ECONOMETRIC ANALYSIS OF ASSET PRICE BUBBLES

By

Shuping Shi and Peter C. B. Phillips

June 2022

COWLES FOUNDATION DISCUSSION PAPER NO. 2331

COWLES FOUNDATION FOR RESEARCH IN ECONOMICS
YALE UNIVERSITY
Box 208281
New Haven, Connecticut 06520-8281

http://cowles.yale.edu/

Econometric Analysis of Asset Price Bubbles*

Shuping Shi[†] and Peter C.B. Phillips[††]

[†]Macquarie University [††]Yale University, University of Auckland, University of Southampton, Singapore Management University

Abstract

In the presence of bubbles, asset prices consist of a fundamental and a bubble component, with the bubble component following an explosive dynamic. The general idea for bubble identification is to apply explosive root tests to a proxy of the unobservable bubble. Three notable proxies are the real asset prices, log price-payoff ratios, and estimated non-fundamental components. The rationale for all three proxy choices rests on the definition of bubbles, which has been presented in various forms in the literature. This chapter provides a theoretical framework that incorporates several definitions of bubbles (and fundamentals) and offers guidance for selecting proxies. For explosive root tests, we introduce the recursive evolving test of Phillips et al. (2015b,c) along with its asymptotic properties. This procedure can serve as a real-time monitoring device and has been shown to outperform several other tests. Like all other recursive testing procedures, the PSY algorithm faces the issue of multiplicity in testing that contaminates conventional significance values. To address this issue, we propose a multiple-testing algorithm to determine appropriate test critical values and show its satisfactory performance in finite samples by simulations. To illustrate, we conduct a pseudo real-time bubble monitoring exercise in the S&P 500 stock market from January 1990 to June 2020. The empirical results reveal the importance of using a good proxy for bubbles and addressing the multiplicity issue.

Keywords: Bubbles; econometrics identification; market fundamental; explosive root; multiplicity; S&P 500 composite index

1 Introduction

The history of financial bubbles demonstrates the potential of these financial phenomena to inflict harm on real economic activity. During the expansion phase of a financial bubble

*Shi acknowledges support from the Australian Research Council Discovery Projects funding scheme (project No.: DP190102049). Phillips acknowledges support from the Kelly Fund at the University of Auckland and the NSF under Grant No. SES 18-50860. Shuping Shi email: shuping.shi@mq.edu.au. Peter C. B. Phillips email: peter.phillips@yale.edu.

a typical consequence is resource misallocation as funds are directed towards asset speculation rather than productive enterprises (Hirano et al., 2015; Caballero and Krishnamurthy, 2006). This process of diversion raises the risk to lending institutions of foreclosure in the event of a bubble collapse and the risk of economic downturn or even general economic crisis in its aftermath (Case et al., 2000). The demise of the hedge fund Long Term Capital Management in 1998 during the Russian financial crisis, the Dot-Com bubble collapse over 2001-2002, and the subsequent 2008 subprime mortgage crisis with the bankruptcy of Lehman Brothers and the advent of the great recession provide three recent examples. In recognition of the major economic implications of financial bubbles, policymakers have emphasized the urgency of research designed to deepen understanding of financial bubble phenomena and to develop measures to address them in order to avoid further crises or to ameliorate their consequences and to maintain economic and financial stability. Correspondingly, there has been a significant surge of academic interest in financial bubbles among economists and in the development of econometric tools to facilitate bubble identification, particularly in real time so that policy measures may be implemented to avert the more serious economic and financial consequences of collapse. There is now a wide literature and for recent econometric research in the field readers are directed to Phillips and Yu (2011); Phillips et al. (2011); Homm and Breitung (2012); Phillips et al. (2015b,c); Lee and Phillips (2016); Harvey et al. (2013, 2016, 2020); Phillips and Shi (2018, 2020); Shi and Phillips (2020); Chen et al. (2020); Pavlidis et al. (2016, 2017); Wang and Yu (2019); Hu and Oxley (2017), among others.

There is a general agreement among economists seeking to model asset prices that in the presence of bubbles asset prices are driven by two separate components, one involving economic fundamentals that relate to the activity and another, the bubble component, carrying the impact of speculative interest. Both these components are theoretical constructs and need precise formulation for empirical research, typically related to particular applications. Various definitions appear in the literature. For example, the definitions presented in Shi (2017) and Shi and Phillips (2020) for housing markets are derived from underlying accounting identities involving returns. These are often expressed in logarithmic form, with the fundamental component being a function of payoffs received from the particular asset (e.g., dividends in the case of equities or rents in the case of housing) and the real interest rates. On the other hand, the definitions used in Phillips et al. (2011) and Phillips et al. (2015b) are based on the present value identity assuming a constant discount factor and expressed in either levels or logarithmic form and the fundamental component is related only to the payoffs of the asset. The relationships among the various definitions in use in empirical research are not always apparent. This chapter provides a framework that embodies many of these definitions and reveals their connections.

A major challenge for bubble identification is the unobservability of the bubble component in asset prices. This component has complex dynamics, exhibiting explosive behavior during the expansion phase followed by decline or collapse in the bubble downturn. The most common solution to the problem of the unobservable bubble component is to monitor its impact through the analysis of realized asset prices or prices adjusted for some observed returns. During the expansionary phase, explosive unit root tests may be applied to simple proxies of the bubble component such as real asset prices or log price-payoff ratios. This approach is successful because in the presence of an expansionary bubble its explosive dynamic typically dominates the time series behavior of asset prices. Nonetheless, the presence of market fundamentals, which are commonly believed to be at most I(1), can compromise the performance of explosive root tests on these simple proxies for the bubble component (Shi and Phillips, 2020). An alternative solution is to estimate the fundamentals separately and remove them from the proxies before applying explosive root tests. While the first approach is easy to implement and has enjoyed widespread use in the literature, the second approach has been investigated only recently (Shi, 2017; Shi et al., 2020; Shi and Phillips, 2020) and has not attracted as much attention, especially in equity markets. The empirical application of the present work applies both approaches to the S&P 500 stock market and reveals the importance of controlling for market fundamentals in identifying bubbles in equity markets.

This chapter focuses on the econometric technique of bubble detection that utilizes the recursive evolving explosive root test of Phillips et al. (2015b,c, PSY hereafter). The PSY procedure utilizes historical information and by applying the detector to the data as it evolves through the observed sample the procedure serves as a real-time monitoring device as well as a general purpose detector. The approach also accounts for the periodically collapsing feature of bubbles and provides consistent estimates for bubble origination and termination dates even when there are multiple bubble episodes within the same sample period. This procedure has been shown to outperform several alternative methods (Phillips et al., 2011; Homm and Breitung, 2012) and has been applied to many different markets, including equity, real estate, and commodity markets. See Wöckl (2019) and Shi and Phillips (2020) for recent surveys of the field.

Although all statistical tests have a non-zero probability of drawing false-positive conclusions, a complex multiplicity issue arises when a particular test is implemented many times on the same data set, as it is in recursive investigations of episodic financial bubbles and their initiation and termination dates. In such applications the probability of making at least one false-positive conclusion rises rapidly toward 100% as the number of tests increases. In the content of real-time bubble monitoring, the PSY procedure is implemented each period as new information arrives. So the multiplicity issue cannot be ignored in empirical testing for financial bubbles. We illustrate the problem by Monte Carlo simulations. To address the issue we propose the use of a multiple-testing algorithm for the computation of suitable critical values for the PSY test. With the use of multiple-testing critical values, the performance of the real-time bubble monitoring procedure is shown to be satisfactory in terms of its family-wise error rates under the null hypothesis of no bubbles and to be successful in its detection rates and estimation accuracy of bubble origination and termination dates under the alternative of bubbles being present in the data.

The rest of the chapter is organized as follows. Section 2 discusses bubble definitions.

Section 3 outlines bubble identification strategies. The PSY explosive root test and its asymptotic properties are introduced in Section 4. Section 5 overviews the multiplicity issue that arises in recursive testing, introduces the multiple-testing algorithm for critical value computation, and explores the finite sample performance of the associated bubble monitoring procedure. The empirical application to the the S&P 500 is reported in Section 6. Section 7 concludes.

2 Bubble Definitions

A standard theory-based approach to the definition of a rational bubble starts from the accounting identity of asset returns (R_{t+1}) over the period of $[t, t+1]$ as

$$1 + R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t}, \tag{1}$$

where P_t is the price of a financial asset and D_t is the payoff received from the asset, such as rental income for houses and dividend income for stocks. Taking conditional expectations at period t and re-organization of equation (1) leads to the standard no arbitrage condition

$$P_t = \rho_t \mathbb{E}_t \left(P_{t+1} + D_{t+1} \right) \text{ with } \rho_t = \frac{1}{1 + \mathbb{E}_t \left(R_{t+1} \right)}, \tag{2}$$

where the quantity $\rho_t \in (0,1)$ is typically referred to as the (time-varying) discount factor.

