OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 76, 3 (2014) 0305–9049
doi: 10.1111/obes.12026

Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behaviour*

PETER C. B. PHILLIPS†, SHUPING SHI‡ and JUN YU§

†*Cowles Foundation, Yale University, University of Auckland, University of Southampton & Singapore Management University, New Haven, CT, USA (e-mail: peter.phillips@yale.edu)*
‡*Research School of Finance, Actuarial Studies and Applied Statistics, The Australian National University, ACT, 0200, Australia (e-mail: shuping.shi@anu.edu.au)*
§*Sim Kee Boon Institute for Financial Economics, School of Economics and Lee Kong Chian School of Business, Singapore Management University, Singapore 178903, (e-mail: yujun@smu.edu.sg)*

Abstract

This article aims to provide some empirical guidelines for the practical implementation of right-tailed unit root tests, focusing on the recursive right-tailed ADF test of Phillips *et al.* (2011b). We analyze and compare the limit theory of the recursive test under different hypotheses and model specifications. The size and power properties of the test under various scenarios are examined and some recommendations for empirical practice are given. Some new results on the consistent estimation of localizing drift exponents are obtained, which are useful in assessing model specification. Empirical applications to stock markets illustrate these specification issues and reveal their practical importance in testing.

I. Introduction

In left-tailed unit root testing, results are often sensitive to model formulation. In effect, the maintained hypothesis or *technical lens* through which the properties of the data are explored can influence outcomes in a major way. Formulating a suitable maintained hypothesis is particularly difficult in the presence of non-stationarity because of the different roles that parameters can play under the null hypothesis of a unit root and the alternative of stationarity or trend stationarity. Many of these issues of formulation have already been extensively studied in the literature on left-tailed unit root testing.

Suppose, for example, that the null hypothesis is difference stationarity and the alternative is stationarity. In a commonly used regression (Fuller, 1995)

*We thank two anonymous referees and the editor for helpful suggestions. We are grateful to Heather Anderson, Farshid Vahid, Hashem Pesaran, and Tom Smith for helpful discussions. Phillips acknowledges support from the NSF under Grant No. SES 09-56687. Shi acknowledges the Financial Integrity Research Network (FIRN) for funding support. Yu thanks the Singapore Ministry of Education for Academic Research Fund under grant number MOE2011-T2-2-096.

JEL Classification numbers: C15, C22.

$$R_1 : y_t = \alpha + \delta y_{t-1} + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t, \quad \varepsilon_t \overset{\text{i.i.d}}{\sim} \left(0, \sigma^2\right), \tag{1}$$

the null $\delta = 1$ is tested against the alternative $\delta < 1$ and the formulation (implicitly) allows for a non-zero mean in y_t under the stationary alternative even though α may be zero under the null. This regression is empirically more appealing than a regression without an intercept because of that flexibility. But if the null is difference stationarity and the alternative is trend stationarity, then the regression model (1) is inappropriate because an empirical trend may be misinterpreted as evidence of a unit root, leading to the augmented formulation

$$R_2 : y_t = \alpha_0 + \alpha_1 t + \delta y_{t-1} + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t, \qquad \varepsilon_t \overset{\text{i.i.d}}{\sim} (0, \sigma^2), \tag{2}$$

where we can test the null $\delta = 1$ against the alternative $\delta < 1$, even if $\alpha_1 = 0$ under the null. Use of the maintained hypothesis R_2 allows for both a unit root with drift ($\alpha_0 \neq 0$ and $\alpha_1 = 0$) under the null and trend stationarity ($\alpha_0 \neq 0$ and $\alpha_1 \neq 0$) under the alternative. Similar issues, of course, arise with more complex maintained hypotheses that allow for trend breaks and other deterministic components. The regression model of a left-tailed unit root test (against stationary or trend stationary alternatives) needs to nest the alternative hypothesis.[1]

Right-tailed unit root tests are also of empirical interest, particularly in detecting explosive or mildly explosive alternatives. For example, to find evidence of financial bubbles, Diba and Grossman (1988) applied right-tailed unit root tests to the fully sampled data. Phillips, Wu and Yu (2011b, PWY hereafter) suggested sequential implementations of right-tailed unit root tests to recursive subsamples; see also Phillips and J Yu (2011). As in left-tailed unit root testing, the formulation of the null and alternative hypotheses and the regression model specification are important in right-tailed tests. Different suggestions appear in the literature and no empirical guidelines have yet been offered. For example, Diba and Grossman used the regression model (2) whereas PWY employed model (1). Further, Diba and Grossman did not allow for bubble crashes in the alternative whereas various collapse mechanisms were considered in both Evans (1991) and P.C.B Phillips and J. Yu (2009, Unpublished Manuscript).

The present article examines appropriate ways of formulating regressions for right-tailed unit root tests to assess empirical evidence for explosive behaviour. To the best of our knowledge, this is the first time these formulation issues have been discussed in the literature. The discussion here focuses on the test procedures in PWY. Other tests for explosive behaviour are possible and many of these have been evaluated in simulations by Homm and Breitung (2012). Their simulations show that, while *ex post* analysis of the full sample data favors Chow type unit root tests for the detection of break points in the transition between unit root and explosive behaviour, recursive tests such as those in PWY perform well in early detection of such transitions and are preferable in this anticipative role as a monitoring system. Homm and Breitung (2012) also confirm that the PWY tests are more robust in the detection of multiple bubble episodes than the other tests they

[1] Similar arguments can be found in Dickey, William and Millar (1986).

considered. The primary intent of PWY was to develop recursive procedures that could assess whether Greenspan's remark on financial exuberance had empirical content at the time he made that statement in December 1995. It is in this context as an early warning device in market surveillance that the PWY tests were developed. Tests for bubbles using adaptive rolling window (RW) methods have also been developed in concurrent work of Phillips, Shi and Yu (2011a). RW methods are particularly useful in detecting and dating multiple bubble episodes. The specification issues raised in the present article apply equally well to all of these other break tests for financial exuberance.

The PWY test and the other tests discussed above belong to a class of reduced form approaches to bubble detection. Conventional unit root tests against stationary alternatives are similar reduced form methods with the explicit purpose of testing shock persistence: evidence of a unit root in data is interpreted as evidence of persistent shocks rather than transitory shocks, with obvious economic implications. In such tests the focus is accordingly very often on the null hypothesis, although we also have tests such as KPSS (Kwaitkowski *et al.*, 1992) of the null hypothesis of stationarity. In right sided unit root tests, the main focus is usually on the alternative hypothesis because we are particularly interested in whether movements in the data reflect evidence of exuberance or departure from fundamentals which commonly embody martingale rather than submartingale characteristics. In financial markets for instance, economic surveillance teams in central banks are now interested in whether data are indicative of market excesses or mispricing in relation to fundamentals; investors are interested in timing investment decisions in relation to market behaviour; and, most especially, regulators and policy makers are concerned to have early indication of any market excesses so that there is time for policy response. To the extent that right sided unit root tests are informative about mildly explosive or submartingale behavior in the data, these tests are useful as a form of market diagnostic or warning alert.

The rest of the article is organized as follows. Section II discusses appropriate choices for the null and alternative hypotheses and the formulation of the fitted regression model, suggesting a new model with a localized drift process whose magnitude may be determined empirically. Section III derives the limit distributions of the ADF statistic. Section IV discusses the explosive model of Evans (1991) for the alternative hypothesis. The sequential right-tailed ADF test, along with its finite sample and limit distributions, are explored in Section V. Section VI reports size and power properties for the sequential right-tailed ADF test. Using the proposed model formulations we apply the test to Nasdaq and S&P 500 market data in section VII. Section VIII concludes. The finite sample distribution of the ADF statistic and proofs of Proposition 1 are given in Appendix S1. Appendix A develops new limit theory for the consistent estimation of the localizing exponent in the drift process. Supporting Information is available in the original version (Phillips, Shi and Yu, 2012).

