma$	sPSY	$\bar{s}$PSY
0.1	0.4	0.6	2	0	0.072	0.105	0.087	0.108	0.061
				1	0.068	0.091	0.079	0.098	0.070
				2	0.068	0.084	0.076	0.092	0.082
			4	0	0.183	0.292	0.263	0.244	0.107
				2	0.181	0.268	0.242	0.222	0.136
				4	0.209	0.263	0.251	0.214	0.176
			6	0	0.466	0.547	0.527	0.487	0.262
				3	0.469	0.524	0.512	0.464	0.306
				6	0.486	0.518	0.513	0.462	0.371
			8	0	0.720	0.754	0.743	0.705	0.538
				4	0.715	0.735	0.728	0.684	0.560
				8	0.722	0.731	0.730	0.679	0.591
			10	0	0.856	0.869	0.861	0.840	0.765
				5	0.854	0.861	0.856	0.832	0.776
				10	0.855	0.861	0.855	0.831	0.782
0.3	0.6	0.8	2	0	0.099	0.139	0.126	0.121	0.050
				1	0.097	0.126	0.119	0.111	0.063
				2	0.104	0.120	0.117	0.105	0.078
			4	0	0.334	0.403	0.382	0.336	0.117
				2	0.334	0.384	0.368	0.315	0.165
				4	0.353	0.380	0.374	0.309	0.219
			6	0	0.634	0.656	0.643	0.602	0.427
				3	0.628	0.641	0.631	0.587	0.461
				6	0.644	0.637	0.639	0.581	0.496
			8	0	0.808	0.815	0.806	0.779	0.686
				4	0.803	0.806	0.800	0.767	0.700
				8	0.805	0.804	0.800	0.763	0.713
			10	0	0.895	0.896	0.894	0.870	0.837
				5	0.889	0.893	0.888	0.867	0.842
				10	0.892	0.890	0.890	0.866	0.845
0.5	0.8	1.0	2	0	0.120	0.164	0.145	0.139	0.050
				1	0.119	0.151	0.140	0.128	0.055
				2	0.126	0.145	0.140	0.123	0.072
			4	0	0.433	0.474	0.457	0.406	0.192
				2	0.428	0.456	0.448	0.390	0.212
				4	0.448	0.453	0.457	0.387	0.264
			6	0	0.710	0.726	0.715	0.669	0.543
				3	0.705	0.717	0.710	0.656	0.558
				6	0.709	0.716	0.710	0.654	0.590
			8	0	0.858	0.863	0.858	0.838	0.778
				4	0.857	0.859	0.854	0.829	0.785
				8	0.858	0.857	0.856	0.829	0.803
			10	0	0.932	0.935	0.931	0.918	0.894
				5	0.931	0.931	0.929	0.914	0.896
				10	0.931	0.931	0.929	0.913	0.902

Notes: τ_1 and τ_2 denote the sample fraction at which the explosive period begins and ends; τ_3 denotes the end of the collapse regime; c_1 and c_2 denote the locally explosive and collapse magnitudes; $\sigma(r)$ denotes the volatility function.

Table III. Local asymptotic power of nominal 0.05-level tests

					$\sigma(r)=S(r,1,1/6,0.4,30)$					$\sigma(r)=S(r,1,1/3,0.4,30)$				
τ_1	τ_2	τ_3	c_1	c_2	PSY_σ	PSY^*_σ	UPSY_σ	$s\text{PSY}$	$\bar{s}\text{PSY}$	PSY_σ	PSY^*_σ	UPSY_σ	$s\text{PSY}$	$\bar{s}\text{PSY}$
0.1	0.4	0.6	2	0	0.074	0.115	0.096	0.113	0.061	0.074	0.113	0.092	0.111	0.061
				1	0.152	0.092	0.139	0.100	0.151	0.085	0.094	0.090	0.097	0.099
				2	0.321	0.103	0.296	0.127	0.276	0.145	0.086	0.131	0.097	0.152
			4	0	0.233	0.343	0.314	0.290	0.124	0.216	0.330	0.297	0.278	0.119
				2	0.515	0.318	0.496	0.298	0.438	0.323	0.294	0.325	0.257	0.267
				4	0.667	0.523	0.648	0.489	0.572	0.475	0.293	0.447	0.280	0.404
			6	0	0.554	0.605	0.588	0.548	0.344	0.535	0.591	0.574	0.534	0.320
				3	0.765	0.631	0.749	0.615	0.673	0.625	0.568	0.610	0.517	0.517
				6	0.837	0.778	0.828	0.745	0.766	0.718	0.579	0.700	0.570	0.620
			8	0	0.783	0.800	0.790	0.757	0.637	0.771	0.787	0.779	0.746	0.616
				4	0.886	0.835	0.881	0.816	0.835	0.814	0.778	0.803	0.742	0.736
				8	0.910	0.883	0.905	0.861	0.869	0.853	0.787	0.842	0.775	0.786
			10	0	0.894	0.896	0.893	0.875	0.821	0.884	0.891	0.886	0.867	0.810
				5	0.943	0.920	0.941	0.910	0.916	0.910	0.886	0.904	0.862	0.856
				10	0.954	0.940	0.951	0.921	0.927	0.921	0.892	0.916	0.875	0.875
0.3	0.6	0.8	2	0	0.472	0.519	0.493	0.443	0.229	0.257	0.321	0.294	0.273	0.085
				1	0.481	0.501	0.487	0.423	0.302	0.266	0.301	0.292	0.250	0.135
				2	0.545	0.497	0.533	0.425	0.414	0.306	0.294	0.310	0.246	0.208
			4	0	0.796	0.808	0.795	0.770	0.677	0.656	0.681	0.662	0.624	0.481
				2	0.795	0.795	0.786	0.761	0.703	0.659	0.665	0.656	0.610	0.516
				4	0.811	0.793	0.803	0.766	0.738	0.679	0.660	0.668	0.609	0.570
			6	0	0.904	0.908	0.903	0.878	0.855	0.836	0.841	0.834	0.811	0.751
				3	0.900	0.905	0.899	0.874	0.861	0.831	0.835	0.828	0.804	0.766
				6	0.906	0.903	0.902	0.878	0.872	0.838	0.834	0.832	0.805	0.779
			8	0	0.951	0.951	0.950	0.942	0.926	0.915	0.921	0.915	0.896	0.872
				4	0.949	0.949	0.947	0.939	0.930	0.914	0.918	0.913	0.888	0.877
				8	0.950	0.948	0.948	0.940	0.934	0.916	0.917	0.913	0.890	0.884
			10	0	0.977	0.977	0.977	0.971	0.967	0.953	0.953	0.951	0.947	0.929
				5	0.978	0.977	0.977	0.970	0.970	0.950	0.950	0.948	0.942	0.931
				10	0.978	0.977	0.977	0.971	0.972	0.951	0.950	0.948	0.942	0.934
0.5	0.8	1.0	2	0	0.587	0.616	0.595	0.556	0.353	0.343	0.406	0.386	0.350	0.114
				1	0.584	0.610	0.591	0.547	0.371	0.340	0.390	0.373	0.332	0.134
				2	0.607	0.607	0.603	0.545	0.434	0.369	0.383	0.386	0.329	0.201
			4	0	0.844	0.847	0.841	0.826	0.747	0.719	0.738	0.721	0.692	0.560
				2	0.843	0.842	0.838	0.820	0.753	0.717	0.729	0.715	0.682	0.573
				4	0.850	0.840	0.843	0.818	0.776	0.730	0.727	0.723	0.680	0.612
			6	0	0.938	0.939	0.933	0.924	0.901	0.875	0.878	0.871	0.855	0.807
				3	0.936	0.934	0.932	0.922	0.903	0.874	0.875	0.868	0.852	0.806
				6	0.938	0.933	0.934	0.921	0.908	0.875	0.873	0.871	0.850	0.820
			8	0	0.974	0.974	0.973	0.965	0.956	0.942	0.943	0.942	0.927	0.908
				4	0.973	0.972	0.971	0.963	0.956	0.941	0.941	0.940	0.924	0.911
				8	0.974	0.972	0.972	0.964	0.957	0.941	0.940	0.940	0.925	0.914
			10	0	0.987	0.986	0.987	0.981	0.980	0.971	0.970	0.970	0.966	0.956
				5	0.987	0.986	0.986	0.980	0.979	0.971	0.970	0.970	0.964	0.960
				10	0.988	0.986	0.987	0.980	0.980	0.972	0.970	0.970	0.965	0.961