2.1 Log Linear Approximation

Let $p_t = \log(P_t)$, $d_t = \log(D_t)$, and $r_t = \log(1 + R_t)$. The return identity (1) can be rewritten using notations p_t, d_t and r_t as

$$r_{t+1} = \log(P_{t+1} + D_{t+1}) - \log P_t = p_{t+1} - p_t + \log(1 + e^{d_{t+1} - p_{t+1}}). \tag{3}$$

Applying a Taylor series expansion of $\log(1 + e^{d_{t+1} - p_{t+1}})$ at the sample mean of $p_t - d_t$ (i.e., $\overline{p-d}$) as in Campbell and Shiller (1988), we obtain the (approximate) relation

$$r_{t+1} = \kappa + \delta p_{t+1} + (1-\delta) d_{t+1} - p_t, \tag{4}$$

where

$$\delta = \frac{1}{1 + \exp(\bar{d} - \bar{p})} \quad \text{and} \quad \kappa = -\delta \log \delta - (1-\delta) \log(1-\delta).$$

By recursive substitution of (4) and taking conditional expectations with respect to information available at period t, we can write the difference $p_t - d_t$ in the form of a sum of a fundamental component f_t and a bubble component b_t as

$$p_t - d_t = f_t + b_t, \tag{5}$$

where the two components are written as

$$f_t = \frac{\kappa}{1-\delta} + \sum_{k=0}^{\infty} \delta^k \mathbb{E}_t(\Delta d_{t+1+k} - r_{t+1+k}), \tag{6}$$

$$b_t = \lim_{i\to\infty} \delta^i \mathbb{E}_t(p_{t+i}). \tag{7}$$

Equations (6) and (7) provide convenient definitions for the market fundamental component and the rational bubble component of the log price dividend ratio in (5).

Fundamentals f_t are determined in terms of a discounted present value involving dividend growth and returns. The bubble component involves the eventual discounted present expectation of the asset price. In the absence of a bubble, the transversality condition holds and $b_t = \lim_{i\to\infty} \delta^i \mathbb{E}_t(p_{t+i}) = 0$, so that eventual discounted expected value is zero. In the presence of speculative sentiment, the transversality condition does not apply and the bubble component b_t satisfies the sub-martingale property

$$\mathbb{E}_t\left(b_{t+1}\right) = \lim_{i\to\infty} \mathbb{E}_t\left(\delta^i p_{t+1+i}\right) = \frac{1}{\delta} \lim_{i\to\infty} \mathbb{E}_t\left(\delta^{i+1} p_{t+1+i}\right) = \frac{1}{\delta} b_t. \tag{8}$$

Since $1/\delta > 1$, equation (8) reveals that the speculative bubble component b_t is expected to be larger in the next period than it is today, i.e., $\mathbb{E}_t\left(b_{t+1}\right) > b_t$. In other words, the variable b_t follows an explosive process. Many of the latest econometric techniques for bubble identification are based on this unique explosive feature of bubbles. See, e.g., Diba and Grossman (1988); Phillips et al. (2011, 2015b); Pavlidis et al. (2016).

The log linear approximation of $\log(1+e^{d_{t+1}-p_{t+1}})$ typically relies on the presumption that the sample mean $\overline{p-d}$ converges to the population mean of $p_t - d_t$. In the case that p_t and d_t are integrated and with a cointegration vector $[1,-1]$, the approximation is valid. In the presence of bubbles, $p_t - d_t$ is itself explosive and the log linear approximation is only valid under certain conditions. Explosive behavior in the time series $p_t - d_t$ leads to an explosive sample mean $\overline{p-d}$ so that $\delta = \frac{1}{1+\exp(\overline{d-p})} \to 1$, thereby potentially compromising the validity of the representation (6). Interestingly, Lee and Phillips (2016) show that both the log linear approximation and the present value identity (5) remain valid as long as the duration of the bubble is asymptotically negligible.

2.2 Constant Discount Factor

It is often convenient to take the special case of a constant discount factor, i.e., $\rho_t = \rho \in (0,1)$. By recursive substitution, equation (2) then becomes

$$P_t = \sum_{i=0}^{\infty} \rho^i \mathbb{E}_t\left(D_{t+i}\right) + \lim_{i\to\infty} \mathbb{E}_t\left(\rho^i P_{t+i}\right). \tag{9}$$

In this event, fundamentals and bubbles are sometimes defined as the two components

$$F_t = \sum_{i=0}^{\infty} \rho^i \mathbb{E}_t \left(D_{t+i} \right) \text{ and } B_t = \lim_{i \to \infty} \mathbb{E}_t \left(\rho^i P_{t+i} \right). \tag{10}$$

Since the discount rate $\rho \in (0,1)$, B_t continues to satisfy a submartingale property where

$$\mathbb{E}_t \left(B_{t+1} \right) = \frac{1}{\rho} B_t. \tag{11}$$

Further, under a constant discount factor we have $\mathbb{E}_t(r_{t+1+k}) = -\log \rho$ by virtue of the definition of ρ_t and r_{t+1}. We can then write equation (6) in simpler form as

$$f_t = \frac{\kappa + \log \rho}{1 - \delta} + \sum_{k=0}^{\infty} \delta^k \mathbb{E}_t(\Delta d_{t+1+k}), \tag{12}$$

which is another commonly used definition of fundamentals. In the absence of bubbles and when the growth rate of dividends Δd_{t+1+k} is stationary, then $p_t - d_t = f_t$ is I(0). For instance, if Δd_t is stationary autoregressive, satisfying $\Delta d_t = \alpha \Delta d_{t-1} + \zeta_t$ with $|\alpha| < 1$ and $\zeta_t \sim iid(0, \sigma_\zeta^2)$, then

$$f_t = \frac{\kappa + \log \rho}{1 - \delta} + \frac{\alpha}{1 - \alpha\delta} \sum_{s=0}^{\infty} \alpha^s \zeta_{t-s}, \tag{13}$$

is a stationary linear autoregressive process. So p_t and d_t are cointegrated. But in the presence of a bubble process b_t, the dynamics of $p_t - d_t$ are dominated by b_t and hence become explosive. Evidence of explosive dynamics in the log price-to-rent ratio can therefore serve as a signal for the presence of a speculative bubble component in the generating mechanism.

2.3 Time-varying Discount Factor

The log gross return r_{t+1} is assumed to equal the return of a baseline asset g_{t+1} plus a risk premium φ_{t+1} (Campbell and Shiller, 1988; Shi, 2017), so that

$$r_{t+1} = g_{t+1} + \varphi_{t+1}.$$

The baseline return can be 3-month treasury bill rates in the case of equities or real mortgage interest rates in the case of housing. It is further assumed that the time-varying risk premium fluctuates around a long-term risk premium level φ (Shi, 2017; Shi et al., 2020), so that

$$\varphi_{t+1} = \varphi + \varepsilon_{t+1},$$

where ε_{t+1} is a zero mean disturbance. The expected log return of the asset is the sum of the expected return of the baseline asset and the expected long-run risk premium, i.e.,

$$\mathbb{E}_t\left(r_{t+1}\right) = \varphi + \mathbb{E}_t\left(g_{t+1}\right).$$

By definition, the logarithmic discount factor is $\log \rho_t = -\mathbb{E}_t\left(r_{t+1}\right)$, and under this assumed model structure, the fundamental component in (6) then becomes

$$f_t = \frac{\kappa - \varphi}{1-\delta} + \sum_{k=0}^{\infty} \delta^k \mathbb{E}_t(\Delta d_{t+1+k} - g_{t+1+k}), \tag{14}$$

which serves as the basis for the fundamental-bubble decomposition used in Section 3.2.

2.4 Collapsing Bubbles

Historical evidence indicates that bubbles are typically transitory with alternate stages of expansion and collapse in asset prices. Various models have been proposed to capture such dynamics. Examples are the early probabilistic bubble generating mechanisms of Blanchard and Watson (1982) and Evans (1991) and a recent behavioral model of Phillips (2016). In Blanchard and Watson (1982) bubbles are generated according to the model

$$B_{t+1} = \begin{cases} (\pi\rho)^{-1} B_t + \varepsilon_{t+1} & \text{with probability } \pi \\ \varepsilon_{t+1} & \text{with probability } 1-\pi \end{cases}, \tag{15}$$

where ε_{t+1} is an error term. In this model the probability of a bubble in each time period is set *a priori* to π and the probability of a bubble collapse is $1-\pi$. The expansion rate of the bubble is the autoregressive parameter $(\pi\rho)^{-1}$ during the bubble expansion phase. The conditional expectation of B_{t+1} is given by

$$\mathbb{E}_t\left(B_{t+1}\right) = \mathbb{E}_t\left\{\pi\left[(\pi\rho)^{-1} B_t + \varepsilon_{t+1}\right] + (1-\pi)\,\varepsilon_{t+1}\right\} = \rho^{-1} B_t,$$

So, the bubble generating process satisfies the submartingale property $\mathbb{E}_t\left(B_{t+1}\right) > B_t$.