II. Formulating hypotheses and the fitted regression

The literature on right-tailed unit root testing has employed several different specifications for the null hypothesis. In PWY the null hypothesis is

$$H_{01}: y_t = y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \overset{\text{i.i.d}}{\sim} (0, \sigma^2),$$

so that Δy_t has mean zero and y_t has no deterministic trend. The *iid* error assumption may be relaxed but is convenient to retain for expository purposes in discussing the specification issues that are our main concern in the present article. In contrast to H_{01}, Diba and Grossman (1988) used the null model

$$H_{03}: y_t = \tilde{\alpha} + y_{t-1} + \varepsilon_t, \qquad \text{with a constant intercept } \tilde{\alpha},$$

so that y_t has deterministic trend behaviour when $\tilde{\alpha} \neq 0$. Under this null, the behaviour of y_t is dominated by the deterministic drift $\tilde{\alpha}t$.

A model that bridges these two null hypotheses involves a weak (local to zero) intercept with the form

$$H_{02}: y_t = \tilde{\alpha}T^{-\eta} + y_{t-1} + \varepsilon_t \qquad \text{with } \eta \geqslant 0. \tag{3}$$

Here y_t has an array formulation, the mean of Δy_t is $\tilde{\alpha}T^{-\eta} = O(T^{-\eta})$, and y_t has a deterministic drift of the form $\tilde{\alpha}t/T^{\eta}$ whose magnitude depends on the sample size and the localizing parameter η. When $\eta > 0$, this drift term is small relative to a linear trend. The null model H_{02} becomes H_{01} when $\eta \to \infty$ and H_{03} when $\eta \to 0$.

Localized drift models of the type (3) are useful in allowing for intermediate cases where there may be drift in the data but it may not be the dominating component. Much financial data over short and medium terms are of this type. As $y_t = \tilde{\alpha}t/T^{\eta} + \sum_{j=1}^{t} \varepsilon_j + y_0$, it is apparent that the drift is small in relation to the stochastic trend when $\eta > \frac{1}{2}$ and equal to or stronger than the stochastic trend when $\eta \in [0, \frac{1}{2}]$. When $\eta = \frac{1}{2}$, the standardized output $T^{-1/2}y_t$ behaves asymptotically like a Brownian motion with drift, which is appropriate in modelling some macroeconomic and financial time series. Moreover, empirical considerations often indicate that the constant is sample size or frequency dependent. For instance, as suggested by a referee, an annual return of 8% in stock prices corresponds to a monthly drift of $0.08/12 = 0.0067$ or a daily drift of $0.08/250 = 0.00032$.

The localizing coefficient η is not a choice or control parameter that is selected by the empirical investigator. Nonetheless, equation (3) may be formulated and on prior grounds it may be assumed that η lies in the negligible effect region $\eta > 0.5$ or the contributing effect region $\eta \leqslant 0.5$. Importantly, η is not consistently estimable when $\eta > 0.5$ because the drift component is dominated by the stochastic trend. In such cases, estimates of η typically converge to $1/2$, corresponding to the order of the stochastic trend (see Appendix A). Of course in these cases, since the drift effect is negligible relative to the stochastic trend, estimation is of lesser importance but still may be useful in assessing the trend order as stochastic. On the other hand, η is consistently estimable when $\eta \in [0, \frac{1}{2}]$, although only at a slow logarithmic rate when $\eta = \frac{1}{2}$, as shown in Appendix A.

One feature of the null hypothesis (3) is that it allows for an interpretation of the data in terms of random cycles about a trend. When $0 \leqslant \eta < 0.5$ the deterministic trend effect is a dominating characteristic. The null model therefore allows for deterministic drift combined with random wandering behaviour that is associated with a unit root in the system. In this event, the null offers a potential explanation of what might otherwise be regarded as apparent bubble activity in the data as a random cycle carrying persistent shocks about a deterministic trend. It is particularly helpful for econometric tests to afford some discriminatory capability between bubbles and this competing interpretation. Allowing for specifications such as equation (3) gives us the opportunity to do so, as we explain in what follows.

Similarly, different alternative hypotheses have been used in the literature on right-tailed unit root tests. The most obvious ones are the following explosive processes:

$$H_{A1}: y_t = \delta y_{t-1} + \varepsilon_t, \quad \delta > 1, \tag{4}$$

$$H_{A2}: y_t = \tilde{\alpha} + \delta y_{t-1} + \varepsilon_t, \quad \delta > 1, \tag{5}$$

$$H_{A3}: y_t = \tilde{\alpha} + \gamma t + \delta y_{t-1} + \varepsilon_t, \quad \delta > 1. \tag{6}$$

These three models mirror alternatives considered in left-tailed unit root tests where $\delta < 1$.

Explosive processes have a long history. In economics, Hicks (1950) suggested the possibility of explosive cyclical behaviour contained by certain structural floors and ceilings with the cycles arising from multiplier-accelerator dynamics. In statistics, White (1958) and Anderson (1959) studied the asymptotic properties of the least squares (LS) estimator under equation (4). In recent work, Phillips and Magdalinos (2007) suggested mildly explosive processes of the type

$$H_{A4}: y_t = \delta_T y_{t-1} + \varepsilon_t \qquad \text{with } \delta_T = 1 + cT^{-\theta}, \tag{7}$$

where $c > 0$, $\theta \in (0, 1)$, T is the sample size and δ_T is a moving parameter sequence. This model is called mildly explosive because the autoregressive coefficient δ_T is in an explosive region of unity (so that $\delta_T \to 1+$ as $T \to \infty$) that lies beyond the usual 'local to unity' interval where $\delta_T = 1 + c/T$ for which the random wandering limit behaviour of the process is similar to that of the limiting behaviour of a unit root process. Under H_{A4}, the behaviour of y_t resembles that of an explosive time series rather than that of a unit root process.

Model (5) is formulated with a non-zero intercept and produces a dominating deterministic component that has an empirically unrealistic explosive form (P. C. B. Phillips and J. Yu, Unpublished Manuscript 2009, PY hereafter). Similar characteristics apply a fortiori in the case of the inclusion of a deterministic trend term in model (6). These forms are unreasonable for most economic and financial time series and an empirically more realistic description of explosive behaviour is given by models (4) and (7), which are both formulated without an intercept or a deterministic trend.

The empirical regression of the right-tailed unit root test given in Diba and Grossman (1988) is R_2. This regression has both a constant as well as a deterministic trend. Since the presence of either of these two terms is empirically unrealistic when $\delta > 1$, regression R_2 is not suitable for right-tailed unit root testing. By contrast, regression R_1 is empirically more realistic and PWY implemented a right-tailed unit root test using this regression formulation.

In light of the above discussion, we recommend that right-tailed unit root tests may be suitably formulated with a null hypothesis H_{02} and an empirical regression R_1. As H_{02} depends on η, we discuss the asymptotic distribution of the test statistic and examine the size and the power properties of the right-tailed unit root test for different values of η in H_{02}. Simulation findings reported below provide further guidelines for the selection of the null and the regression model with associated test critical values.

We are particularly interested in the robustness of tests to the value of the localizing coefficient η that determines the strength of the drift function in the data. In general, our findings indicate that the recommended test is robust to values of $\eta > 0.5$, that is for models where the drift component in the data is not the dominant data characteristic. When the drift is dominant, i.e. when $\eta \in (0.0.5)$, test results are sensitive to the value of η. It therefore seems sensible in empirical work to report results for a range of values of η so that any such sensitives are evident in the empirical findings. As indicated earlier, the parameter η may also be estimated (see Appendix A), and when $\eta \in (0.0.5]$, consistent estimation of η is possible. These estimation results might be included in the empirical findings.