Notes: τ_1 and τ_2 denote the sample fraction at which the explosive period begins and ends; τ_3 denotes the end of the collapse regime; c_1 and c_2 denote the locally explosive and collapse magnitudes; $\sigma(r)$ denotes the volatility function, with $S(r,\sigma_1,\sigma_2,\tau_\sigma,\gamma)$ denoting a smooth transition function from σ_1 to σ_2 with midpoint τ_σ and speed γ.

DOI: 10.1111/jtsa.12784

Table IV. Local asymptotic power of nominal 0.05-level tests

					$\sigma(r) = S(r, 1, 3, 0.4, 30)$					$\sigma(r) = S(r, 1, 6, 0.4, 30)$				
τ_1	τ_2	τ_3	c_1	c_2	PSY_σ	PSY^*_σ	UPSY_σ	sPSY	$\bar{s}$PSY	PSY_σ	PSY^*_σ	UPSY_σ	sPSY	$\bar{s}$PSY
0.1	0.4	0.6	2	0	0.070	0.093	0.081	0.101	0.063	0.068	0.086	0.080	0.095	0.061
				1	0.066	0.086	0.075	0.094	0.063	0.063	0.081	0.075	0.091	0.062
				2	0.064	0.081	0.072	0.088	0.064	0.060	0.077	0.070	0.087	0.061
			4	0	0.142	0.249	0.220	0.204	0.099	0.121	0.219	0.190	0.182	0.092
				2	0.135	0.234	0.203	0.194	0.101	0.115	0.206	0.179	0.172	0.091
				4	0.134	0.226	0.196	0.185	0.111	0.111	0.198	0.173	0.167	0.092
			6	0	0.365	0.470	0.449	0.409	0.194	0.294	0.426	0.400	0.361	0.171
				3	0.357	0.451	0.435	0.396	0.204	0.286	0.412	0.387	0.347	0.174
				6	0.356	0.447	0.432	0.393	0.213	0.283	0.408	0.383	0.343	0.173
			8	0	0.630	0.676	0.659	0.623	0.414	0.539	0.624	0.605	0.563	0.333
				4	0.623	0.660	0.647	0.610	0.418	0.532	0.611	0.594	0.552	0.335
				8	0.620	0.658	0.646	0.604	0.423	0.531	0.610	0.591	0.547	0.335
			10	0	0.805	0.825	0.816	0.783	0.652	0.742	0.777	0.767	0.734	0.560
				5	0.795	0.815	0.806	0.776	0.654	0.736	0.769	0.758	0.725	0.558
				10	0.795	0.814	0.805	0.772	0.655	0.735	0.766	0.755	0.721	0.558
0.3	0.6	0.8	2	0	0.070	0.088	0.082	0.086	0.050	0.066	0.074	0.073	0.074	0.053
				1	0.070	0.081	0.073	0.079	0.054	0.065	0.067	0.066	0.069	0.054
				2	0.068	0.073	0.069	0.073	0.057	0.063	0.061	0.063	0.065	0.057
			4	0	0.153	0.201	0.180	0.172	0.065	0.119	0.149	0.137	0.131	0.059
				2	0.146	0.182	0.167	0.157	0.076	0.113	0.136	0.121	0.117	0.067
				4	0.149	0.178	0.165	0.148	0.091	0.112	0.131	0.119	0.112	0.071
			6	0	0.334	0.407	0.384	0.339	0.123	0.235	0.293	0.271	0.246	0.089
				3	0.331	0.386	0.366	0.314	0.155	0.228	0.274	0.252	0.223	0.104
				6	0.339	0.381	0.368	0.308	0.189	0.227	0.270	0.249	0.215	0.124
			8	0	0.576	0.606	0.597	0.550	0.356	0.422	0.473	0.464	0.407	0.195
				4	0.566	0.588	0.582	0.528	0.388	0.413	0.455	0.443	0.383	0.218
				8	0.570	0.584	0.580	0.522	0.409	0.419	0.449	0.441	0.378	0.243
			10	0	0.760	0.774	0.766	0.733	0.622	0.634	0.652	0.646	0.596	0.433
				5	0.752	0.758	0.752	0.715	0.642	0.628	0.639	0.635	0.571	0.454
				10	0.754	0.754	0.751	0.711	0.649	0.625	0.636	0.632	0.568	0.467
0.5	0.8	1.0	2	0	0.093	0.111	0.103	0.096	0.054	0.090	0.102	0.098	0.089	0.057
				1	0.090	0.102	0.097	0.090	0.054	0.087	0.094	0.091	0.085	0.056
				2	0.088	0.096	0.096	0.087	0.061	0.087	0.087	0.087	0.083	0.058
			4	0	0.266	0.311	0.287	0.264	0.114	0.239	0.279	0.262	0.234	0.112
				2	0.263	0.290	0.277	0.244	0.125	0.229	0.261	0.249	0.216	0.119
				4	0.272	0.285	0.281	0.241	0.149	0.239	0.257	0.254	0.214	0.136
			6	0	0.533	0.565	0.546	0.506	0.347	0.503	0.524	0.514	0.469	0.317
				3	0.534	0.546	0.537	0.488	0.362	0.491	0.503	0.497	0.449	0.323
				6	0.544	0.543	0.544	0.482	0.394	0.503	0.501	0.505	0.444	0.353
			8	0	0.744	0.752	0.749	0.705	0.622	0.715	0.719	0.718	0.671	0.593
				4	0.739	0.744	0.743	0.691	0.632	0.712	0.707	0.712	0.656	0.605
				8	0.744	0.744	0.746	0.690	0.655	0.715	0.705	0.713	0.656	0.623
			10	0	0.864	0.868	0.865	0.843	0.808	0.846	0.855	0.850	0.819	0.773
				5	0.864	0.865	0.862	0.838	0.809	0.846	0.848	0.846	0.809	0.781
				10	0.863	0.865	0.863	0.838	0.818	0.847	0.847	0.847	0.809	0.792

Notes: τ_1 and τ_2 denote the sample fraction at which the explosive period begins and ends; τ_3 denotes the end of the collapse regime; c_1 and c_2 denote the locally explosive and collapse magnitudes; $\sigma(r)$ denotes the volatility function, with $S(r, \sigma_1, \sigma_2, \tau_\sigma, \gamma)$ denoting a smooth transition function from σ_1 to σ_2 with midpoint τ_σ and speed γ.

DOI: 10.1111/jtsa.12784
Journal of Time Series Analysis published by John Wiley & Sons Ltd.