Unlike Blanchard and Watson (1982), Evans (1991) considers a bubble model with three stages. At the initial stage, the bubble expands at rate of ρ^{-1}. After exceeding the threshold b, the expansion accelerates with a rate of $(\pi\rho)^{-1}$ and during this period faces a probability of collapse. The collapse mechanism is governed by a Bernoulli process θ_t, which takes the value one with probability π and zero otherwise. More specifically,

$$B_{t+1} = \begin{cases} \rho^{-1} B_t \varepsilon_{B,t+1} & \text{if } B_t < b \\ \left[\upsilon + (\pi\rho)^{-1}\theta_{t+1}\left(B_t - \rho\upsilon\right)\right]\varepsilon_{B,t+1} & \text{if } B_t \geq b \end{cases}. \tag{16}$$

The error term in this model is multiplicative, rather than additive, and takes the form of $\varepsilon_{B,t+1} = \exp\left(v_{t+1} - w^2/2\right)$ with $v_{t+1} \overset{i.i.d}{\sim} \mathcal{N}\left(0, w^2\right)$, which ensures that

$$\mathbb{E}_t\left(B_{t+1}\right) = \rho^{-1} B_t.$$

The approach taken in Phillips (2016) constructs a behavioral model of asset markets that produces episodes of exuberance, collapse and recovery. These episodes emerge because of subjective investor-based rational expectations about the impact of fundamentals on the market price. Investors are assumed to have heterogeneous sentiments about the market, allowing them to be exuberant or cautious or strict followers of fundamentals. This is achieved via boundary conditions that describe their respective views of the expected market impact of the same economic fundamentals. Equilibrium solution paths take multiple forms, depending on the dominating market sentiment. This rational expectations model of asset pricing produces market bubbles when exuberant sentiment dominates and then gives way to cautionary and negative sentiment.

3 Bubble Identification

As indicated earlier a significant challenge to empirical bubble identification is the unobservability of the bubble component in the decomposition of asset prices into bubble and fundamental components. The problem is in some respects analogous to another common time series identification problem, the separation of trend and cycle components in nonstationary data. Various econometric approaches have been proposed to address this challenge of unobservable components and the main ideas are reviewed in what follows.

3.1 Strategy I: Use Simple Proxies

Consider the fundamental and bubble definitions given in equations (9)-(10). In the absence of bubbles, the transversality condition $\lim_{i\to\infty} \mathbb{E}_t \left(\rho^i P_{t+i}\right) = 0$ is assumed to hold and the asset price in (9) is at most I(1), as the dividend process D_t is generally believed to be integrated of order one. On the other hand, in the presence of speculation, P_t is dominated by the bubble component and accordingly follows an explosive process during the expansionary phase of the bubble. Similarly, from (5)-(7), $p_t - d_t$ is at most I(1) in the absence of speculative behavior but becomes explosive when speculation overtakes the market. An immediate solution to the unobservable bubble component problem is to apply explosive root tests to P_t or $p_t - d_t$ for bubble identification directly in the observable price or price/divident ratio data. The approach seeks to detect the explosive dyamics in P_t or $p_t - d_t$ that is induced by the underlying bubble component.

The accuracy of this approach to bubble identification can be compromised when the fundamental component (which is taken as residual noise in this approach) is highly persistent. Since the fundamental process F_t is a weighted linear function of the dividend process D_{t+i}, which is itself generally believed to be $I(1)$, F_t can be expected to manifest persistence. The fundamental component f_t of $p_t - d_t$ given in (13) is related to the growth rate of dividends Δd_{t+1+k}, which may be taken as stationary. In other words, under the constant discount factor assumption, f_t in (13) is less persistent than F_t. Therefore, we focus on $p_t - d_t$ for subsequent discussions on bubble identification. Nevertheless, when the discount factor is allowed to be time-varying, the fundamental of $p_t - d_t$ in (14) is also related to returns of a baseline asset, which may well not be stationary. In fact, most time series of real interest rates are either highly persistent or fractionally integrated, and may be considered nonstationary (Rose, 1988; Mishkin, 1995; Rapach and Weber, 2004; Sun and Phillips, 2004; Phillips, 2005) .

3.2 Strategy II: Decomposition

To control the impact of market fundamentals on tests designed to detect explosive root behavior in time series such as $p_t - d_t$, Shi (2017) proposed to remove the residual noise element by decomposing the observed series $p_t - d_t$ into fundamental and non-fundamental components. The explosive root tests could then be conducted directly on the empirically estimated non-fundamental component.

Two different approaches have been suggested for performing this decomposition. Campbell and Shiller (1988) introduced a structural decomposition method based on (14). The decomposition involves three steps: (i) forecasting the future stream $\Delta d_{t+1+k} - g_{t+1+k}$ using a VAR model; (ii) calibrating parameters κ, φ, and δ; and (iii) combining the results from the first two steps in equation (14) to obtain an estimate of f_t. This approach requires asset prices and dividends or rents data to be available in their actual dollar values.

The second approach is a reduced form method proposed recently in Shi and Phillips (2020). Working with the reduced form involves fewer steps than the structural approach, which has the added advantage of reducing the potential impact of estimation or calibration errors on the final outcome. Additionally, the price and dividend or rent series can be measured in either index form or dollar values, which significantly improves the feasibility of this approach in practice. The following discussion gives an overview of the method.

The reduced form econometric model of the key variables is assumed to have the following general form

$$y_t = \alpha + \beta X_t + \varepsilon_t, \tag{17}$$

$$X_t = d_X + \rho_X X_{t-1} + u_{Xt} \tag{18}$$

where y_t is the first difference of $p_t - d_t$ and the regressor X_t includes payoffs of the asset, d_t, returns of a baseline asset, g_t, and other relevant observable fundamentals. The regressor X_t is $k \times 1$ and its components are assumed to have varying unknown degrees of persistence, including I(1), near I(1), or forms of mild stationarity. A convenient way of formulating this is to assume that the autoregressive coefficient matrix in (18) allows for such possibilities via a general specification of the form

$$\rho_X = I_k + \frac{C_X}{T^\alpha}, \tag{19}$$

where I_k is a $k \times k$ identity matrix, $C_X = \text{diag}(c_{X1}, ..., c_{Xk})$ with $c_{Xi} \leq 0$, and $\alpha \in (0, 1]$. The errors ε_t in the regression equation (17) are assumed to be generated by the following mechanism

$$\varepsilon_t = \sum_{i=1}^{q} \phi_i \varepsilon_{t-i} + \sigma_t \eta_t \tag{20}$$

$$\sigma_t^2 = c_0 + \sum_{l=1}^{m} \omega_l \varepsilon_{t-l}^2 + \sum_{j=1}^{n} \gamma_j \sigma_{t-j}^2, \tag{21}$$

where $\eta_t \overset{i.i.d}{\sim} (0,1)$, allowing for both weak dependence and conditional heteroskedasticity. The simple case of no serial correlation and homoskedasticity is included when $\phi_i = \omega_l = \gamma_j = 0$ for all (i,l,j).

The system (17)-(18) is similar to typical predictive regression formulations. But the present goal is to estimate the fundamental component rather than to forecast $p_t - d_t$. It is therefore important in the formulation not to include variables that are contaminated by speculative behavior in the regression. For example, consumer sentiment, which measures how optimisitc or pessimistic consumers feel about the market, could be one of the driving forces behind speculative market behavior. Further, the time-varying risk premium (i.e., $\varphi_{t+1} = r_{t+1} - g_{t+1}$) is a linear function of asset returns r_{t+1} which may itself embody consequences of a bubble component. So, neither consumer sentiment nor time-varying risk premia are considered for inclusion in the regression. On the other hand, the real interest rate, which reflects borrowing costs, is believed to be a key fundamental driving factor for asset prices, in addition to expected future payoffs. For housing markets, prices may be related to other variables that affect housing demand (e.g., population, immigration, employment, personal income) and housing supply (Shi, 2017). Even though most of these variables manifest varying degrees of persistence, the specification (18)-(19) for X_t suitably accommodates them all.

When regressors are highly persistent, it is well known that least squares estimation leads to bias and non-standard limit theory, invalidating inference by conventional t and Wald tests. Endogeneity is a further complication, typically arising from joint determination of y_t and X_t, omitted variables, or errors in variables. Various methods have been proposed to address these problems but most of these are practical only in the case of a scalar X_t, which is unrealistic in most applications. One method that addresses all of these complications is the IVX instrumental variable estimation method, proposed in Phillips and Magdalinos (2009). This method has many advantages, a particularly important one in many applications being its use of endogenous instruments, which avoids the need to find external instruments that satisfy suitable orthogonality and relevance conditions. Kostakis et al. (2015) extended this method to account for potential heteroskedasticity and a further extension to deal more successfully with serially correlated errors than the original IVX method was given by Yang et al. (2020).