III. Full-sample right-tailed unit root tests

Right-tailed unit root tests, like their left-tailed counterparts, have asymptotic distributions that depend on the specification of the null hypothesis and the regression model. As discussed above, a suitable regression model for right-tailed testing is R_2 and an empirically reasonable null is a unit root process with a drift of the form $\tilde{\alpha}T^{\eta}$, arising from H_{02}. The right-tailed unit root test discussed in this section is the ADF test applied to the full sample. Other unit root tests can be studied in exactly the same manner. Note that the magnitude of the drift is inversely related to parameter η. Accordingly, we assume that the data generating process is given by the model

$$\Delta y_t = \alpha + \beta y_{t-1} + \sum_{k=1}^{p-1} \phi_k \Delta y_{t-k} + \varepsilon_t, \quad \varepsilon_t \overset{\text{i.i.d.}}{\sim} (0, \sigma^2), \tag{8}$$

with $\alpha = \tilde{\alpha}T^{-\eta}$ and null hypothesis $H_{02} : \beta = 0$. The assumption of i.i.d errors may be relaxed in deriving the limit theory under both the null (Phillips, 1987; Phillips and Solo, 1992) and the alternative (Lee, 2011) but is retained for convenience here.[2]

Lemma 1. Under the model (8) and null hypothesis H_{02} with $\eta > 0.5$, the asymptotic distribution of the ADF t-statistic is

$$\mathrm{ADF} \xrightarrow{L} \frac{\frac{1}{2}[W^2(1) - 1] - W(1)\int_0^1 W(s)\,\mathrm{d}s}{\left\{\int_0^1 W^2(s)\,\mathrm{d}s - \left[\int_0^1 W(s)\,\mathrm{d}s\right]^2\right\}^{1/2}} := F_1(W), \tag{9}$$

where W is a standard Wiener process and $\xrightarrow{L}$ denotes convergence in distribution; If H_{02} holds with $0 \leqslant \eta < 0.5$, then the asymptotic ADF distribution is

$$\mathrm{ADF} \xrightarrow{L} \left[\int_0^1 s\mathrm{d}W(s) - \int_0^1 W(s)\mathrm{d}s\right]\left(\int_0^1 s^2\,\mathrm{d}s\right)^{-1/2} := F_2(W). \tag{10}$$

The proof is subsumed within the proof of proposition 1.

Remark 1. The asymptotic ADF distribution when $\eta > 0.5$ is identical to that for the PWY formulation despite the inclusion of an intercept in the null model. The intercept does not affect the limit distribution because the implied drift in the process has smaller order than the stochastic trend.

[2] The finite sample distributions of the ADF statistic have a similar pattern to Figure 2.

Remark 2. Suppose the null hypothesis is specified as H_{03}. The asymptotic ADF distribution in this case[3] is identical to that of the case when $0 \leqslant \eta < 0.5$ [equation(10)]. Here the implied drift has higher order of magnitude and behaves like a linear deterministic trend.

Remark 3. The asymptotic ADF distribution when $\eta = 0.5$ is

$$\mathrm{ADF} \xrightarrow{L} (D_\sigma - A_\sigma C_\sigma)(B_\sigma - A_\sigma^2)^{-1/2}, \tag{11}$$

with $A_\sigma = \frac{1}{2}\tilde{\alpha} + \sigma \int_0^1 W(s)\mathrm{d}s, B_\sigma = \frac{1}{3}\tilde{\alpha}^2 + \sigma^2 \int_0^1 W(s)^2\mathrm{d}s + 2\tilde{\alpha}\sigma \int_0^1 W(s)s\mathrm{d}s, C_\sigma = W(1)$ and $D_\sigma = \tilde{\alpha}[W(1) - \int_0^1 W(s)\mathrm{d}s] + \frac{1}{2}\sigma[W(1)^2 - 1]$. Importantly, the limit theory in this case depends on the nuisance parameters $\tilde{\alpha}$ and σ, so it is not invariant unless we include a trend in the regression or adjust for the trend in some other way (as, for example in Schmidt and Phillips, 1992).

IV. Specifications for explosive behaviour

Two specifications for the alternative hypothesis, both formulated without an intercept or a deterministic trend, are given by model (4) and model (7) in Section II. Neither model has structural breaks. But as argued in Evans (1991, p. 924) 'bubbles do not appear to be empirically plausible unless there is a significant chance that they will collapse after reaching high levels.' This argument is consistent with other models of explosive processes such as the business cycle model of Hicks (1950), where each cycle has an explosive expansion phase and a subsequent downswing due to disinvestment proceeding at the rate of deterioration of capital. Thus, more complete specification of the alternative hypothesis requires the inclusion of a downswing or bubble collapse process. This section considers a simple time series model proposed by Evans (1991).

The DGP proposed by Evans (1991) consists of a market fundamental component P_t^f, which follows a random walk process

$$P_t^f = \tilde{u} + P_{t-1}^f + \sigma_f \varepsilon_t, \varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0,1), \tag{12}$$

and a periodically collapsing explosive bubble component such that

$$B_{t+1} = \rho^{-1} B_t \varepsilon_{B,t+1}, \qquad \text{if } \; B_t < b \tag{13}$$

$$B_{t+1} = [\zeta + (\pi\rho)^{-1}\theta_{t+1}(B_t - \rho\zeta)]\varepsilon_{B,t+1}, \qquad \text{if } \; B_t \geqslant b, \tag{14}$$

where $\rho^{-1} > 1$ and $\varepsilon_{B,t} = \exp(y_t - \tau^2/2)$ with $y_t \overset{\text{i:i:d}}{\sim} N(0, \tau^2)$. θ_t follows a Bernoulli process which takes the value 1 with probability π and 0 with probability $1 - \pi$. ζ is the remaining size after the bubble collapse. The bubble component has the property that $\mathbb{E}_t(B_{t+1}) = \rho^{-1} B_t$. By construction, the bubbles collapse completely in a single period when triggered by the Bernoulli process realization.

The market fundamental equation (12) may be obtained by combining a random walk dividend process with a Lucas asset pricing equation as follows

[3] The asymptotic ADF distribution under this case is well documented in the unit root literature; see Phillips (1987) and Phillips and Perron (1988)

TABLE 1

Parameter settings

	$\tilde{u}$	σ_f	P_0^f	ρ	b	B_0	π	ζ	τ	κ
Yearly	0.740	7.869	41.195	0.952	1	0.50	0.85	0.50	0.05	20
Monthly	0.131	3.829	94.122	0.985	1	0.50	0.85	0.50	0.05	150

$$D_t = \mu + D_{t-1} + \varepsilon_{Dt}, \qquad \varepsilon_{Dt} \overset{\text{i.i.d.}}{\sim} N(0, \sigma_D^2), \tag{15}$$

$$P_t^f = \frac{\mu\rho}{(1-\rho)^2} + \frac{\rho}{1-\rho} D_t, \tag{16}$$

where μ is the drift of the dividend process and σ_D^2 the variance of the dividend innovations. The drift ($\tilde{u}$) of the market fundamental process is $\mu\rho(1-\rho)^{-1}$ and the standard deviation is $\sigma_f = \sigma_D \rho(1-\rho)^{-1}$. In Evans (1991), the parameter values for μ and σ_D^2 were matched to the sample mean and sample variance of the first differences of real S&P500 dividends from 1871 to 1980. The value for the discount factor ρ is equivalent to a 5% annual interest rate. So the parameter settings in Evans (1991) correspond to a yearly frequency. In accordance with our empirical application, we consider a parameter set calibrated to monthly data. Correspondingly, the parameters μ and σ_D^2 are determined as the sample mean and sample variance of the monthly first differences of real Nasdaq dividends as described in the application section (normalized to unity at the beginning of the sample period). These are $\mu = 0.0020$ and $\sigma_D^2 = 0.0034$, respectively. The discount factor equals 0.985. We can then calculate the values of $\tilde{u}$, σ_f, P_0^f based on those of μ, σ_D^2, D_0.