Table V. Local asymptotic power of nominal 0.05-level tests

					$\sigma(r) = S(r, 1, 1/6, 0.8, 30)$					$\sigma(r) = S(r, 1, 1/3, 0.8, 30)$				
τ_1	τ_2	τ_3	c_1	c_2	PSY_σ	PSY^*_σ	UPSY_σ	$s\text{PSY}$	$\bar{s}\text{PSY}$	PSY_σ	PSY^*_σ	UPSY_σ	$s\text{PSY}$	$\bar{s}\text{PSY}$
0.1	0.4	0.6	2	0	0.072	0.105	0.087	0.108	0.061	0.072	0.105	0.087	0.108	0.061
				1	0.068	0.091	0.079	0.098	0.070	0.068	0.091	0.079	0.098	0.070
				2	0.068	0.084	0.076	0.092	0.082	0.068	0.084	0.076	0.092	0.082
			4	0	0.183	0.292	0.263	0.244	0.107	0.183	0.292	0.263	0.244	0.107
				2	0.181	0.268	0.242	0.222	0.136	0.181	0.268	0.242	0.222	0.136
				4	0.209	0.263	0.251	0.214	0.176	0.209	0.263	0.251	0.214	0.176
			6	0	0.466	0.547	0.527	0.487	0.262	0.466	0.547	0.527	0.487	0.262
				3	0.469	0.524	0.512	0.464	0.306	0.469	0.524	0.512	0.464	0.306
				6	0.486	0.518	0.513	0.462	0.371	0.486	0.518	0.513	0.462	0.371
			8	0	0.720	0.754	0.743	0.705	0.538	0.720	0.754	0.743	0.705	0.538
				4	0.715	0.735	0.728	0.684	0.560	0.715	0.735	0.728	0.684	0.560
				8	0.722	0.731	0.730	0.679	0.591	0.722	0.731	0.730	0.679	0.591
			10	0	0.856	0.869	0.861	0.840	0.765	0.856	0.869	0.861	0.840	0.765
				5	0.854	0.861	0.856	0.832	0.776	0.854	0.861	0.856	0.832	0.776
				10	0.855	0.861	0.855	0.831	0.782	0.855	0.861	0.855	0.831	0.782
0.3	0.6	0.8	2	0	0.099	0.139	0.126	0.121	0.050	0.099	0.139	0.126	0.121	0.050
				1	0.100	0.126	0.120	0.111	0.066	0.098	0.126	0.119	0.111	0.065
				2	0.107	0.120	0.120	0.105	0.086	0.106	0.120	0.120	0.105	0.083
			4	0	0.334	0.403	0.383	0.336	0.117	0.334	0.403	0.383	0.336	0.117
				2	0.338	0.384	0.369	0.315	0.176	0.337	0.384	0.368	0.315	0.172
				4	0.363	0.381	0.378	0.309	0.241	0.357	0.381	0.377	0.309	0.235
			6	0	0.634	0.656	0.643	0.602	0.427	0.634	0.656	0.643	0.602	0.427
				3	0.633	0.641	0.632	0.587	0.469	0.631	0.641	0.632	0.587	0.467
				6	0.652	0.637	0.643	0.581	0.513	0.650	0.637	0.641	0.581	0.507
			8	0	0.808	0.815	0.806	0.779	0.686	0.808	0.815	0.806	0.779	0.686
				4	0.803	0.806	0.800	0.767	0.706	0.803	0.806	0.800	0.767	0.705
				8	0.806	0.804	0.800	0.764	0.717	0.805	0.804	0.800	0.763	0.716
			10	0	0.895	0.896	0.895	0.870	0.837	0.895	0.896	0.895	0.870	0.837
				5	0.890	0.893	0.889	0.866	0.844	0.890	0.893	0.889	0.866	0.844
				10	0.893	0.890	0.891	0.867	0.846	0.893	0.890	0.891	0.867	0.846
0.5	0.8	1.0	2	0	0.133	0.189	0.164	0.159	0.051	0.129	0.182	0.161	0.153	0.051
				1	0.294	0.165	0.270	0.146	0.203	0.160	0.160	0.168	0.139	0.103
				2	0.514	0.169	0.487	0.212	0.387	0.288	0.157	0.264	0.140	0.205
			4	0	0.487	0.521	0.504	0.458	0.275	0.477	0.508	0.494	0.444	0.255
				2	0.696	0.504	0.677	0.495	0.574	0.550	0.486	0.531	0.431	0.400
				4	0.809	0.697	0.797	0.661	0.722	0.661	0.485	0.643	0.473	0.542
			6	0	0.754	0.771	0.755	0.725	0.633	0.740	0.757	0.738	0.710	0.610
				3	0.877	0.786	0.869	0.777	0.808	0.785	0.744	0.771	0.702	0.699
				6	0.910	0.881	0.905	0.853	0.867	0.847	0.753	0.838	0.745	0.771
			8	0	0.888	0.894	0.887	0.868	0.819	0.879	0.885	0.878	0.862	0.809
				4	0.936	0.908	0.933	0.906	0.910	0.899	0.879	0.895	0.864	0.861
				8	0.955	0.938	0.952	0.928	0.931	0.922	0.887	0.917	0.880	0.890
			10	0	0.953	0.953	0.953	0.936	0.917	0.951	0.950	0.947	0.932	0.914
				5	0.977	0.964	0.975	0.954	0.958	0.955	0.945	0.950	0.933	0.934
				10	0.982	0.972	0.980	0.963	0.964	0.964	0.947	0.961	0.936	0.944

Notes: τ_1 and τ_2 denote the sample fraction at which the explosive period begins and ends; τ_3 denotes the end of the collapse regime; c_1 and c_2 denote the locally explosive and collapse magnitudes; $\sigma(r)$ denotes the volatility function, with $S(r, \sigma_1, \sigma_2, \tau_\sigma, \gamma)$ denoting a smooth transition function from σ_1 to σ_2 with midpoint τ_σ and speed γ.