The basic idea of the IVX method is to construct instrumental variables from X_t that are less persistent than X_t when X_t is highly persistent. For completeness, details of the

IVX estimation method and the internal instrument construction process are provided in Appendix A. The following discussion shows how the method is implemented in the present context.

Let $\widehat{\alpha}$ and $\widehat{\beta}$ be the IVX estimates of α and β in (17) and the corresponding fitted regression equation be

$$\widehat{y}_t = \widehat{\alpha} + \widehat{\beta} X_{t-1},$$

so that $\widehat{y}_t$ is the regression estimate of y_t, the first difference of $p_t - d_t$. Using this regression equation, the estimated market fundamental component, $\hat{f}_t$, and estimated non-fundamental component, $\hat{e}_t$, are obtained as follows

$$\hat{f}_t = \hat{f}_1 + \sum_{i=2}^{t} \hat{y}_t \quad \text{and} \quad \hat{e}_t = p_t - d_t - \hat{f}_t \tag{22}$$

so that $\hat{f}_t$ is a partial sum of the regression fitted values $\widehat{y}_t$ taken from the initialization $\hat{f}_1 = p_1 - d_1$, thereby providing an estimate of the fundamental component of $p_t - d_t$. The residual component $\hat{e}_t = p_t - d_t - \hat{f}_t$ therefore contains estimation errors as well as components that cannot be explained by the fundamentals. The residuals $\hat{e}_t$ can be expected to mirror explosive behavior in the bubble component of $p_t - d_t$ when bubbles are present in the data. Having removed the fundamental component $\hat{f}_t$ from $p_t - d_t$, the residual term $\hat{e}_t$ should be less contaminated by noise from fundamentals and serve as a better proxy for the bubble component.

4 The PSY Explosive Root Test

The overarching goal of bubble identification procedures is to detect the presence of speculative behavior in each period using only historical data. The PSY explosive root test is applied to historical data x_t, where x_t constitutes different time series data depending on the particular bubble identification strategy. For the two strategies considered above we have

$$x_t = \begin{cases} p_t - d_t & \text{Strategy I} \\ \hat{e}_t & \text{Strategy II} \end{cases}.$$

The null hypothesis is that there is no bubble present in period t and the alternative hypothesis is that observation t carries a speculative bubble component. Since x_t is at most I(1) in the absence of bubbles, the null hypothesis can be translated into a unit root model given by the equation

$$\mathcal{H}_0 : x_t = c_T + x_{t-1} + u_t, \tag{23}$$

where the intercept satisfies $c_T = c_0 T^{-\gamma}$ for some constant c_0, with power coefficient $\gamma > 1/2$, and T being the total number of observations. The setting $\gamma > 1/2$ ensures that the drift induced by the intercept c_T is asymptotically negligible as $T \to \infty$.[1] See Phillips et al. (2014) for more discussion.

Under the alternative, x_t is driven by a bubble mechanism and follows a mildly explosive process (Phillips and Magdalinos, 2007) such that

$$\mathcal{H}_A : x_t = \delta_T x_{t-1} + u_t \text{ with } \delta_T = 1 + cT^{-\alpha}, \tag{24}$$

where $c > 0$ and $\alpha \in (0, 1)$. Detecting the presence of a bubble component in the generating mechanism at period t is therefore equivalent to distinguishing a martingale process (23) (possibly with negligible drift) from a mildly explosive process (24) at a specific data point.

4.1 The PWY Data Generating Process

Blanchard and Watson (1982) and Evans (1991) used stochastic mechanisms for bubble generation in which bubbles are turned on and off according to a certain probabilistic rule (i.e., equations (15) and (16)). Alternative methods, like the reduced form methods considered earlier, involve direct empirical specifications with deterministic break dates that can be estimated econometrically. Several methods of this type have been considered in the literature and employed for the econometric analysis of bubble origination and termination dates (Phillips et al., 2011; Phillips and Yu, 2009; Phillips et al., 2015b,c; Phillips and Shi, 2018).

The data generating process proposed by Phillips et al. (2011) is for x_t and takes the form

$$x_t = \begin{cases} x_{t-1} + \varepsilon_t & t \in [1, t_e) \cup [t_c + 1, T] \\ \delta_T x_{t-1} + \varepsilon_t & t \in [t_e, t_c - 1] \\ X^*_{T_c} & t = t_c \end{cases}, \tag{25}$$

where $\varepsilon_t \overset{i.i.d}{\sim} (0, \sigma^2)$, t_e is the bubble origination date, t_c is the date of bubble collapse, and δ_T is specified as in (24). The variable x_t follows a unit root process in the normal market regimes $t \in [1, t_e) \cup [t_c + 1, T]$ and is mildly explosive with autoregressive coefficient δ_T in the bubble expansion regime $t \in [t_e, t_c - 1]$. The rate of bubble expansion is controlled by magnitude of the power parameter α and that of the localizing coefficient $c > 0$.[2] When the bubble episode collapses it does so in a single period and adjusts to the level around that of the bubble origination level with a small perturbation, i.e., $X^*_{t_c} = X_{t_e} + O_p(1)$.

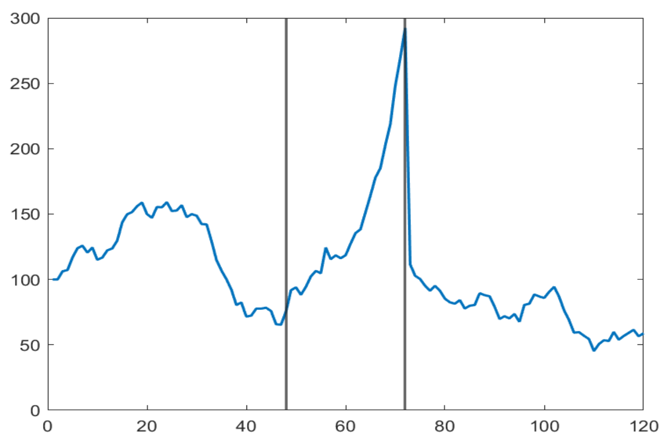
This data generating process can be generalized to allow for multiple bubbles as in Phillips et al. (2015b,c) and to include a small drift in the normal regime or allow for various

[1] Suppose the T equidistant observations span over N years in total. The distance between two observations is Δ, taking values $\Delta = \{1/256, 1/52, 1/12\}$, respectively, for daily, weekly, and monthly observations. We have $T = N/\Delta$ and $c_T = c_0 N^{-\gamma} \Delta^{\gamma}$. In finite samples, the magnitude of the drift depends on the sampling frequency Δ. Suppose the time period N is fixed. The higher the sampling frequency (or equivalently the smaller the value of Δ), the smaller the drift value c_T.

[2] Smaller α and larger c lead to greater values of the autoregressive coefficient $\delta_T > 1$ and hence stronger explosive behavior.

rates of bubble collapse as in Phillips and Shi (2018). Figure 1 displays a typical realization of the data generating process (25). The parameter settings are: $x_0 = 100$, $c = 1$, $\sigma = 6.79$, $\alpha = 0.6$, $t_e = \lfloor 0.4T \rfloor$, $t_c = \lfloor 0.6T \rfloor$, and $T = 120$, as in Phillips et al. (2015b), where the floor function $\lfloor . \rfloor$ delivers the integer part of the argument. As is apparent from the trajectory shown, when the bubble process begins there is little indication of any change in its character or the generating mechanism. But this quickly changes as the bubble takes hold and the explosive behavior becomes manifest. The abrupt collapse at the termination of the bubble is the consequence of the single period transition to the region of the process at the point of bubble origination.

Figure 1: A typical realization of the data generating process (25) proposed in Phillips et al. (2011). The two vertical lines indicate the locations of the bubble origination and bubble termination dates.



4.2 The PSY Recursive Evolving Algorithm

Policy makers, regulators, financial institutions and individual investors are all intensely interested in learning whether market conditions have changed from normal to abnormal behavior. From an econometric perspective, this interest translates into learning from the data whether observations of x_s are generated by a martingale process under normal efficient market conditions as in (23) or whether speculative behavior has overtaken the market with a mildly explosive dynamic as in (24). The recursive algorithms suggested in Phillips et al. (2011, 2015b,c) deliver test procedures that provide a mechanism for answering this important question of whether a subtle change has occurred in the generating process that may ultimately lead to significant changes in market behavior that carry attendant risks of market re-adjustment or collapse.