The settings of the parameters in the bubble components [equations (13)–(14)] are the same as those in Evans (1991). The asset price P_t is equal to the sum of the market fundamental component and the bubble component, namely $P_t = P_t^f + \kappa B_t$, where κ controls the relative magnitudes of these two components. The parameter settings are given in Table 1 for *yearly* and *monthly* data.

V. The sup ADF test

Evans (1991) argued that right-tailed unit root tests, when applied to the full sample, have little power to detect periodically collapsing bubbles and demonstrated this effect in simulations. The low power of standard unit root tests is due to the fact that periodically collapsing bubble processes behave rather like an $I(1)$ process or even a stationary linear autoregressive process when the probability of bubble collapse is non-negligible.

To overcome the problem identified in Evans, PWY proposed a sup ADF (SADF) statistic to test for the presence of explosive behaviour in a full sample. In particular, the methods rely on forward recursive regressions coupled with sequential right-sided unit root tests. Similar methods using other test procedures like the cumulative sum (CUSUM) test have been considered recently in Homm and Breitung (2012). Generalized versions of the sup ADF test that employs variable rolling windows are also possible (see Phillips *et al.* 2011a).

These sequential tests assess period by period evidence for unit root behaviour against explosive alternatives. If the right-tailed ADF test is employed in each period, the test statistic proposed by PWY is the sup value of the corresponding ADF sequence. In this setup, the alternative hypothesis of the test therefore includes the periodically collapsing explosive behaviour. The null hypotheses are exactly the same as that for the right-tailed unit root test in equation (3).

Suppose r is the window size of the regression (in proportion to the full sample size) for the right-tailed unit root test. In the sup ADF test, the window size r expands from r_0 to 1 through recursive calculations. The smallest window size r_0 is selected to ensure that there are sufficient observations to initiate the recursion. The number of observations in the regression is $T_r = [Tr]$, where $[\cdot]$ signifies the integer part of its argument and T is the total number of observations.

The fitted regression model for the sup ADF test is R_1. The corresponding ADF t-statistic is denoted by ADF_r. To test for the existence of bubbles, inferences are based on the sup ADF statistic $\text{SADF}(r_0) = \sup_{r\in[r_0,1]} \text{ADF}_r$. This notation highlights the dependence of SADF on the initialization parameter r_0.

Limit distribution of sup ADF

Proposition 1. Under the model (8) and null hypothesis H_{02} with $\eta > 0.5$, the asymptotic distribution of the sup ADF statistic is

$$\text{SADF}(r_0) \xrightarrow{L} \sup_{r\in[r_0,\,1]} \left\{ \frac{\frac{1}{2}r[W(r)^2 - r] - \int_0^r W(s)\mathrm{d}s W(r)}{r^{1/2}\left\{ r\int_0^r W(s)^2\mathrm{d}s - \left[\int_0^r W(s)\mathrm{d}s\right]^2 \right\}^{1/2}} \right\} := F_3(W, r_0), \quad (17)$$

If H_{02} holds with $0 \leqslant \eta < 0.5$, the sup ADF statistic converges to

$$\text{SADF}(r_0) \xrightarrow{L} \sup_{r\in[r_0,1]} \left\{ \left[\int_0^r s\mathrm{d}W(s) - \int_0^r W(s)\mathrm{d}s\right] \left(\int_0^r s^2\mathrm{d}s\right)^{-1/2} \right\} := F_4(W, r_0). \quad (18)$$

Remark 4. The asymptotic ADF_r distribution when $0 \leqslant \eta < 0.5$ is

$$\text{ADF}_r \xrightarrow{L} \left[\int_0^r s\mathrm{d}W(s) - \int_0^r W(s)\,\mathrm{d}s\right] \left(\int_0^r s^2\mathrm{d}s\right)^{-1/2}, \quad (19)$$

which is distributed as standard normal. Suppose $r_A, r_B \in [r_0, 1]$ and $r_A \neq r_B$, the asymptotic ADF_{r_A} distribution and the asymptotic ADF_{r_B} distribution are correlated since both are functions of the same standard Wiener process.

Remark 5. The asymptotic distribution of the SADF statistic when $\eta = 0.5$ is

$$\text{SADF}(r_0) \xrightarrow{L} \sup_{r\in[r_0,1]} [r^{-1/2}(rD_{r,\sigma} - A_{r,\sigma}C_{r,\sigma})(rB_{r,\sigma} - A_{r,\sigma}^2)^{-1/2}],$$

with $A_{r,\sigma} = \frac{1}{2}\tilde{\alpha}r^2 + \sigma\int_0^r W(s)\mathrm{d}s$, $B_{r,\sigma} = \frac{1}{3}\tilde{\alpha}^2r^3 + \sigma^2\int_0^r W(s)^2\mathrm{d}s + 2\tilde{\alpha}\sigma\int_0^r W(s)s\mathrm{d}s$, $C_{r,\sigma} = W(r)$ and $D_{r,\sigma} = \tilde{\alpha}[rW(r) - \int_0^r W(s)\mathrm{d}s] + \frac{1}{2}\sigma[W(r)^2 - r]$. Similar to the ADF statistic, in this case the limit theory depends on the nuisance parameters $\tilde{\alpha}$ and σ.

Figures 1(a,b) examine the sensitivity of the asymptotic distributions of SADF when $\eta > 0.5$ and $0 \leqslant \eta < 0.5$ with respect to r_0. The distributions are obtained from 20000 replications, approximating the Wiener process by partial sums of standard normal variates with 5000 steps, and using a standard Gaussian kernel density estimate and Silverman rule for bandwidth choice (Silverman, 1986). The smallest window size r_0 is set to $\{0.2, 0.15, 0.10, 0.05\}$.

Figure 1(a) displays the asymptotic distributions when $\eta > 0.5$ (i.e. $F_3(W, r_0)$) while Figure 1(b) is for the case $0 \leqslant \eta < 0.5$ (i.e. $F_4(W, r_0)$). Under both cases, the asymptotic distributions of the SADF statistic move sequentially to the right as r_0 decreases.[4] In addition, the asymptotic distribution $F_4(W, r_0)$ has larger values for the 90%, 95% and 99% quantiles. For example, the 95% critical values of $F_3(W, r_0)$ with $r_0 = \{0.2, 0.15, 0.10, 0.05\}$ are respectively $1.39, 1.44, 1.50, 1.57$ while those of $F_4(W, r_0)$ are respectively $2.79, 2.84, 2.90, 3.00$. Obviously, the critical values are sensitive to r_0 and this needs to be taken into account in empirical work. The smallest window size r_0 is a choice parameter and will typically be selected closer to zero when the early part of the sample is of interest and needs to be included in the recursion, subject to degrees of freedom in the ADF regression.