DOI: 10.1111/jtsa.12784

Table VI. Local asymptotic power of nominal 0.05-level tests

					$\sigma(r) = S(r, 1, 3, 0.8, 30)$					$\sigma(r) = S(r, 1, 6, 0.8, 30)$				
τ_1	τ_2	τ_3	c_1	c_2	PSY_σ	PSY^*_σ	UPSY_σ	$s\text{PSY}$	$\bar{s}\text{PSY}$	PSY_σ	PSY^*_σ	UPSY_σ	$s\text{PSY}$	$\bar{s}\text{PSY}$
0.1	0.4	0.6	2	0	0.072	0.105	0.087	0.108	0.061	0.072	0.105	0.087	0.108	0.061
				1	0.068	0.091	0.079	0.098	0.070	0.068	0.091	0.079	0.098	0.070
				2	0.068	0.084	0.076	0.092	0.082	0.068	0.084	0.076	0.092	0.082
			4	0	0.183	0.292	0.263	0.244	0.107	0.183	0.292	0.263	0.244	0.107
				2	0.181	0.268	0.242	0.222	0.136	0.181	0.268	0.242	0.222	0.135
				4	0.209	0.263	0.251	0.214	0.176	0.208	0.263	0.251	0.214	0.176
			6	0	0.466	0.547	0.527	0.487	0.262	0.466	0.547	0.527	0.487	0.262
				3	0.469	0.524	0.512	0.464	0.306	0.469	0.524	0.512	0.464	0.306
				6	0.486	0.518	0.513	0.462	0.371	0.486	0.518	0.513	0.462	0.370
			8	0	0.720	0.754	0.743	0.705	0.538	0.720	0.754	0.743	0.705	0.538
				4	0.715	0.735	0.728	0.684	0.560	0.715	0.735	0.728	0.684	0.559
				8	0.722	0.731	0.730	0.679	0.591	0.722	0.731	0.730	0.679	0.591
			10	0	0.856	0.869	0.861	0.840	0.765	0.856	0.869	0.861	0.840	0.765
				5	0.854	0.861	0.856	0.832	0.776	0.854	0.861	0.856	0.832	0.776
				10	0.855	0.861	0.855	0.831	0.782	0.855	0.861	0.855	0.831	0.782
0.3	0.6	0.8	2	0	0.099	0.139	0.126	0.121	0.050	0.099	0.139	0.126	0.121	0.050
				1	0.096	0.127	0.118	0.111	0.059	0.096	0.127	0.117	0.112	0.057
				2	0.098	0.120	0.115	0.107	0.072	0.096	0.121	0.114	0.108	0.068
			4	0	0.333	0.401	0.382	0.336	0.117	0.333	0.401	0.382	0.336	0.117
				2	0.333	0.384	0.367	0.317	0.150	0.331	0.385	0.365	0.317	0.141
				4	0.340	0.379	0.369	0.309	0.201	0.337	0.379	0.366	0.309	0.178
			6	0	0.634	0.655	0.642	0.602	0.426	0.633	0.655	0.642	0.601	0.424
				3	0.626	0.641	0.630	0.587	0.446	0.626	0.640	0.629	0.588	0.438
				6	0.637	0.637	0.634	0.580	0.473	0.633	0.635	0.632	0.581	0.457
			8	0	0.808	0.814	0.806	0.778	0.686	0.808	0.814	0.806	0.777	0.685
				4	0.803	0.806	0.799	0.768	0.692	0.801	0.805	0.799	0.767	0.685
				8	0.805	0.804	0.800	0.764	0.704	0.804	0.804	0.798	0.763	0.700
			10	0	0.895	0.896	0.894	0.870	0.837	0.895	0.896	0.893	0.870	0.837
				5	0.889	0.893	0.888	0.867	0.840	0.889	0.892	0.887	0.866	0.837
				10	0.892	0.890	0.889	0.865	0.843	0.891	0.890	0.887	0.865	0.841
0.5	0.8	1.0	2	0	0.106	0.146	0.129	0.126	0.047	0.098	0.131	0.120	0.116	0.046
				1	0.102	0.137	0.125	0.119	0.047	0.096	0.126	0.118	0.113	0.045
				2	0.100	0.134	0.122	0.117	0.049	0.094	0.124	0.115	0.109	0.043
			4	0	0.366	0.414	0.395	0.351	0.135	0.313	0.376	0.353	0.318	0.111
				2	0.356	0.401	0.383	0.338	0.136	0.305	0.365	0.344	0.308	0.108
				4	0.352	0.396	0.378	0.334	0.141	0.303	0.359	0.339	0.305	0.110
			6	0	0.640	0.668	0.651	0.605	0.433	0.585	0.622	0.599	0.548	0.355
				3	0.631	0.658	0.641	0.594	0.431	0.580	0.613	0.591	0.541	0.353
				6	0.629	0.654	0.638	0.592	0.433	0.578	0.611	0.588	0.538	0.353
			8	0	0.822	0.833	0.822	0.791	0.708	0.772	0.797	0.776	0.740	0.636
				4	0.817	0.827	0.815	0.782	0.708	0.768	0.791	0.770	0.733	0.633
				8	0.816	0.825	0.813	0.781	0.709	0.765	0.785	0.767	0.731	0.633
			10	0	0.908	0.909	0.905	0.890	0.847	0.885	0.888	0.884	0.865	0.801
				5	0.905	0.907	0.903	0.888	0.846	0.883	0.887	0.881	0.860	0.800
				10	0.904	0.906	0.903	0.888	0.848	0.882	0.886	0.881	0.859	0.800

Notes: τ_1 and τ_2 denote the sample fraction at which the explosive period begins and ends; τ_3 denotes the end of the collapse regime; c_1 and c_2 denote the locally explosive and collapse magnitudes; $\sigma(r)$ denotes the volatility function, with $S(r, \sigma_1, \sigma_2, \tau_\sigma, \gamma)$ denoting a smooth transition function from σ_1 to σ_2 with midpoint τ_σ and speed γ.

(late small upward shift in volatility) and $S(r, 1, 6, 0.8, 30)$ (late large upward shift in volatility) the results are once more similar to the homoskedastic case, where PSY_σ^* represents the best performing test.

On the basis of our asymptotic simulations, what is clear is that for any given DGP, the better performing of the new PSY_σ and PSY_σ^* tests dominates the better performing of the sPSY and $\bar{s}$PSY tests, with gains of up to about 0.13 (outside of very low power cases where very small losses of up to 0.01 are observed). Between PSY_σ and PSY_σ^*, there is no clear winner unless we are prepared to take a stance on a particular form of bubble and/or volatility DGP setting being present in the data. This might seem counterintuitive since PSY_σ involves fitting what might be considered a redundant intercept term. In reality though, these matters are not easily resolved because PSY_σ and PSY_σ^* are based on the double-supremum of $\text{DF}_\sigma(\lambda_1, \lambda_2)$ and $\text{DF}_\sigma^*(\lambda_1, \lambda_2)$, and the magnitude and locations of the double-supremum involves a very complex (essentially intractable) interaction of the values of the parameters $c_1, c_2, \tau_1, \tau_2, \tau_3$ and the volatility path $\sigma(r)$. Given that each test offers power gains over the other for some areas of the parameter constellation considered, and the particular parameter settings would be unknown to a practitioner, it makes sense to consider a procedure that aims to harness the higher power that is available from PSY_σ and PSY_σ^* in any particular DGP setting by employing a *union of rejections* strategy. We detail this approach in the next section.

5. A UNION OF REJECTIONS BASED STRATEGY

We now consider a union of rejections testing strategy based on inference from both PSY_σ and PSY_σ^*, in line with the initial work on this approach in Harvey *et al.* (2019) in the context of left-tailed unit root testing under uncertainty regarding the presence or otherwise of a linear trend. Specifically, denoting the asymptotic ξ level null critical value of PSY_σ by cv_ξ (from the $G_{0,0,\sigma}$ distribution) and that of PSY_σ^* by cv_ξ^* (from the $G_{0,0,\sigma}^*$ distribution), a union of rejections strategy can be written as the decision rule

$$\text{Reject } H_0 \text{ if } \left\{\text{PSY}_\sigma > \psi_\xi \text{cv}_\xi \text{ or } \text{PSY}_\sigma^* > \psi_\xi \text{cv}_\xi^*\right\},$$

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

$$\text{UPSY}_\sigma := \max\left(\text{PSY}_\sigma, \frac{\text{cv}_\xi}{\text{cv}_\xi^*}\text{PSY}_\sigma^*\right),$$

the decision rule is then equivalent to

$$\text{Reject } H_0 \text{ if } \text{UPSY}_\sigma > \psi_\xi \text{cv}_\xi.$$

An application of the continuous mapping theorem along with the results in Theorem 2 shows that

$$\text{UPSY}_\sigma \Rightarrow \max\left(G_{c_1,c_2,\sigma}, \frac{\text{cv}_\xi}{\text{cv}_\xi^*}G_{c_1,c_2,\sigma}^*\right).$$

Note that there is no need to explicitly calculate the scaling constant ψ_ξ as, for a given value of $\text{cv}_\xi^s/\text{cv}_\xi$, all we require is the critical value $\text{cv}_\xi^U := \psi_\xi \text{cv}_\xi$ which can be obtained directly from the limiting null distribution of UPSY_σ; that is, $\max(G_{0,0,\sigma}, (\text{cv}_\xi/\text{cv}_\xi^*)G_{0,0,\sigma}^*)$. Asymptotic and finite sample critical values for cv_ξ^U are provided in Table I.