The PSY procedure employes an evolving recursion to compute the test statistic. The recursion starts by generating a sequence of augmented Dickey-Fuller (ADF) statistics from subsamples of the data. Let t_1 and t_2 be the starting and ending dates of a subsample such that $t_2 > t_1 \geq 1$ giving a window size $t_w = t_2 - t_1 + 1 \geq t_{\min}$, where $t_{\min}$ is the minimum window size required to run the regression. The fitted ADF regression equation is

$$\Delta x_t = \hat{\beta}_0 + \hat{\beta}_1 x_{t-1} + \sum_{j=1}^{K} \hat{\psi}_j \Delta x_{t-i} + \hat{\varepsilon}_t, \tag{26}$$

where t runs from t_1 to t_2, K is the lag order, and $\hat{\varepsilon}_t$ is the residual. The regression is fitted by ordinary least squares. The t-statistic for the significance of $\hat{\beta}_1$ is referred to as the ADF statistic and is denoted ADF_{t_1,t_2}. This statistic is computed in an evolving fashion over the data from varying points of initialization.

The goal in this evolving recursion is to determine whether the data generating process has a bubble component. As attention moves through the sample observations, suppose the current observation of interest is s. In this evolving recursion, the end dates of all the subsamples are fixed on the observation of interest (in this case $t_2 = s$), whereas the starting date varies from the first observation to $s - t_{\min} + 1$. Consider, for example, the sample path displayed in Figure 1 and suppose $s = 60$. The recursive algorithm together with its attendant calculations are illustrated in Figure 2. The PSY test statistic, denoted by $PSY_s(t_{\min})$, is computed as the extremum of the individual ADF statistics ADF_{t_1,t_2}, viz.,

$$PSY_s(t_{\min}) = \max_{\substack{t_2 = s, \\ t_1 \in [1, s - t_{\min} + 1]}} \{ADF_{t_1,t_2}\}.$$

Let $cv_s(\beta_T)$ be the $100\beta_T\%$ critical value of the $PSY_s(t_{\min})$ statistic.[3] The null hypothesis is rejected and observation s is taken as driven by a speculative bubble if the critical value is exceeded, that is if

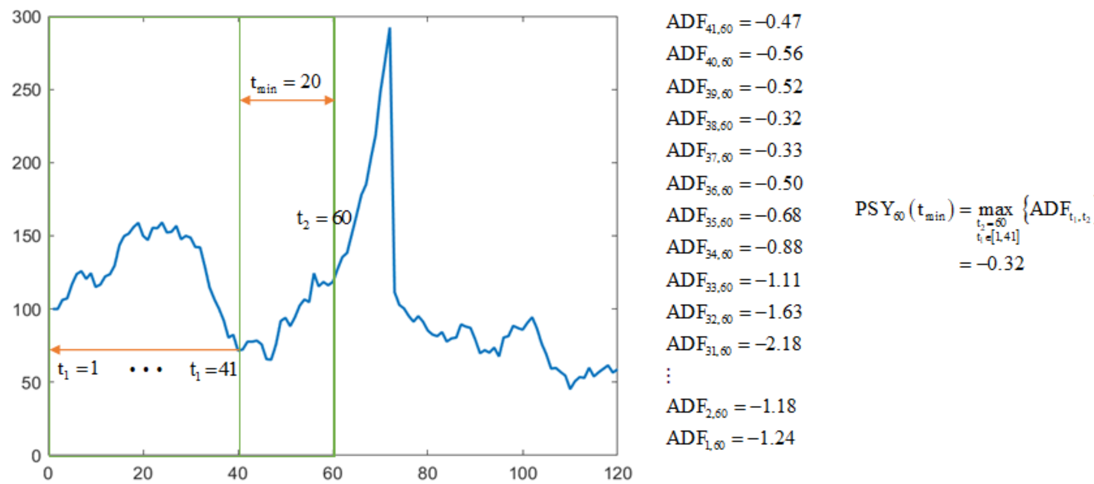
$$PSY_s(t_{\min}) > cv_s(\beta_T).$$

Otherwise, we conclude that the null is not rejected and there is no bubble component in the data at period s. The feasible range for s runs from $t_{\min}$ to T, the lower point of the range $t_{\min}$ being determined by the requirement of a minimum window size for the ADF regressions. The first chronological observation at which the PSY statistic exceeds the critical value is taken as the estimated bubble origination date, so that

$$\hat{t}_e = \inf_{s \in [t_{\min}, T]} \{PSY_s(t_{\min}) > cv_s(\beta_T)\}.$$

[3] Critical values of the PSY statistic are computed numerically by simulation, evolve according to the position of the time period s, and depend on the setting of the minimum window size $t_{\min}$. For ease of notation they are written here as $cv_s(\beta_T)$, although $cv_s(\beta_T, t_{\min})$ would be more precise.

Figure 2: The recursive evolving algorithm: $t_2 = s$ and $t_1 \in [1, s - t_{\min} + 1]$.



In similar fashion, the termination date of the bubble is estimated by the first chronological observation after $\hat{t}_e$ that falls below its corresponding critical value, i.e.,

$$\hat{t}_c = \inf_{s \in [\hat{t}_e, T]} \{PSY_s(t_{\min}) < cv_s(\beta_T)\}.$$

The nuisance parameter $t_{\min}$ is usually set according to the rule recommended in Phillips et al. (2015b) on the basis of extensive simulations, viz., $t_{\min} = \lfloor 0.01T + 1.8\sqrt{T} \rfloor$.

4.3 Asymptotic Theory

The properties of the evolving recursion are analyzed in Phillips et al. (2015b,c) under both the null and alternative hypotheses using asymptotic theory as the sample size $T \to \infty$ and finite sample simulations. This section reviews the limit properties of the PSY procedure and the dating estimators that emerge from it. We use the notation $\Rightarrow$ to signify weak convergence, $:=$ to represent definitional equivalence, $\sim_a$ to indicate distributional equivalence, $\to^p$ to represent convergence in probability, and τ to denote fractional observations (i.e., $\tau_1 = t_1/T$, $\tau_2 = t_2/T$, $\tau_w = t_w/T$, $\tau_{\min} = t_{\min}/T$, $\tau = s/T$, $\tau_e = t_e/T$, and $\tau_c = t_c/T$).

The error term u_t in $\mathcal{H}_0$ may be weakly dependent and is assumed to satisfy the following linear process conditions to facilitate the development of the asymptotics (Phillips and Solo, 1992).

Assumption 4.1 *Let* $u_t = \psi(L)\varepsilon_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$, *where* L *is the lag operator, the summability condition* $\sum_{j=0}^{\infty} j|\psi_j| < \infty$ *holds and* $\{\varepsilon_t\}$ *is an i.i.d. sequence with zero mean, variance* σ^2, *and* $\mathbb{E}|\varepsilon_t|^{4+\delta} < \infty$ *for some* $\delta > 0$.

Under the null hypothesis of (23) and Assumption 4.1, Phillips et al. (2015b) show that

$$ADF_{t_1,t_2} \Rightarrow \frac{\frac{1}{2}\tau_w \left[W(\tau_2)^2 - W(\tau_1)^2 - \tau_w\right] - \int_{\tau_1}^{\tau_2} W(r)\,dr\,[W(\tau_2) - W(\tau_1)]}{\tau_w^{1/2}\left\{\tau_w \int_{\tau_1}^{\tau_2} W(r)^2\,dr - \left[\int_{\tau_1}^{\tau_2} W(r)\,dr\right]^2\right\}^{1/2}} := \Upsilon_{\tau_1,\tau_2},$$

where $W(.)$ is a standard Wiener process. By virtue of continuous mapping arguments, the PSY statistic is found to have the following pivotal asymptotic distribution (Phillips et al., 2015b):

$$\begin{aligned} PSY_s(t_{\min}) &\Rightarrow \sup_{\substack{\tau_2=\tau,\\ \tau_1\in[0,\tau_2-\tau_{\min}]}} \left\{ \frac{\frac{1}{2}\tau_w \left[W(\tau_2)^2 - W(\tau_1)^2 - \tau_w\right] - \int_{\tau_1}^{\tau_2} W(r)\,dr\,[W(\tau_2) - W(\tau_1)]}{\tau_w^{1/2}\left\{\tau_w \int_{\tau_1}^{\tau_2} W(r)^2\,dr - \left[\int_{\tau_1}^{\tau_2} W(r)\,dr\right]^2\right\}^{1/2}} \right\} \\ &:= \Upsilon_\tau(\tau_{\min}). \end{aligned}$$

The asymptotic $100\beta_T\%$ critical value of $PSY_s(t_{\min})$ is taken as the $100(1-\beta_T)^{th}$ percentile of the distribution $\Upsilon_\tau(\tau_{\min})$.