The finite sample distribution of sup ADF

The finite sample distribution of the SADF statistic depends on the sample size T, the value of the drift in the null hypothesis (depending on T and η) and the smallest window size r_0. Figure 2 describes the finite sample distributions of the SADF statistic obtained by kernel density estimation for $T = 400$, $r_0 = 0.1$, $\tilde{\alpha} = 1$,[5] and $\eta = \{1, 0.8, 0.6, 0.5, 0.4, 0.2, 0\}$. The bold solid lines are the asymptotic distributions and the dotted lines are the finite sample distributions. For given T and r_0 the finite sample distribution moves towards $F_3(W, 0.1)$

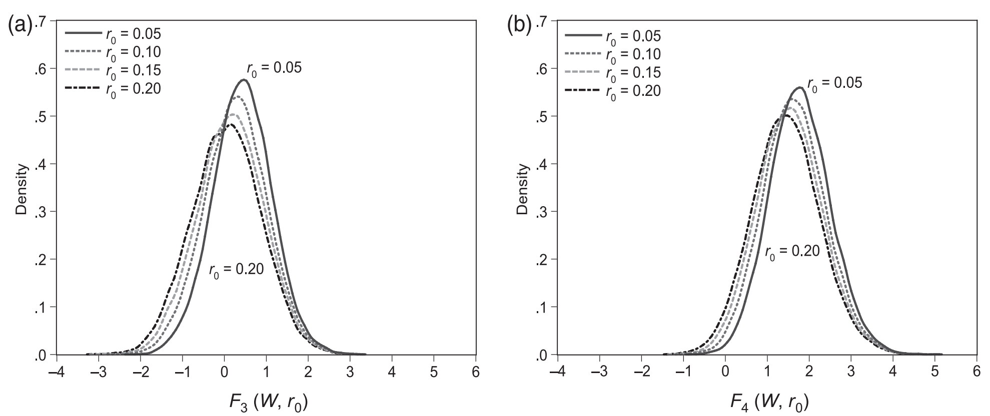

Figure 1. The asymptotic distributions of the SADF statistic with $r_0 = \{0.20, 0.15, 0.10, 0.05\}$

[4] Intuitively, when r_0 is smaller, the feasible range of r (i.e. $[r_0, 1]$) becomes wider and hence the parameter space of the distribution of $\lim_{T\to\infty} \mathrm{ADF}_r$ expands. The asymptotic SADF distribution, which applies the sup function to the aforementioned distribution, then moves sequentially towards the right as r_0 decreases.

[5] With this normalization, $\alpha_T = T^{-\eta}$ and the localization exponent η is the sole determinant of drift magnitude. See the discussion in Appendix A.

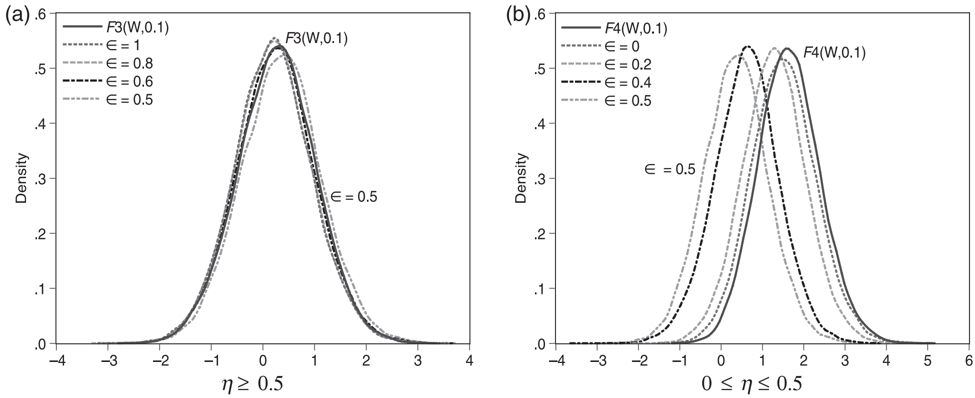

Figure 2. The finite sample distributions of the SADF statistic when $T = 400, r_0 = 0.1$ and $\eta = \{1, 0.8, 0.6, 0.5, 0.4, 0.2, 0\}$

as η increases and shifts towards $F_4(W, 0.1)$ as η decreases. An obvious separation occurs when $\eta = 0.5$. The discrepancies among the finite sample distributions are negligible with $\eta \in \{0.6, 0.8, 1\}$ but become larger for $\eta \in \{0.4, 0.2, 0\}$.

The finite sample SADF distribution is invariant to η when $\eta > 0.5$ but varies significantly with η when η is less than 0.5.

VI. Size and power comparison

The 90%, 95% and 99% quantiles of the asymptotic distributions of the SADF statistic when $\eta > 0.5$ and $0 \leqslant \eta < 0.5$ (i.e. $F_3(W, r_0)$ and $F_4(W, r_0)$) are presented in Table 2. As before, critical values are obtained by simulations with 20,000 replications of Wiener processes in terms of partial sums of standard normal variates with 5,000 steps.

Table 3 gives sizes for the SADF test based on nominal asymptotic critical values with sample sizes $T = 100, 200$ and 400. The nominal size is 5%. The DGP is specified according to the respective null hypotheses with $\tilde{\alpha} = 1$, $\eta = \{1, 0.8, 0.6, 0.4, 0.2, 0\}$. The number of replications is 20,000. The lag order is determined by BIC with maximum lag length 12. The smallest window size has 40 observations. Table 3 shows that for all cases with $\eta > 0.5$ there is no obvious size distortion when using the asymptotic critical values,[6] whereas there are significant size distortions for some cases with $0 \leqslant \eta < 0.5$. In particular, we note that there is little size distortion for the case $\eta = 0$ but size distortion becomes progressively more severe when the value of η increases to 0.5. For example, the size of the SADF test is 0.045, 0.021 and 0.002 for $\eta = \{0, 0.2, 0.4\}$ respectively when the sample size $T = 400$ and $r_0 = 0.1$.[7] The jump in size between the two cases $\eta = 0.4$ and $\eta = 0.6$ is caused by the discontinuity in the asymptotic theory between $\eta \geqslant 0.5$ and $0 \leqslant \eta < 0.5$, as shown in Figure 2(a),(b). The finite sample distributions approach $F_4(W, r_0)$ as $\eta \to 0$

[6]There are significant size distortions when using the significance test proposed by Campbell and Perron (1991) (with the maximum lag length 12) to determine the lag order. For example, the size of the SADF test when $\eta = 1$ is 0.119, 0.122 and 0.122 for $T = \{100, 200, 400\}$ respectively.

[7]We observe similar patterns of size distortion when $r_0 = 0.4$ for all sample sizes. However, when T is large, there is some advantage to using a small value for r_0 so that the sup ADF test does not miss any opportunity to capture an explosive phase in the data, as discussed earlier in the text.

TABLE 2

Asymptotic critical values of the SADF statistic (against explosive alternative)

	$F_3(W,r_0)$			$F_4(W,r_0)$		
	90%	95%	99%	90%	95%	99%
$r_0=0.4$	0.86	1.17	1.77	2.29	2.60	3.21
$r_0=0.2$	1.10	1.39	1.95	2.47	2.79	3.40
$r_0=0.1$	1.23	1.51	2.04	2.62	2.90	3.48

Note: Asymptotic critical values are obtained using 20,000 replications and partial sums with 5,000 steps.

TABLE 3

Sizes of the SADF test (using asymptotic critical values). The data generating process is specified according to the respective null hypothesis. The nominal size is 5%

	$\eta>0.5$			$0\leqslant\eta<0.5$		
	$\eta=1$	$\eta=0.8$	$\eta=0.6$	$\eta=0.4$	$\eta=0.2$	$\eta=0$
$T=100$ and $r_0=0.4$	0.043	0.046	0.056	0.004	0.019	0.039
$T=200$ and $r_0=0.2$	0.045	0.046	0.056	0.003	0.020	0.044
$T=400$ and $r_0=0.1$	0.048	0.048	0.056	0.002	0.021	0.045

Note: Size calculations are based on 20,000 replications.

which causes the reduction in the size distortion observed in Table 3 as $\eta\to 0$. Note that in all these cases for $0\leqslant\eta<0.5,$ the distortion is towards conservative tests.