The asymptotic local power results for UPSY_σ are also given in Tables II–VI. Throughout, we see that the local powers of UPSY_σ track very closely the better power that is available from PSY_σ and PSY_σ^* individually. Indeed, across the whole of Tables II–VI, the mean power loss for UPSY_σ compared to the better of PSY_σ and PSY_σ^* is

DOI: 10.1111/jtsa.12784

only 0.010 with a standard deviation of less than 0.008, and the largest power deficit is only 0.033. Given that the performance of the union of rejections strategy is so close to that of the best of PSY_σ and PSY^*_σ, it is unsurprising to see that UPSY_σ dominates the better performing of the sPSY and $\bar{s}$PSY tests across almost all parameter settings (the only exceptions being cases where all tests have very low power).

6. FINITE SAMPLE SIMULATIONS

We now turn to an examination of the finite sample properties of the tests. Our simulations are based on the model in (1)- (3) with $T = 200$. We set $\mu = 0$ (without loss of generality) and $u_1 = \epsilon_1$, where $\epsilon_t = \sigma_t z_t$ with the z_t generated as NIID(0, 1) random variates. Here the limit volatility functions $\sigma(r)$ are discretised to $\sigma_t(t/T)$ over 200 steps. In the non-parametric variance estimator $\hat{\sigma}^2_t$ used in constructing x_t for PSY_σ and PSY^*_σ we again employ the Gaussian kernel for $K(\cdot)$ and set the bandwidth to $h = 0.1T^{-0.25}$. We use finite sample critical values for all tests, i.e. those in Table I for PSY_σ, PSY^*_σ and UPSY_σ, and critical values simulated in the same way for sPSY and $\bar{s}PSY$. Table S1 reports 0.05-level finite sample sizes and powers over the same set of volatility patterns as for the asymptotic power analysis of Tables II–VI; for the powers we consider the same constellation of bubble/collapse parameter settings as Tables II–VI.

First we note that throughout Table S1, the sizes of all tests are close to the nominal level across the different volatility patterns. Some very modest over-size is observed in the case of PSY_σ, up to 0.058, but even this feature is largely absent when considering the union of rejections strategy UPSY_σ, which has a maximum size of 0.053. The sPSY and $\bar{s}$PSY sizes are invariant to σ_t here, since the statistics are based on $\text{sign}(\Delta y_t) = \text{sign}(\epsilon_t) = \text{sign}(z_t)$ under the null.

Turning to finite sample power, it is clear from a comparison of Tables II–VI and S1 that, in the main, the finite sample rejection frequencies bear a very close resemblance to the corresponding local asymptotic results. There are some individual settings for which the correspondence deviates from this general pattern, but these are rare. Hence, the overall patterns of results, the rankings of the tests for different volatility patterns and bubble/collapse timings, and the magnitudes of the relative power differences are largely the same as in the local asymptotic case. In particular, the powers of UPSY_σ are very close to the better power that is available from PSY_σ and PSY^*_σ, with the mean power loss for UPSY_σ compared to the better of PSY_σ and PSY^*_σ only 0.007 across the Table S1 results, with a SD of less than 0.008, and a largest power deficit of 0.037. In these finite sample results, UPSY_σ power continues to dominate the better performing of the sPSY and $\bar{s}$PSY tests across almost all parameter settings, as with the corresponding asymptotic local power results in Tables II–VI.

In addition to these finite sample results that use i.i.d errors, we also investigate finite sample size robustness under departures from this assumption. First, we consider the case where z_t is a conditionally heteroskedastic GARCH(1,1) process, with $z_t = \sqrt{h_t}\eta_t$, where $\eta_t \sim \text{NIID}(0, 1)$, and $h_t = 0.1 + 0.1z^2_{t-1} + 0.8h_{t-1}$ (with $h_0 = z_0 = 0$). Sizes are simulated for the same volatility functions as considered in Table S1, using simulated finite sample critical values obtained under conditional and unconditional homoskedasticity. The results are presented in Table S2 for $T = 200$ and $T = 400$. Reliable finite sample size is generally observed for the new procedures across the different DGPs, particularly for PSY^*_σ. Some modest over-size is displayed, but this reduces with the sample size, as would be expected. As in Table S1, the sPSY and $\bar{s}$PSY sizes are exactly 0.05 since here $\text{sign}(\Delta y_t) = \text{sign}(\eta_t)$.

The second departure from i.i.d errors that we consider is serial correlation in ϵ_t, using the AR(1) specification $\epsilon_t = \theta\epsilon_{t-1} + \sigma_t z_t$. In line with the discussion in Remark 3.3, we augment the subsample regression (6), and its without-intercept equivalent respectively, with one lagged difference term Δx_{t-1} and construct x_t as in (5) continuing to use the first-differences based estimator given in (4).[5] The comparator sign-based tests are also adjusted for serial correlation using the recursive prewhitening method outlined in HLZ. Finite sample critical values for the lag-augmented/prewhitened tests are simulated using NIID(0, 1) errors. The results are presented in Table S3 for $T = 200$ and $T = 400$. We see that the new procedures generally display decent finite sample size control under these serial correlation settings. Some modest over-size is seen, particularly when $\theta = 0.4$ and

[5] We also considered the analogous tests based on the residual-based alternative to the first-differences-based estimator discussed in Remark 3.3, but found these to display inferior finite sample performance to the reported tests.

DOI: 10.1111/jtsa.12784 Journal of Time Series Analysis published by John Wiley & Sons Ltd.

$T = 200$, although the over-size reduces for $T = 400$ as expected. The sign-based comparator procedures also display some over-size, with the degree broadly in line with that observed for the new tests.

7. CONCLUSIONS

In this article, we have proposed modified versions of the seminal bubble detection methods developed in Phillips *et al.* (2015) that work effectively in the presence of non-stationary volatility. The modification purges unconditional heteroskedasticity from the data under the null by re-scaling the first differences of the data by a kernel estimate of volatility and then recumulating. The procedures developed in Phillips *et al.* (2015) are then applied to this recumulated data, rather than the original data. Simulations indicate that our new tests perform well relative to extant bubble detection tests that allow for non-stationary volatility. A union of rejections procedure based on versions of our statistics from regressions with and without an intercept was found to perform especially well across a wide range of simulation settings.

ACKNOWLEDGEMENTS

We are grateful to the Co-Editor, Morten Nielsen, and two anonymous referees for their helpful and constructive comments on an earlier version of this article.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

REFERENCES

Beare B. 2018. Unit root testing with unstable volatility. *Journal of Time Series Analysis* **39**:816–835.

Boswijk HP, Zu Y. 2018. Adaptive wild bootstrap tests for a unit root with non-stationary volatility. *The Econometrics Journal* **21**:87–113.

Boswijk HP, Zu Y. 2022. Adaptive testing for cointegration with nonstationary volatility. *Journal of Business & Economic Statistics* **40**:744–755.

Brillinger DR. 2001. *Time Series: Data Analysis and Theory* Society for Industrial and Applied Mathematics, Philadelphia.

Calvo-Gonzalez O, Shankar R, Trezzi R. 2010. Are commodity prices more volatile now? World Bank Policy Research Working Paper No. 5460.

Cavaliere G, Nielsen MØ, Taylor AMR. 2022. Adaptive inference in heteroskedastic fractional time series models. *Journal of Business and Economic Statistics* **40**:50–65.