Next, we review limiting properties of the PSY statistic under the data generating process (25), which has a single bubble episode. For ease of notation, we refer to the two normal periods (i.e., $[1,t_e)$ and $[t_c+1,T]$) as N_0 and N_1 and the bubble period $[t_e,t_c-1]$ as B. The asymptotic properties of ADF_{t_1,t_2} depend on the location of the subsample. There are four possible scenarios:

$$\begin{aligned} (1)\ t_1 &\in N_0 \text{ and } t_2 \in N_0, \\ (2)\ t_1 &\in N_0 \text{ and } t_2 \in B, \\ (3)\ t_1 &\in N_0 \text{ and } t_2 \in C, \\ (4)\ t_1 &\in N_0 \text{ and } t_2 \in N_1. \end{aligned}$$

Except for the first episode N_0, all the other cases involve structural breaks. For case (2) the subsample begins in the normal regime N_0 with a structural break introducing the bubble component to the data generating mechanism and the bubble regime B ends just before the collapse period (3). In regime B the regression is dominated by bubble dynamics with a mildly explosive autoregression. In this regime, rather than converging to a random variable as in case (1), the ADF statistic diverges to positive infinity at rate $O_p\left(T^{1-\alpha/2}\right)$. For cases (3) and (4), the sample period includes both expansion and collapse stages of the bubble. The initial strong reverting autoregressive pattern within the subsample causes the ADF statistic to diverge to negative infinity at rate $O_p\left(T^{(1-\alpha)/2}\right)$, which signals the collapse regime C. More specifically, Phillips et al. (2015c) show that under the DPG (25),

$$ADF_{t_1,t_2} \sim_a \begin{cases} \Upsilon_{\tau_1,\tau_2} & t_1 \in N_0 \text{ and } t_2 \in N_0 \\ T^{1-\alpha/2} \frac{\tau_w^{3/2}}{\sqrt{2(\tau_e-\tau_1)}} \to +\infty & t_1 \in N_0 \text{ and } t_2 \in B \\ -T^{(1-\alpha)/2} \sqrt{\frac{1}{2} c\tau_w} \to -\infty & t_1 \in N_0 \text{ and } t_2 \in C \\ -T^{(1-\alpha)/2} \sqrt{\frac{1}{2} c\tau_w} \to -\infty & t_1 \in N_0 \text{ and } t_2 \in N_1 \end{cases} . \tag{27}$$

By continuous mapping arguments the asymptotic behavior of the PSY statistic in the various locations is given by

$$PSY_s\left(t_{\min}\right) \sim_a \begin{cases} \Upsilon_\tau\left(\tau_{\min}\right) & s \in N_0 \\ T^{1-\alpha/2} \sup_{\tau_1 \in [0,\tau-\tau_{\min}]} \left\{ \frac{(\tau-\tau_1)^{3/2}}{\sqrt{2(\tau_e-\tau_1)}} \right\} \to +\infty & s \in B \\ -T^{(1-\alpha)/2} \left(\frac{c\tau}{2}\right)^{1/2} \to -\infty & s \in C \\ -T^{(1-\alpha)/2} \left(\frac{c\tau}{2}\right)^{1/2} \to -\infty & s \in N_1 \end{cases} .$$

The PSY statistic is asymptotically equivalent to a random variable (i.e., $O_p(1)$) when the observation of interest s falls in the normal regime N_0. It diverges to positive infinity at rate $O_p(T^{1-\alpha/2})$ when s is in the bubble regime and diverges to negative infinity when s falls in regimes C and N_1.

Consistency of the bubble origination and termination date estimators, viz., $\hat{t}_e/T \to^p \tau_e$ and $\hat{t}_c/T \to^p \tau_c$, follows under the following simple condition on the respective expansion rate of the critical value $cv_\tau(\beta_T)$ and the power coefficient α that controls the mildly explosive rate in the bubble regime

$$\frac{1}{cv_\tau\left(\beta_T\right)} + \frac{cv_\tau\left(\beta_T\right)}{T^{1-\alpha/2}} \to 0.$$

We assume that $\beta_T \to 0$ and hence $cv_\tau\left(\beta_T\right) \to \infty$ as $T \to \infty$, to eliminate Type I errors asymptotically, and the second condition requires the rate of divergence to be slower than $O_p\left(T^{1-\alpha/2}\right)$.

Readers are referred to Phillips et al. (2015b,c,a) for proofs of these asymptotic results, generalizations that include multiple bubble episodes, and further discussion, including extensive simulation findings of the properties of these procedures in finite samples. Phillips and Shi (2018) provide further asymptotic results under DGPs with various rates of bubble collapse and Chen et al. (2020) give an extension in which the DGP includes multiple assets in which the dynamics are driven by common factors and the asymptotic performance of the PSY procedure in detecting a common factor bubble is analyzed.

5 The Multiplicity Issue

In all hypothesis testing the prospect of making a false decision is always present with non-zero probabilities of a false positive (a Type I error controlled by size) and a false negative (a Type II error of failing to reject the null, determined by test power). The basic principle underlying Neyman-Pearson testing is to cap size (i.e., control size to some preassigned level β, say) and use a test that maximizes power against the alternative of interest. The multiplicity issue is concerned with the failure to cap size at the required level when performing multiple hypothesis tests. This problem is universal when repeated testing occurs, as in recursive test procedures. In the content of bubble identification, it refers to the situation where a speculative bubble is falsely identified in repeated testing of the same null hypothesis. For each hypothesis test, suppose the probability of making a Type I error is β_T which is often set to be 5% and referred to as the 'nominal size'. Dependence of the size on the sample size T is often convenient for asymptotic theory to ensure test consistency, as seen above. The multiplicity issue arises when the test is implemented many times, as it is when the test is employed recursively over subsets of the sample observations. The presence of Type I error distortion among a large number of such tests is almost guaranteed unless corrective measures are taken.

Suppose a real-time analysis is conducted on a sample containing T observations with the PSY procedure. The PSY statistic is computed for each observation running from t_{min} to T and inferences of bubble existence are drawn for each of the $T - t_{min} + 1$ individual observations. The probability of making a false positive conclusion for each test may be set to 5%. But collectively the probability of making at least one false positive conclusion among the $T - t_{min} + 1$ tests (a.k.a, the family wise error rate, FWER) is close to 100% if the number of tests is large. Such 'false discovery' can grossly distort inference and potentially mislead investors and policymakers. It is therefore important to address the issue.

5.1 Multiple Testing Algorithm

The goal in dealing with multiplicity is to quantify the degree of Type I error accumulation that arises when many hypothesis tests are performed and with this information control size by suitable adjustment. Suppose the window width used in multiple testing for the presence of a bubble is T_w. The focus in controlling multiplicity is then on samples that are of the effective size T_w. We propose the following procedure for simulating finite sample critical values to employ in multiple testing. The approach is in the same spirit as bootstrapping methods that have been developed to address data mining and multiplicity in recursive testing (White, 2000; Shi et al., 2020; Shi and Phillips, 2020). Let $T^* = t_{min} + T_w - 1$.

- Step 1: generate a sample path with T^* observations from the model

$$x_t^{(m)} = x_{t-1} + \varepsilon_t.$$

- Step 2: compute the PSY statistic sequence $\left\{PSY_s^{(m)}(t_{\min})\right\}_{s=t_{\min}}^{T^*}$ from the simulated data series.

- Step 3: compute
$$MPSY^{(m)} = \max_{s\in[t_{\min},T^*]} \left\{ PSY_s^{(m)}(t_{\min}) \right\}.$$

- Step 4: repeat Step 1-3 M^* times.

- Step 5: take the 95^{th} percentile of $\left\{MPSY^{(m)}\right\}_{m=1}^{M^*}$ as the critical value of the PSY test, denoted by $cv^*(0.05)$.

There are two differences between the standard procedure for simulating finite sample critical values and the algorithm proposed above. First, in Step 1 instead of generating T observations as in the standard procedure T^* observations are generated so that the resulting PSY statistic sequence from the simulated data is of the correct dimension T_w. Second, in Step 5 the 95^{th} percentile is taken from the distribution of $MPSY^{(m)}$ instead of $PSY_s^{(m)}$ for each $s \in [t_{\min}, T]$ as in the standard procedure. This step ensures that the false rejection probability for the observation with the largest PSY statistic within the control window is 5%.

5.2 Simulations

We compare the family wise error rate of the PSY test with the standard and the multiple testing procedures, labelled respectively as 'PSY (standard)' and 'PSY (multiple)', for computing critical values. The minimum window size of the PSY procedure is set according the rule of Phillips et al. (2015b) and the lag order K of the ADF model is set to one. The data are generated under the null hypothesis (23) over a 10-year period with sampling frequencies of quarterly, monthly, and weekly (i.e., $T = \{40, 120, 520\}$). For each parameter constellation, the simulation is repeated $M = 2,000$ times.