To assess test power we assume the DGP is Evans (1991) periodically collapsing explosive process, with both yearly and monthly parameters settings (see Table 1). The sample sizes considered for those two parameters settings are $T=\{100,200\}$ and $T=\{100,200,400\},$ respectively. For each parameter and sample size setting, we calculate powers of the sup ADF test under four different specifications in the null hypothesis: $\eta>0.5,$[8] $\eta=0.4$, $\eta=0.2$ and $\eta=0$, all with $\tilde{\alpha}=1$. The powers for cases $\eta>0.5$ and $\eta=0$ are calculated from the 95% quantile of $F_3(W,r_0)$ and $F_4(W,r_0),$ respectively (Table 2). The power calculations for $\eta=0.4$ and $\eta=0.2$ are based on the 95% quantiles of the finite sample distributions (Table 4). These calculations are intended to show how power may depend on $\eta,$ while noting that η is not a choice parameter and is typically unknown to the empirical investigator. The number of replications in all simulations reported below is 20,000.

From Table 5, test power evidently increases with sample size. Under the yearly parameter setting and $T=200$, power for $\eta>0.5$, $\eta=0.4$, $\eta=0.2$ and $\eta=0$ is 20%, 21%, 21% and 22% higher than when $T=100$. Power for $\eta>0.5$ is always higher than when $0\leqslant\eta<0.5,$ as might be expected because a significant deterministic trend $(0\leqslant\eta<0.5)$ provides an alternate mechanism for capturing the variation in periodically explosive data.

[8] This is due to the observation that as long as η is greater than 0.5, the discrepancy among the finite sample critical values of the SADF statistic is negligible.

TABLE 4

The finite sample critical values of the SADF statistic (against explosive alternative)

	$\eta=0.4$			$\eta=0.2$		
	90%	95%	99%	90%	95%	99%
$T=100$ and $r_0=0.4$	1.27	1.61	2.28	1.86	2.20	2.85
$T=200$ and $r_0=0.2$	1.50	1.81	2.44	2.13	2.45	3.05
$T=400$ and $r_0=0.1$	1.63	1.91	2.46	2.29	2.59	3.17

Note: The finite sample critical values are obtained by simulation with 20,000 replications.

TABLE 5

Powers of the SADF test under Evans (1991) periodically collapsing explosive behaviour

	$\eta>0.5$	$\eta=0.4$	$\eta=0.2$	$\eta=0$
Yearly parameter settings				
$T=100$ and $r_0=0.4$	0.43	0.36	0.27	0.22
$T=200$ and $r_0=0.2$	0.63	0.57	0.48	0.44
Monthly parameter settings				
$T=100$ and $r_0=0.4$	0.58	0.49	0.34	0.26
$T=200$ and $r_0=0.2$	0.75	0.67	0.53	0.48
$T=400$ and $r_0=0.1$	0.86	0.81	0.71	0.68

Note: Power calculations are based on 20,000 replications.

In addition, when $0 \leqslant \eta < 0.5$, power decreases as $\eta \to 0$. From the lower panel of Table 5 (monthly parameters settings), when $T=400$, for instance, the power of the test is 86% when $\eta > 0.5$ and then declines from 81% to 68% as η changes from 0.4 to zero.

VII. Empirics

We conduct an empirical application of the sup ADF test to the Nasdaq and S&P 500. The Nasdaq composite index and the Nasdaq dividend yield are sampled from February 1973 to July 2009 (constituting 438 observations), obtained from DataStream International. The consumer price index, which is used to convert stock prices and dividends into real series, is downloaded from the Federal Reserve Bank of St Louis. The real S&P 500 stock price index and the real S&P 500 stock price index dividend are obtained from Robert Shiller's website.

Table 6 displays the SADF statistics for the logarithmic real Nasdaq index and the logarithmic real Nasdaq dividend, along with respective critical values for $\eta > 0.5$, $\eta=0.4$, $\eta=0.2$ and $\eta=0$ using the normalized form of the localized drift parameter $\alpha_T=\tilde{\alpha}T^{-\eta}$ with $\tilde{\alpha}=1$ under the null (2). Table 7 presents the SADF statistic and respective critical values for the S&P 500 price-to-dividend ratio. The lag order is determined by BIC with maximum lag length 12. The smallest fractional window r_0 is set to be 0.1. As in the simulation experiments, we use asymptotic critical values for the specifications $\eta > 0.5$ and $\eta=0$ and finite sample critical values for the specifications $\eta=0.4$ and $\eta=0.2$. Although the parameter η is generally unknown to the investigator, we report critical values for

TABLE 6

The sup ADF test of the NASDAQ stock market

	SADF statistic			
Log Real NASDAQ Index	2.56			
Log Real NASDAQ Dividend	−1.07			
	$\eta > 0.5$	$\eta = 0.4$	$\eta = 0.2$	$\eta = 0$
90%	1.23	1.63	2.29	2.62
95%	1.51	1.91	2.59	2.90
99%	2.04	2.51	3.16	3.48

Note: Critical values of the sup ADF test under the specification of $\eta = 0.4$ and $\eta = 0.2$ are obtained by simulations with 20,000 replications and sample size 438. The smallest fractional window r_0 is set to be 0.1.

TABLE 7

The sup ADF test of the S&P 500 stock market

Price-to-dividend ratio	*SADF statistic*			
	3.44			
$\eta > 0.5$	$\eta = 0.4$	$\eta = 0.2$	$\eta = 0$	
90%	1.23	1.70	2.43	2.62
95%	1.51	1.99	2.73	2.90
99%	2.04	2.56	3.26	3.48

Note: Critical values of the sup ADF test under the specification of $\eta = 0.4$ and $\eta = 0.2$ are obtained by simulations with 20,000 replications and sample size 1,680. The smallest fractional window r_0 is set to be 0.1.

these cases so that the robustness of the empirical results can be assessed against model specifications with various values of η. The finite sample critical values for the Nasdaq and the S&P 500 markets are obtained from simulations with 20,000 replications and sample size 438 and 1,680 respectively.

As is evident by Table 6, for the logarithmic real Nasdaq index, we reject the unit root null hypothesis in favour of the explosive alternative at the 10% significance level under model specifications with $\eta > 0.5$, $\eta = 0.4$ and $\eta = 0.2$, whereas we fail to reject the null hypothesis at the 10% significance level under the specification of $\eta = 0$ (although the difference between the test statistic and the critical value is very small in this case). Furthermore, we cannot reject the null hypothesis of a unit root at the 10% significance level for the logarithmic real Nasdaq dividend under all specifications considered.