Chang Y, Park JY. 2002. On the asymptotics of ADF tests for unit roots. *Econometric Reviews* **21**:431–447.

Goncalves S, Kilian L. 2007. Asymptotic and bootstrap inference for $AR(\infty)$ processes with conditional heteroskedasticity. *Econometric Reviews* **26**:609–641.

Hall P, Heyde CC. 1980. *Martingale Limit Theory and Its Applications* Academic Press, New York.

Harvey DI, Leybourne SJ, Sollis R, Taylor AMR. 2016. Tests for explosive financial bubbles in the presence of non-stationary volatility. *Journal of Empirical Finance* **38**:548–574.

Harvey DI, Leybourne SJ, Zu Y. 2019. Testing explosive bubbles with time-varying volatility. *Econometric Reviews* **38**:1131–1151.

Harvey DI, Leybourne SJ, Zu Y. 2020. Sign-based unit root tests for explosive financial bubbles in the presence of nonstationary volatility. *Econometric Theory* **36**:122–169.

McMillan DG, Wohar ME. 2011. Structural breaks in volatility: the case of UK sector returns. *Applied Financial Economics* **21**:578–595.

Novak J. 2014. *Three lectures on free probability*. In *Random Matrix Theory, Interacting Particle Systems, and Integrable Systems*, Vol. 65, Deift P, Forrester P (eds.). Cambridge University Press, New York; 309–383.

Phillips PCB, Wu Y, Yu J. 2011. Explosive behavior in the 1990s Nasdaq: when did the exuberance escalate asset values? *International Economic Review* **52**:201–226.

Phillips PCB, Shi S-P, Yu J. 2015. Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P500. *International Economic Review* **56**:1043–1078.

Rapach DE, Strauss JK, Wohar ME. 2008. *Forecasting stock return volatility in the presence of structural breaks*. In *Forecasting in the Presence of Structural Breaks and Model Uncertainty*. Frontiers of Economics and Globalization, Vol. 3, Rapach DE, Wohar ME (eds.). Emerald, Bingley, UK; 381–416.

Vivian A, Wohar ME. 2012. Commodity volatility breaks. *Journal of International Financial Markets, Institutions and Money* **22**:395–422.
Xu K-L, Phillips PCB. 2008. Adaptive estimation of autoregressive models with time-varying variance. *Journal of Econometrics* **142**:265–280.
Zivot E, Andrews DWK. 1992. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business and Economic Statistics* **10**:251–270.

APPENDIX A

A.1. Lemma 1 and Proof

We begin by stating and proving a preparatory lemma that will subsequently be required for the proof of Theorem 1.

Lemma 1. Let the conditions of Theorem 1 hold. Then,

$$\sum_{j=2}^{T}\left(\hat{\sigma}_j^2-\sigma_j^2\right)^2 = O_p(h^{-1}) + o_p(T).$$

Proof. Defining $\tilde{\sigma}_t^2 := \sum_{j=2}^{T} w_{t,j}(\sigma_j z_j)^2$ and $\overline{\sigma}_t^2 := \sum_{j=2}^{T} w_{t,j}\sigma_j^2$, notice first that

$$\sum_{j=2}^{T}\left(\hat{\sigma}_j^2-\sigma_j^2\right)^2 \leqslant 3\left(\sum_{j=2}^{T}\left(\hat{\sigma}_j^2-\tilde{\sigma}_j^2\right)^2+\sum_{j=2}^{T}\left(\tilde{\sigma}_j^2-\overline{\sigma}_j^2\right)^2+\sum_{j=2}^{T}\left(\overline{\sigma}_j^2-\sigma_j^2\right)^2\right).$$

Using the same argument as used by Harvey *et al.* (2019) in deriving equation (9) in the proof of their Theorem 1, we obtain that $\sum_{j=2}^{T}(\hat{\sigma}_j^2-\tilde{\sigma}_j^2)^2 = O_p(T^{-1})$. By Lemma S.1 (b) of CNT, $\sum_{j=2}^{T}(\tilde{\sigma}_j^2-\overline{\sigma}_j^2)^2 = O_p(h^{-1})$. Using the argument as in the proof of Lemma S.1 (a) of CNT and lemma A(l) of Xu and Phillips (2008), we have that $\sum_{j=2}^{T}\left(\overline{\sigma}_j^2-\sigma_j^2\right)^2 = o_p(T)$. In summary, the first term is dominated by the second and third terms, while the relative magnitude of the second and the third terms is indeterminate, and so $\sum_{j=2}^{T}\left(\hat{\sigma}_j^2-\sigma_j^2\right)^2 = O_p(h^{-1}) + o_p(T)$. ■

A.2. Proof of Theorem 1

Under H_1 we can write

$$T^{-1/2}x_{\lfloor Tr\rfloor} = T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{u_j-u_{j-1}}{\hat{\sigma}_j} = T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\hat{\sigma}_j} + T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{\sigma_j z_j}{\hat{\sigma}_j}.$$

To establish the limit of $T^{-1/2}x_{\lfloor Tr\rfloor}$ stated in (8), we will first show that for $r > 0$,

$$\sup_{0\leqslant r\leqslant 1}\left|T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\hat{\sigma}_j} - T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\sigma_j}\right| = o_p(1), \quad (A.1)$$

and

$$\sup_{0\leqslant r\leqslant 1}\left|T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{\sigma_j z_j}{\hat{\sigma}_j} - T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}z_j\right| = o_p(1), \quad (A.2)$$

and then derive

$$T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\sigma_j}+T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}z_j \Rightarrow \begin{cases} W(r) & r\leqslant\tau_1 \\ W(r)+c_1\int_{\tau_1}^{r}\frac{V_1(s)}{\sigma(s)}\mathrm{d}s & \tau_1<r\leqslant\tau_2 \\ W(r)+c_1\int_{\tau_1}^{\tau_2}\frac{V_1(s)}{\sigma(s)}\mathrm{d}s-c_2\int_{\tau_2}^{r}\frac{V_2(s)}{\sigma(s)}\mathrm{d}s & \tau_2<r\leqslant\tau_3 \\ W(r)+c_1\int_{\tau_1}^{\tau_2}\frac{V_1(s)}{\sigma(s)}\mathrm{d}s-c_2\int_{\tau_2}^{\tau_3}\frac{V_2(s)}{\sigma(s)}\mathrm{d}s & r>\tau_3 \end{cases}. \quad \text{(A.3)}$$

From (A.1) to (A.3), (8) then follows.