Let $I_s^{(m)}$ be the bubble indicator for observation s from the m^{th} simulated data series, taking value one when $PSY_s^{(m)}$ is above the critical value and zero otherwise. The family wise error rate is defined as
$$FWER = \frac{1}{M}\sum_{m=1}^{M} \mathbf{1}\left(\sum_{s=t_{\min}}^{T} I_s^{(m)} > 0\right),$$
where $\mathbf{1}(.)$ is an indication function taking value one when the condition is satisfied and zero otherwise.

Table 1 shows the family wise error rates of the 'PSY (standard)' and 'PSY (multiple)' tests. For the PSY (multiple), the FWER is controlled within the whole sample period ($T_w = T$). Evidently, with the use of standard critical values, the FWER is lowest when $T = 40$ but still badly distorted at 55%, and then rises rapidly to 93% when T increases to 520. In sharp contrast, the FWER of the multiple testing procedure is stable around 5% under all settings.

Table 1: The family wise error rates of the PSY procedure

FWER / Sample Size	Quarterly	Monthly	Weekly
	$T = 40$	$T = 120$	$T = 520$
PSY (standard)	0.55	0.78	0.93
PSY (multiple)	0.04	0.06	0.05

It is important to recognize that there is a trade-off between the use of FWER and Type II errors. Fewer Type I errors typically result in more Type II errors, meaning that there is greater chance of failure to identify small bubbles or to detect a bubble at its later stage of development. To examine the performance of the PSY procedure under the alternative, we consider the bubble generating process (25). Parameters are set as in Figure 1. It is standard in the literature to report the power of a test, which is the probability of rejecting the null hypothesis when it is false. This is equivalent to the FWER in multiple testing but the computation is made under the alternative. However, the power statistic does not reveal the estimation accuracy of bubble locations and does not preclude Type I errors. An important example of the latter is the inconsistency that arises when the test falsely rejects the null for observation $t \in [1, t_e - 1]$ but fails to detect bubbles in $[t_e, t_c - 1]$. Although this is an undesirable scenario, it would still be counted towards test power. We therefore use two alternative measures to measure success under the alternative: (i) the successful detection rate (SDR); and (ii) the averaged delays in the estimated bubble origination and collapse dates.

The SDR measures the probability of the estimated bubble origination date falling between the true start and collapse dates, given by

$$SDR = \frac{1}{M} \sum_{m=1}^{M} \mathbf{1} \left(t_e \leq \hat{t}_e^{(m)} \leq t_c \right).$$

Averaged delay in dating the bubble is computed by the following two measures, one for origination and one for collapse

$$\frac{1}{M} \sum_{m=1}^{M} (\hat{t}_e^{(m)} - t_e) \text{ and } \frac{1}{M} \sum_{m=1}^{M} (\hat{t}_c^{(m)} - t_c).$$

Evidently, the control of Type I errors comes at a cost. From Table 2, the succesful detection rate of PSY (multiple) is lower (75% versus 84%) and the bias in the estimated bubble origination date is larger (12.20 months versus 7.56 months), although there is a slight improvement in the identification accuracy of the bubble termination date. In practice, one could take an intermediate position between the two procedures, depending on the purpose and subject to individual preferences. That is, instead of controlling the Type I error over the entire 10-year's period, it may be more realistic and useful to manage

Table 2: The successful detection rates and the biases of the estimated bubble origination and collasping dates. Numbers in the parenthes are standard deviations.

$T = 120$	SDR	$\frac{1}{M}\sum_{m=1}^{M}(\hat{t}_e^{(m)} - t_e)$	$\frac{1}{M}\sum_{m=1}^{M}(\hat{t}_c^{(m)} - t_c)$
PSY (standard)	0.84	7.56 (4.99)	0.79 (1.72)
PSY (multiple)	0.75	12.20 (5.33)	0.77 (1.71)

it over a smaller window, e.g., 1 year, 2 years, or 5 years. As such, one would expect a smaller delay should there be a bubble within the window. This approach to testing and estimation is analogous to using a loss function that takes size and power into account, rather than strictly following the principle of capping size.

The performance of the multiple testing procedure under alternative settings of T_w is investigated in Table 3. The data generating processes is the same as before. In addition to FWER, we report the average numbers of rejections (ANR) under the null, i.e.,

$$ANR = \frac{1}{M}\sum_{m=1}^{M}\sum_{s=t_{\min}}^{T} I_s^{(m)},$$

which complements the FWER rate as a measure for Type I errors. The performance of the test under the alternative is again evaluated by SDR and by the average delays in identifying bubble start and collapse dates.

Although the FWER remains sizeable (44%) when the control window is one year, by comparison with the 78% of the standard procedure (Table 1), it is already a dramatic improvement. The average number of false rejections within each sample path is 1.67, with a standard deviation of 3.26. On the other hand, we observe a slight deterioration in the bubble identification accuracy. The average delay in identifying the bubble origination date is slightly longer (8.76 observations), while the SDR is the same as that of the standard procedure (84%). As the control window T_w increases from 1 year to 5 years, the family-wise error rate declines from 44% to 10%, and the average number of false rejections decreases from 1.67 to 0.22. Meanwhile, the successful detection rate reduces to 77%, and the delay in identifying bubble origination increases to 11.53 months. We do not observe obvious changes in the estimation accuracy of the bubble termination date.

In practice, the choice of the control window T_w is subjective. The selection of a larger control window results in more conservative testing with a lower possibility of making false-positive conclusions. In contrast, a small control window enables a higher probability of detecting the presence of a bubble as well as earlier detection.

6 Empirical Application: S&P 500 Stock Market

As an illustration, we conduct a pseudo real-time analysis monitoring for a bubble in the S&P 500 stock market at the monthly frequency, starting from January 1990 and

Table 3: The finite sample performance of the PSY test with the multiple testing procedure for critical values. The control window T_w is set to be 1 year, 2 years, and 5 years. The sample size $T = 120$. Numbers in the parentheses are standard deviations.

$T = 120$	Null		Alternative		
	FWER	# of Rejections	SDR	$\frac{1}{M}\sum_{m=1}^{M}(\hat{t}_e^{(m)} - t_e)$	$\frac{1}{M}\sum_{m=1}^{M}(\hat{t}_c^{(m)} - t_c)$
T_w: 1 years	0.44	1.67 (3.26)	0.84	8.76 (5.21)	0.85 (1.32)
T_w: 2 years	0.22	0.57 (1.53)	0.80	10.44 (5.37)	0.78 (1.71)
T_w: 5 years	0.10	0.22 (0.84)	0.77	11.53 (5.33)	0.84 (1.36)

terminating in June 2020. The sample period covers the dot-com expansion in the late 1990s and the 2007-2009 subprime mortgage crisis period. Figure 3 displays the logarithmic price-dividend ratios of the S&P 500 composite index. The price and dividend series are obtained from Robert Shiller's website.[4] The data series experienced a decade-long expansion at the beginning of the sample period and peaked in March 2000. The expansion is followed by a market downturn that leads to a reversion to the market's 1997 level in 2003. The data exhibit a V-shape dynamic around the subprime mortgage crisis period.

Figure 3: The log price-dividend ratios of the S&P 500 composite index and the estimated residual component from January 1990 to June 2020.

We apply the two bubble identification strategies detailed in Section 3 to the market. The first strategy is to apply the PSY explosive root test to the log price-dividend ratios. The second is to estimate the fundamental component of the log price-dividend ratios and apply the PSY test to the non-fundamental/residual component. The decomposition is

[4]See `http://www.econ.yale.edu/~shiller/data.htm`.

based on equations (17)-(18), where y_t is the first difference of $p_t - d_t$ and X_t includes real dividends and 3-month treasury bill rates. The real dividend series is from Robert Shiller's website and 3-month treasury bill rates are from the FRED database.[5] The regression model is estimated using IVX detailed in Appendix A. The estimated fundamental component is plotted along with the log price-dividend ratios in Figure 3. We observe a general upward trend in the data series. The residual component is computed as the difference between the log price-dividend ratios and the estimated fundamentals and captures fluctuations around the trend.

For the PSY procedure, the minimum window contains 38 observations, which is set according to the rule of $t_{\min} = 0.01T + 1.8\sqrt{T}$. Monitoring starts from February 1993 onwards. Critical values are obtained from the standard procedure in the first row of Figure 4 and the multiple testing procedure in the second and third rows, with the control window being two years and five years respectively (i.e., $T_w = \{24, 60\}$). The black lines are the x_t series, and the green shaded areas are periods when the PSY statistic exceeds its critical value. The bubble signal is switched on for two sample periods: the late 1990s and 2008-2009. While the first episode is associated with upward expansion of the market, the second episode overlaps with a market plunge. When using the PSY (standard) test on the residuals, this test also identifies a market downturn between 2002-2003. Although the PSY test was initially designed for identifying explosive bubbles, Phillips and Shi (2019) show that this procedure can also detect crises. For analysis and further detail on crisis identification readers may refer to Phillips and Shi (2019).