In other words, for model specifications with $\eta > 0.5$, $\eta = 0.4$ and $\eta = 0.2$, we find evidence of explosive behaviour in the Nasdaq stock market using the sup ADF test. However, if the null hypothesis is specified as

$$H_{03}: y_t = 1 + y_{t-1} + \varepsilon_t, \quad \varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0, \sigma^2), \tag{20}$$

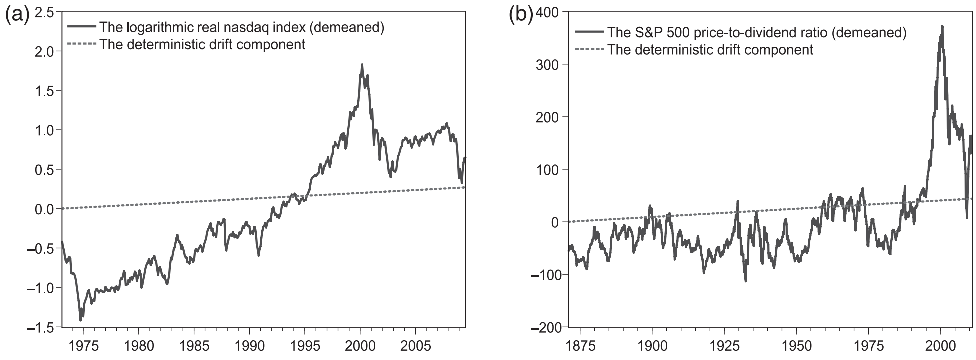

Figure 3. The deterministic drift component. (a) the logarithmic real NASDAQ index; (b) the S&P 500 price-to-dividend ratio

(i.e. the specification corresponding to $\eta = 0$), the sup ADF finds no evidence of bubble existence in the Nasdaq during the sample period. This null hypothesis implies that the long-term average return of the stock index is 100%, which is obviously unrealistic and can be excluded on prior grounds. That is, for this parameterization of the drift the null specification has a constant proportional growth component in the specification that implies a null generating mechanism $y_t = t + \sum_{s=1}^{t} \varepsilon_s + y_0$ with a strongly dominant linear trend that gives a growth rate of 100% a year. Evidence discussed below indicates that the localizing drift coefficient η for this series is close to 1, providing confirmatory empirical rejection of equation (20).

From Table 7, we reject the unit root null hypothesis against the explosive alternative at the 10% significance level under all model specifications for the S&P 500 price-to-dividend ratio. Accordingly, the SADF test provides strong evidence for the presence of explosive behaviour in both the Nasdaq (confirming the conclusion of PWY) and the S&P 500 stock markets, showing that the evidence is robust to all specifications of the null model (with the exception of extreme models such as equation (20) for the Nasdaq for which there is little prior or empirical support).

As shown in Appendix A, the localizing drift exponent η in the null model can be consistently estimated when $0 \leqslant \eta \leqslant 0.5$. When $\eta > 0.5$ the dominating component is the stochastic trend and estimates of η accordingly converge in probability to 0.5. The empirical estimate obtained by the method of Appendix A (fitted intercept) is $\hat{\eta} = 0.99$ for the logarithmic real Nasdaq index and $\hat{\eta} = 0.30$ for the S&P 500 price-to-dividend ratio. The deterministic drift component (i.e. $\alpha_T t$) of the logarithmic real Nasdaq index and the S&P 500 price-to-dividend ratio is displayed in Figure 3(a,b) respectively. The bubble character of the data are prominent in these figures even with the removal of the deterministic drift.

VIII. Conclusion

This article has investigated various formulations of the null and alternative hypotheses in studying empirical evidence of exuberance in economic and financial time series. The formulations involve different specifications of the regression models used for the construction of empirical tests of exuberance, which are shown to impact both the finite sample and the asymptotic distributions of the tests.

Our findings suggest an empirical model specification for use in practical work. The empirical model does not include a linear deterministic trend in the regression but has a fitted intercept and thereby allows for some deterministic drift in the process under the null hypothesis of a unit root. The drift coefficient is generally unknown to the investigator but may be consistently estimated even when it is local to zero, although the rate of convergence is slow in this case, as discussed in Appendix A. Moreover, empirical findings may be assessed against a wide range of specifications that include drift in the null. The test relies on estimation (or recursive estimation) of the autoregressive coefficient in the model

$$\Delta y_t = \alpha_T + \beta y_{t-1} + \sum_{i=1}^{k} \phi_i \Delta y_{t-i} + \varepsilon_t,$$

where the null hypothesis ($\beta = 0$) allows for an intercept $\alpha_T = \tilde{\alpha} T^{-\eta}$ that is local to zero. The limit distributions of the ADF and SADF statistics are derived for cases where $\eta > 0.5$, $\eta = 0.5$ and $0 \leqslant \eta < 0.5$ with corresponding asymptotic critical values that may be used to assess evidence in support of the null or alternative ($\beta > 0$) contingent on the model specification for a range of values of η. This approach permits the investigator to assess the robustness of the findings to different specifications of the deterministic trend in the model. When $\eta \leqslant 0.5$, the parameter may be consistently estimated using the method given in Appendix A.

Final Manuscript Received: December 2012

References

Anderson, T. W. (1959). ‘On asymptotic distributions of estimates of parameters of stochastic difference equations’, *Annals of Mathematical Statistics*, Vol. 30, pp. 676–687.

Campbell, J. Y. and Perron, P. (1991). ‘Pitfalls and opportunities: what macroeconomists should know about unit roots”, In *National Bureau of Economic Research Macroeconomics Annual*, Vol. 6, MIT Press, pp. 141–201.

Diba, B. T. and Grossman, H. I. (1988). ‘Explosive rational bubbles in stock prices?’ *The American Economic Review*, Vol. 78, pp. 520–530.

Dickey, D. A., William, R. B. and Miller, R. B. (1986). ‘Unit roots in time series models: Tests and implications’, *The American Statistician*, Vol. 40, pp. 12–26.

Evans, G. W. (1991). ‘Pitfalls in testing for explosive bubbles in asset prices’, *The American Economic Review*, Vol. 81, pp. 922–930.

Fuller, W. A. (1995). *Introduction to Statistical Time Series*, New York, John Wiley & Sons, Inc.

Hicks, J. (1950). *A Contribution to the Theory of Trade Cycle*, Oxford, Oxford University Press.

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

Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992). ‘Testing the null hypothesis of stationarity against the alternative of a unit root’, *Journal of Econometrics*, Vol. 54, pp. 159–178.

Lee, J.-H. (2011). *Asymptotics for Explosive Autoregression with Conditional Heteroskedasticity*, Working Paper, Yale University.

Phillips, P. C. B. (1987). ‘Time series regression with a unit root’, *Econometrica*, Vol. 55, pp. 277–301.

Phillips, P. C. B., and Magdalinos, T. (2007). ‘Limit theory for moderate deviations from a unit root’, *Journal of Econometrics*, Vol. 136, pp. 115–130.

Phillips, P. C. B., and Perron, P. (1988). 'Testing for a unit root in time series regression', *Biometrika*, Vol. 75, pp. 335–346.
Phillips, P. C. B. and Solo, V. (1992). 'Asymptotics for linear processes', *The Annals of Statistics*, Vol. 20, pp. 971–1001.
Phillips, P.C.B., and Yu, J. (2011). 'Dating the timeline of financial bubbles during the subprime crisis', *Quantitative Economics*, Vol. 2, pp. 455–491.
Phillips, P. C. B., Shi, S. and Yu, J. (2011a). *Testing for Multiple Bubbles*, Singapore Management University, Working Papers No. 09-2011.
Phillips, P. C. B., Wu, Y. and Yu, J. (2011b). 'Explosive behavior in the 1990s Nasdaq: when did exuberance escalate asset values?' *International Economic Review*, Vol. 52, pp. 201–226.
Phillips, P. C. B., Shi, S. and Yu, J. (2012). *Specification Sensitivity in Right-Tailed Unit Root Testing for Explosive Behavior*, Cowles Foundation for Research in Economics, Discussion Paper, CFDP #1842.
Schmidt, P. and Phillips, P.C.B. (1992). 'LM tests for a unit root in the presence of deterministic trends', *Oxford Bulletin of Economics and Statistics*, Vol. 54, pp. 257–287.
Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.
White, J. S. (1958). 'The limiting distribution of the serial correlation coefficient in the explosive case', *Annals of Mathematical Statistics*, Vol. 29, pp. 1188–1197.