For (A.1), we only demonstrate the result when $r>\tau_3$; for r in other regimes the results can be shown in a similar way and, hence, are omitted. Notice that, from the definition of ρ_j,

$$\begin{aligned}\sup_{r>\tau_3}&\left|T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\hat\sigma_j}-T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\sigma_j}\right| \\ &\leqslant\left|c_1T^{-3/2}\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}u_{j-1}\left(\frac{1}{\hat\sigma_j}-\frac{1}{\sigma_j}\right)\right|+\left|c_2T^{-3/2}\sum_{j=\lfloor T\tau_2\rfloor+1}^{\lfloor T\tau_3\rfloor}u_{j-1}\left(\frac{1}{\hat\sigma_j}-\frac{1}{\sigma_j}\right)\right| \\ &=:A_1+A_2.\end{aligned}$$

Using the Cauchy–Schwarz inequality,

$$\begin{aligned}A_1&\leqslant\left|c_1T^{-3/2}\left(\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}u_{j-1}^2\right)^{1/2}\left(\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}\left(\frac{1}{\hat\sigma_j}-\frac{1}{\sigma_j}\right)^2\right)^{1/2}\right| \\ &\leqslant c_1T^{-\frac{3}{2}}\left(\min_{\lfloor T\tau_1\rfloor+1\leqslant j\leqslant\lfloor T\tau_2\rfloor}|\hat\sigma_j\sigma_j(\hat\sigma_j+\sigma_j)|\right)^{-1}\left(\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}u_{j-1}^2\right)^{\frac{1}{2}}\left(\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}(\hat\sigma_j^2-\sigma_j^2)^2\right)^{\frac{1}{2}}.\end{aligned}$$

Using the argument in the proof of Lemma A(h) and A(j) in Xu and Phillips (2008), we have that $(\min_{\lfloor T\tau_1\rfloor+1\leqslant j\leqslant\lfloor T\tau_2\rfloor}|\hat\sigma_j\sigma_j(\hat\sigma_j+\sigma_j)|)^{-1}=O_p(1)$. It is also straightforwardly seen that $\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}u_{j-1}^2=O_p(T^2)$. Using Lemma 1 we have $\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}\left(\hat\sigma_j^2-\sigma_j^2\right)^2=O_p(h^{-1})+o_p(T)$. In total, $A_1=o_p(1)$. Similarly we also have $A_2=o_p(1)$, and so (A.1) is verified.

To establish (A.2), consider the decomposition

$$\begin{aligned}&T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\left(\frac{\sigma_jz_j}{\hat\sigma_j}-z_j\right) \\ &=T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\left(\frac{\sigma_jz_j}{\hat\sigma_j}-\frac{\sigma_jz_j}{\tilde\sigma_j}\right)+T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\left(\frac{\sigma_jz_j}{\tilde\sigma_j}-\frac{\sigma_jz_j}{\overline{\sigma}_j}\right)+T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\left(\frac{\sigma_jz_j}{\overline{\sigma}_j}-z_j\right) \\ &=:B_1+B_2+B_3.\end{aligned}$$

For the first term, B_1, we have that

$$|B_1|\leqslant T^{-1/2}\max_{1\leqslant j\leqslant\lfloor Tr\rfloor}\left|\frac{\sigma_j}{\hat\sigma_j\tilde\sigma_j(\tilde\sigma_j+\hat\sigma_j)}\right|\left(\sum_{j=2}^{\lfloor Tr\rfloor}(\tilde\sigma_j^2-\hat\sigma_j^2)^2\right)^{1/2}\left(\sum_{j=2}^{\lfloor Tr\rfloor}z_j^2\right)^{1/2}. \quad \text{(A.4)}$$

From the proof of Lemma 1, we have that $\sum_{j=2}^{\lfloor Tr \rfloor}\left(\tilde{\sigma}_j^2 - \hat{\sigma}_j^2\right)^2 = O_p(T^{-1})$. It is also straightforwardly seen that $\sum_{j=2}^{\lfloor Tr \rfloor} z_t^2 = O_p(T)$ and $\max_{2\leqslant j\leqslant \lfloor Tr \rfloor}\left|\frac{\sigma_j}{\hat{\sigma}_j\tilde{\sigma}_j(\tilde{\sigma}_j+\hat{\sigma}_j)}\right| = O_p(1)$. It then follows that $B_1 = o_p(1)$ for all $0 \leqslant r \leqslant 1$. This result can be strengthened to be uniform over the same interval by noticing that the magnitude of the right-hand side of (A.4) is non-decreasing in r, and so $\sup_{0\leqslant r\leqslant 1}|B_1| = o_p(1)$.

Turning next to B_2, using the equality $p^{-1} - q^{-1} = (q-p)q^{-2} + (q-p)^2 p^{-1} q^{-2}$, consider the following decomposition

$$\begin{aligned} B_2 &= T^{-1/2}\sum_{j=2}^{\lfloor Tr \rfloor}\sigma_j z_j\left(1/\tilde{\sigma}_j^2 - 1/\overline{\sigma}_j^2\right) \\ &= T^{-1/2}\sum_{j=2}^{\lfloor Tr \rfloor}\sigma_j z_j\left(\overline{\sigma}_j^2 - \tilde{\sigma}_j^2\right)\overline{\sigma}_j^{-4} + T^{-1/2}\sum_{j=2}^{\lfloor Tr \rfloor}\sigma_j z_j\left(\overline{\sigma}_j^2 - \tilde{\sigma}_j^2\right)^2\overline{\sigma}_j^{-4}\tilde{\sigma}_j^{-2} \\ &=: B_{21} + B_{22}. \end{aligned}$$

Consider first B_{21}. We evaluate

$$\begin{aligned} E|B_{21}|^2 &= T^{-1}E\left(\sum_{j=2}^{\lfloor Tr \rfloor}\sigma_j z_j\left(\sum_{i=2}^{T} w_{j,i}\sigma_i^2\left(z_i^2 - 1\right)\right)\overline{\sigma}_j^{-4}\right)^2 \\ &= T^{-1}E\left(\sum_{j=2}^{\lfloor Tr \rfloor}\sum_{i=2}^{T} w_{j,i}\sigma_j\sigma_i^2\overline{\sigma}_j^{-4} z_j\left(z_i^2 - 1\right)\right)^2 \\ &= T^{-1}E\sum_{j,j'=2}^{\lfloor Tr \rfloor}\sum_{i,i'=2}^{T} w_{j,i}w_{j',i'}\sigma_j\sigma_{j'}\sigma_i^2\sigma_{i'}^2\overline{\sigma}_j^{-4}\overline{\sigma}_{j'}^{-4}\left(z_j z_{j'}\left(z_i^2 - 1\right)\left(z_{i'}^2 - 1\right)\right) \\ &\leqslant T^{-1}\sum_{j,j'=2}^{T}\sum_{i,i'=2}^{T} w_{j,i}w_{j',i'}\sigma_j\sigma_{j'}\sigma_i^2\sigma_{i'}^2\overline{\sigma}_j^{-4}\overline{\sigma}_{j'}^{-4}\left|E\left(z_j z_{j'}\left(z_i^2 - 1\right)\left(z_{i'}^2 - 1\right)\right)\right|. \end{aligned}$$

Using Assumption 1 and applying the standard formula representing joint moments by summation of different products of joint cumulants (see section 1.3 of Novak, 2014 and section 2.3 of Brillinger, 2001), we can show that

$$\sum_{j,j'=2}^{T}\sum_{i,i'=2}^{T}|E(z_j z_{j'}(z_i^2 - 1)(z_{i'}^2 - 1))| = O(T^2).$$

It then follows that $EB_{21}^2 = O(1/(Th^2)) = o(1)$. Notice that the above derivation also goes through for $\sup_{0\leqslant r\leqslant 1}|B_{21}|^2$ and it follows that $\sup_{0\leqslant r\leqslant 1} B_{21} = O_p(1/(Th^2)) = o_p(1)$.

For B_{22}, using the Cauchy–Schwarz inequality for the sum we have

$$\left|T^{-1/2}\sum_{j=2}^{\lfloor Tr \rfloor}\sigma_j z_j\left(\overline{\sigma}_j^2 - \tilde{\sigma}_j^2\right)^2\overline{\sigma}_j^{-4}\tilde{\sigma}_j^{-2}\right| \leqslant \mathbb{C}\left(\frac{1}{T}\sum_{j=2}^{\lfloor Tr \rfloor}(\sigma_j z_j)^2\right)^{1/2}\left(\sum_{j=2}^{\lfloor Tr \rfloor}\left(\overline{\sigma}_j^2 - \tilde{\sigma}_j^2\right)^4\right)^{1/2}.$$

It is straightforwardly seen that $\sup_{0\leqslant r\leqslant 1} T^{-1}\sum_{j=2}^{\lfloor \tau T \rfloor}(\sigma_j z_j)^2 = O_p(1)$. Notice that $\sup_{0\leqslant r\leqslant 1}\sum_{j=2}^{\lfloor Tr \rfloor}\left(\overline{\sigma}_j^2 - \tilde{\sigma}_j^2\right)^4 = \sum_{j=2}^{T}\left(\overline{\sigma}_j^2 - \tilde{\sigma}_j^2\right)^4$, which is $O_p\left(\frac{1}{Th^2}\right)$ by Lemma S.1 (e) accompanying CNT. Therefore we have that $\sup_{0\leqslant r\leqslant 1}|B_{22}| = o_p(1)$. In total, we therefore have that $\sup_{0\leqslant r\leqslant 1}|B_2| = o_p(1)$.

Turning finally to B_3, using the fact that the σ_j and $\overline{\sigma}_j$ are both deterministic, we have that

$$\begin{aligned} E|B_3|^2 &= E\left(T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{1}{\overline{\sigma}_j(\sigma_j+\overline{\sigma}_j)}\left(\overline{\sigma}_j^2-\sigma_j^2\right)z_j\right)^2 \\ &= T^{-1}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{1}{\overline{\sigma}_j^2(\sigma_j+\overline{\sigma}_j)^2}\left(\overline{\sigma}_j^2-\sigma_j^2\right)^2 E z_j^2 \\ &\leqslant \max_{2\leqslant j\leqslant \lfloor Tr\rfloor}\frac{1}{\overline{\sigma}_j^2(\sigma_j+\overline{\sigma}_j)^2}T^{-1}\sum_{j=2}^{\lfloor Tr\rfloor}\left(\overline{\sigma}_j^2-\sigma_j^2\right)^2. \end{aligned}$$

From the proof of Lemma 1, we have that $T^{-1}\sum_{j=2}^{\lfloor Tr\rfloor}\left(\overline{\sigma}_j^2-\sigma_j^2\right)^2 = o(1)$ and so $\max_j \frac{1}{\overline{\sigma}_j^2(\sigma_j+\overline{\sigma}_j)^2}$ is clearly $O(1)$ and hence $E|B_3|^2 = o(1)$. It follows from the Markov inequality that $B_3 = o_p(1)$ and the result is also uniform over $0 \leqslant r \leqslant 1$, and so (A.2) is verified.

We next establish (A.3). First, by a standard MDS FCLT (see, e.g., Hall and Heyde, 1980), we have that $T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor} z_j \Rightarrow W(r)$, $r \in [0,1]$, where $W(r)$ is a standard Brownian motion. For $T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\sigma_j}$, the limit will depend on the regime r lies in. Again here we only give the derivation of the result when $r > \tau_3$; the derivation of the results for r in the other regimes is similar and therefore omitted. When $r > \tau_3$, by approximation of the Riemann integral,

$$\begin{aligned} T^{-1/2}\sum_{j=2}^{\lfloor Tr\rfloor}\frac{(\rho_j-1)u_{j-1}}{\sigma_j} &= c_1 T^{-\frac{3}{2}}\sum_{j=\lfloor T\tau_1\rfloor+1}^{\lfloor T\tau_2\rfloor}\frac{u_{j-1}}{\sigma_j} - c_2 T^{-\frac{3}{2}}\sum_{j=\lfloor T\tau_2\rfloor+1}^{\lfloor T\tau_3\rfloor}\frac{u_{j-1}}{\sigma_j} \\ &\Rightarrow c_1\int_{\tau_1}^{\tau_2}\frac{V_1(s)}{\sigma(s)}\mathrm{d}s - c_2\int_{\tau_2}^{\tau_3}\frac{V_2(s)}{\sigma(s)}\mathrm{d}s. \end{aligned}$$

A.3. Proof of Theorem 2

In the following, we detail the proof for the without-intercept version PSY_σ^*. The with-intercept version result for PSY_σ can be derived analogously and, hence, is omitted.

Using Theorem 1, the stated result for the without-intercept version follows if we can 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 $\hat{\phi}(\lambda_1,\lambda_2)$, we have the following expansion,

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

From (8) it is easily seen that $\sum_{t=\lfloor\lambda_1 T\rfloor+1}^{\lfloor\lambda_2 T\rfloor}\Delta x_t x_{t-1} = O_p(T)$ and $\sum_{t=\lfloor\lambda_1 T\rfloor+1}^{\lfloor\lambda_2 T\rfloor}x_{t-1}^2 = O_p(T^2)$, and we therefore have that

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

DOI: 10.1111/jtsa.12784

Notice also that

$$\sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} (\Delta x_t)^2 = \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} \left(\frac{\Delta u_t}{\hat{\sigma}_t}\right)^2.$$

Using the same argument as we used in deriving (A.1), we therefore have that

$$\frac{1}{\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 1}\left| \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} \left(\frac{\Delta u_t}{\hat{\sigma}_t}\right)^2 - \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} \left(\frac{\Delta u_t}{\sigma_t}\right)^2 \right| = o_p(1).$$

Then, by the definition of u_t, we have that

$$\Delta u_t = \begin{cases} \varepsilon_t & t \leqslant \lfloor \tau_1 T \rfloor \\ (c_1/T)u_{t-1} + \varepsilon_t & \lfloor \tau_1 T \rfloor < t \leqslant \lfloor \tau_2 T \rfloor \\ (-c_2/T)u_{t-1} + \varepsilon_t & \lfloor \tau_2 T \rfloor < t \leqslant \lfloor \tau_3 T \rfloor \\ \varepsilon_t & t > \lfloor \tau_3 T \rfloor \end{cases}.$$

It is then straightforward to show that $(\lfloor \lambda_2 T \rfloor - \lfloor \lambda_1 T \rfloor - 1)^{-1} \sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} \left(\frac{\Delta u_t}{\sigma_t}\right)^2 \xrightarrow{p} 1$. Consequently,

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

We are now in a position to derive the distribution of the DF statistic over the (λ_1, λ_2) sub-sample. To that end,

$$\begin{aligned} \mathrm{DF}^*_\sigma(\lambda_1, \lambda_2) &= \frac{T^{-1}x^2_{\lfloor \lambda_2 T \rfloor} - T^{-1}x^2_{\lfloor \lambda_1 T \rfloor} - T^{-1}\sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor}(\Delta x_t)^2}{2\sqrt{\hat{s}^2(\lambda_1, \lambda_2)T^{-2}\sum_{t=\lfloor \lambda_1 T \rfloor+1}^{\lfloor \lambda_2 T \rfloor} x^2_{t-1}}} \\ &\Rightarrow \frac{\mathcal{X}(\lambda_2)^2 - \mathcal{X}(\lambda_1)^2 - (\lambda_2 - \lambda_1)}{2\sqrt{\int_{\lambda_1}^{\lambda_2} \mathcal{X}(r)^2 \mathrm{d}r}}. \end{aligned} \tag{A.5}$$

The large sample result in (A.5) holds formally only for fixed λ_1, λ_2. However, following the same approach (which is based on the proof strategy adopted by Zivot and Andrews, 1992, to prove their Theorem 1) as that used to establish equation (A.6) on p. 1072 in the proof of Theorem 1 in PSY (pp. 1072–1075), the stated result for the limiting null distribution of the PSY^*_σ statistic can be shown to follow by means of the Continuous Mapping Theorem from the fixed λ_1, λ_2 representation in (A.5).