Appendix A

Estimation of the localizing trend exponent η

Suppose $0 \leqslant \eta \leqslant 0.5$. We develop a procedure for the consistent estimation of η in this case. Consider the null model

$$y_t = \alpha_T t + \sum_{s=1}^{t} u_s + y_0 = \alpha_T t + \xi_t + y_0, \tag{21}$$

where $\alpha_T = \mu T^{-\eta}$ with $\mu \neq 0$, $\xi_t = \sum_{s=1}^{t} u_s$, and $u_t = \psi(L)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$ with $\sum_{j=0}^{\infty} j|\psi_j| < \infty$ and $\varepsilon_t \overset{\text{i.i.d.}}{\sim} (0, \sigma^2)$ with finite fourth moment. The fitted slope regression estimator is

$$\hat{\alpha}_T = \sum_{t=1}^{T} t y_t \Big/ \sum_{t=1}^{T} t^2 = \alpha_T + \sum_{t=1}^{T} t(\xi_t + y_0) \Big/ \sum_{t=1}^{T} t^2,$$

or in the fitted intercept case $\breve{\alpha}_T = \sum_{t=1}^{T} \tilde{t} y_t / \sum_{t=1}^{T} \tilde{t}^2 = \alpha_T + \sum_{t=1}^{T} \tilde{t}(\xi_t + y_0) / \sum_{t=1}^{T} \tilde{t}^2$, where $\tilde{t} = t - T^{-1} \sum_{s=1}^{T} s$. For $y_0 = o_p(\sqrt{T})$, we have by standard methods

$$\Theta_T := \sqrt{T}(\hat{\alpha}_T - \alpha_T) \sim \frac{T^{-5/2} \sum_{t=1}^{T} t\xi_t}{T^{-3} \sum_{t=1}^{T} t^2} \xrightarrow{L} 3\psi(1)\sigma \int_0^1 W(s)s \, \mathrm{d}s =: \Theta,$$

$$\breve{\Theta}_T := \sqrt{T}(\breve{\alpha}_T - \alpha_T) \sim \frac{T^{-5/2} \sum_{t=1}^{T} \tilde{t}\xi_t}{T^{-3} \sum_{t=1}^{T} \tilde{t}^2} \xrightarrow{L} 12\psi(1)\sigma \int_0^1 W(s)\left(s - \frac{1}{2}\right) \mathrm{d}s =: \breve{\Theta}$$

so that $\hat{\alpha}_T - \alpha_T \xrightarrow{p} 0$ and $\breve{\alpha}_T - \alpha_T \xrightarrow{p} 0$. Hence, both $\hat{\alpha}_T$ and $\breve{\alpha}_T$ are $\sqrt{T}$ consistent. The following derivations are identical for both these estimates so we confine attention to $\breve{\alpha}_T$. Take an expanded probability space where $\breve{\Theta}_T \xrightarrow{p} \breve{\Theta}$. In this space we have $\breve{\Theta}_T = \breve{\Theta} + o_p(1)$, and then

$$\breve{\alpha}_T = \alpha_T + T^{-1/2}\breve{\Theta} + o_p(T^{-1/2}). \tag{22}$$

This implies

$$\log|\breve{\alpha}_T| = \log|\alpha_T + T^{-1/2}\breve{\Theta} + o_p(T^{-1/2})|$$
$$= \log\left|\frac{\mu}{T^\eta} + T^{-1/2}\breve{\Theta} + o_p(T^{-1/2})\right|$$
$$= -\eta\log T + \log|\mu| + \frac{T^\eta}{\mu T^{1/2}}\breve{\Theta} + o_p\left(\frac{1}{T^{1/2-\eta}}\right).$$

Hence

$$\frac{\log|\breve{\alpha}_T|}{\log T} = -\eta + \frac{\log|\mu|}{\log T} + \frac{1}{\mu T^{1/2-\eta}\log T}\breve{\Theta} + o_p\left(\frac{1}{T^{1/2-\eta}\log T}\right).$$

We therefore have the rate estimator

$$\breve{\eta}_T = -\frac{\log|\breve{\alpha}_T|}{\log T} \xrightarrow{p} \eta, \tag{23}$$

and

$$T^{1/2-\eta}\log T\left(\breve{\eta}_T - \eta + \frac{\log|\mu|}{\log T}\right) \xrightarrow{L} -\breve{\Theta}/\mu. \tag{24}$$

Both equations (23) and (24) also hold in the original probability space. Note that $\breve{\eta}_T$ is consistent for all $0 \leqslant \eta \leqslant 0.5$ and all $\mu \neq 0$. However, equation (24) indicates that $\breve{\eta}_T$ is consistent but has a second order bias of $-(\log|\mu|/\log T)$. Similar results hold in the no intercept case. In particular, using $\hat{\alpha}_T$ we have the rate estimator

$$\hat{\eta}_T = -\frac{\log\left|\hat{\alpha}_T\right|}{\log T} \xrightarrow{p} \eta, \tag{25}$$

and the corresponding limit distribution is

$$T^{1/2-\eta}\log T\left(\hat{\eta}_T - \eta + \frac{\log|\mu|}{\log T}\right) \xrightarrow{L} -\Theta/\mu. \tag{26}$$

Note that $\breve{\eta}_T$ and $\hat{\eta}_T$ remain consistent for $\eta = \frac{1}{2}$ but at a slow logarithmic rate and with appropriately modified limit variates $\breve{\Theta}$ and Θ.

If the model has no drift component, $y_t = \sum_{s=1}^t u_s + y_0 = \xi_t + o_p(\sqrt{T})$. Similarly, when $\eta > 0.5$, we have $y_t = (\mu/T^\eta)t + \sum_{s=1}^t u_s = \xi_t + o_p(\sqrt{T})$ and in place of equation (22) $\breve{\alpha}_T = \sum_{t=1}^T \tilde{t}(\xi_t + y_0)/\sum_{t=1}^T \tilde{t}^2 = T^{-1/2}\breve{\Theta} + o_p(T^{-1/2})$ in the enlarged space. Then,

$$\log|\breve{\alpha}_T| = \log|T^{-1/2}\breve{\Theta} + o_p(T^{-1/2})| = -\frac{1}{2}\log T + \log|\breve{\Theta}| + o_p(1),$$

and it follows that $\breve{\eta}_T = -(\log|\breve{\alpha}_T|/\log T) \to_p \frac{1}{2}$. The same result holds for $\hat{\eta}_T$ and when $\eta = \frac{1}{2}$. In this case both $\breve{\eta}_T$ and $\hat{\eta}_T$ consistently estimate the dominant stochastic trend order (0.5).

Rate estimators like $\hat{\eta}_T$ and $\breve{\eta}_T$ are affected by scaling just like intercepts and deterministic trends, none of which are invariant under scale transformations of the data. In the present case, the intercept $\alpha_T = \mu/T^\eta$ is influenced by two parameters. As the above analysis shows, the rate parameter $\eta < \frac{1}{2}$ can be consistently estimated and the limit theory equations (24) and (26) show that there is a second order bias effect in the limit distribution that involves $\log|\mu|/\log T$. The specification of the weak trend parameter $\alpha_T = \mu/T^{\eta}$ means that data scaling potentially affects both parameters (μ, η). Normalization such as $|\mu| = 1$ means that scale effects are either explicitly or implicitly carried by estimators of η like $\hat{\eta}$: explicitly if we incorporate the effect directly into the parameterization by changing η (for given T) to preserve the normalization[9] or implicitly in terms of the bias effect when η is estimated, as shown in the limit theory. These potential effects need to be considered in interpreting rate estimates.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. The finite sample distribution of ADF and proofs.

[9] For example, given T and setting $\alpha_T = \mu/T^{\eta} = 1/T^{\eta+\delta}$ implies $\delta = -(\log\mu/\log T)$ and normalization effectively transforms $\eta \longmapsto \eta - (\log\mu/\log T) = \eta'$ so that scale effects are absorbed into the parameter η.