Some further comments are worth mentioning regarding the dot-com bubble episode. First, the identified origination (respectively, termination) date is earlier (later) when the multiplicity issue is not accounted for or when the control window T_w is smaller. This result is consistent with the simulation findings in Section 5, where both the family-wise error rate and the successful detection rate are higher with PSY (standard) or when PSY (multiple) is used with a smaller control window. Second, the identified bubble period is shorter when we control for the impact of market fundamentals using the approach of Shi and Phillips (2020). It suggests that the expansion was first driven by fundamentals followed by speculative sentiment and investment in 1997-1998, in a similar way to the behavioral model of Phillips (2016). Third, the multiple testing procedure indicates that the bubble signal switched off almost two years before the market reached its peak. This finding is coherent with results in Laurent and Shi (2020) and has potent implications for investors who aim to withdraw investments in anticipation of a market collapse.

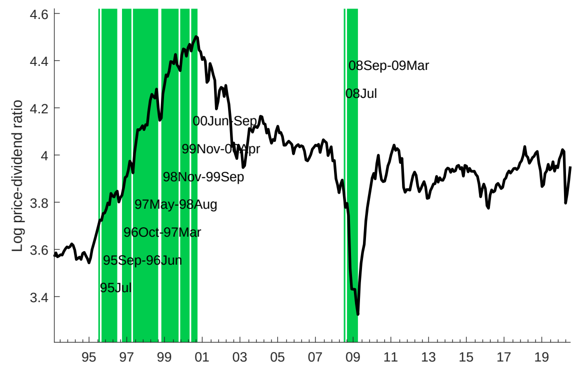
7 Conclusion

In the presence of a financial bubble asset prices may be written as the sum of two compnents, one determined by economic fundamentals and the other following an explosive

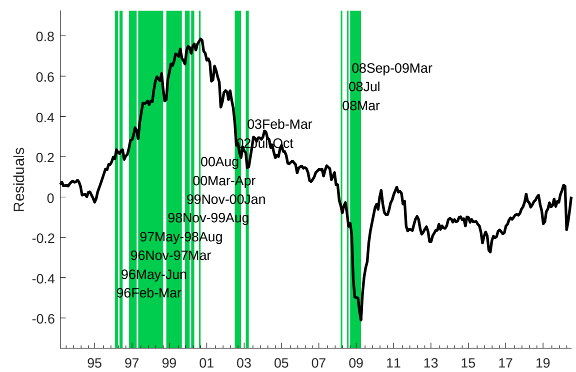
[5] see `https://fred.stlouisfed.org/`.

Figure 4: The PSY explosive root tests. The black lines are the log price-dividend ratios in the first column and the estimated non-fundamental components in the second column. The green shaded areas are periods when the PSY statistic is larger than its critical value.

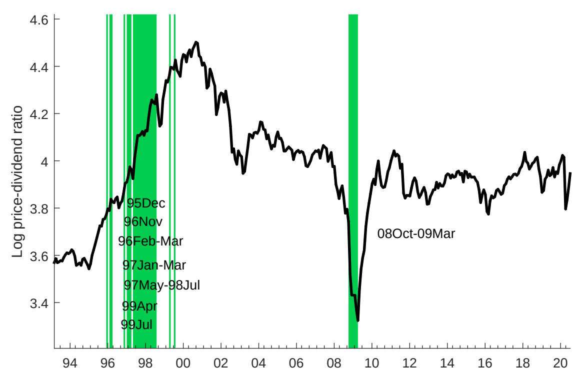
(a) Log price-dividend ratio: PSY (standard)



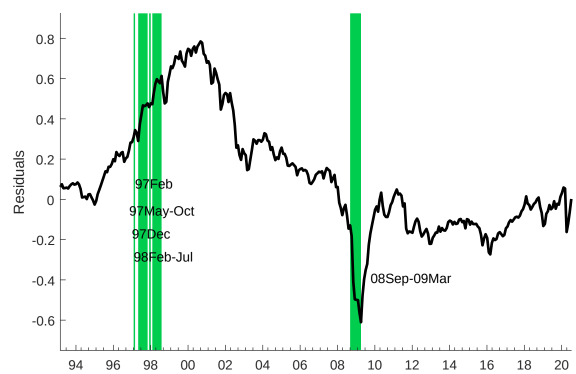
(b) Residual: PSY (standard)



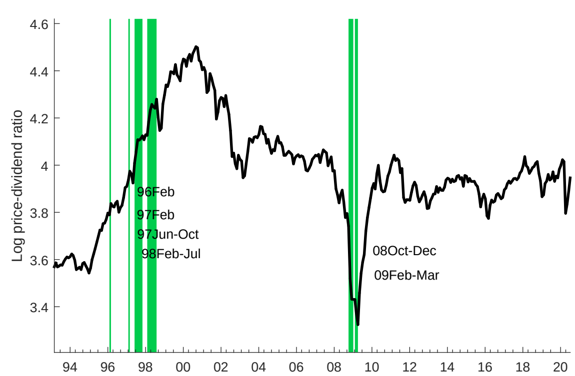
(c) Log price-dividend ratio: PSY (multiple) $T_w = 24$



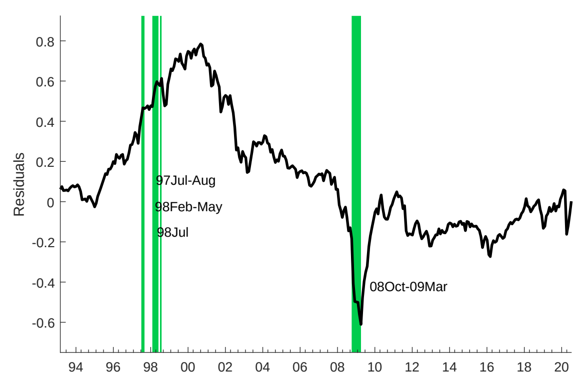
(d) Residual: PSY (multiple) $T_w = 24$



(e) Log price-dividend ratio: PSY (multiple) $T_w = 60$



(f) Residual: PSY (multiple) $T_w = 60$



dynamic that constitutes a bubble. The common strategy for bubble identification in empirical work is to apply explosive autoregressive root tests to proxies for the unobserved bubble. In practice proxies are typically constructed by using asset prices themselves or by standardizing prices in some way (e.g., by taking ratios) using observed fundamentals that are considered relevant to the asset. This paper outlines an alternative approach that provides a theoretical framework for the construction of a bubble series that is simple to use and general enough to apply in the case of many different assets.

The two most commonly used proxies for bubbles are real prices and log price-payoff ratios. During a speculative bubble, both asset prices and log price-payoff ratios are explosive because the dynamics of the bubble component dominate the observed data. Market fundamentals are typically near integrated time series, so these proxies of the bubble component are therefore contaminated by highly persistent series, a feature that inevitably influences the sensitivity and performance of explosive root tests.

A new strategy in the literature is to estimate the fundamental component itself and remove it from the proxy variable before applying explosive root tests. The additional step in this strategy helps to reduce the persistent noise element in the proxy and enable more accurate bubble identification. This approach has led to new bubble test results in several housing markets but has yet to be applied to equities. The present chapter overviews the fundamental-bubble decomposition approach proposed by Shi and Phillips (2020) and provides an empirical application of the bubble identification strategy to the S&P 500 stock market.

Explosive autoregressive root tests play a crucial role in bubble identification. This chapter reviews the popular PSY real-time explosive root test strategy. The PSY approach aims to detect explosive dynamics, accounting for their nonlinearity and using this characteristic to provide a consistent real-time dating strategy. The procedure employs a recursive evolving algorithm that has been shown to outperform other methods. Nevertheless, recursive testing of the type employed by PSY inevitably faces multiplicity problems with size control. These issues are discussed and illustrated by simulations.

To address multiplicity the chapter provides a multiple-testing algorithm to generate critical values for the PSY test that are designed to control family-wise size. Simulation results show that the standard PSY procedure can have extremely high family-wise error rates, whereas the new procedure PSY (multiple) has good size control and satisfactory performance overall in detection rates under bubble alternatives. An empirical application to the S&P500 stock market reveals the importance of controlling for market fundamentals and addressing the multiplicity issue in testing for bubble behavior in equity markets. In particular, controlling for the impact of market fundamentals leads to a shorter Dot-Com bubble period in the 1990s and suggests that the expansive phase was first driven by fundamentals that then led to speculative behavior. Attention to the consequences of multiple testing produced a further material change in existing findings on the Dot-Com bubble by switching off the bubble signal in the data sometime before the market peak.

References

Blanchard, O. J. and M. W. Watson (1982). Bubbles, rational expectations and financial markets. *NBER Working Paper No. 0945*.

Caballero, R. J. and A. Krishnamurthy (2006). Bubbles and capital flow volatility: Causes and risk management. *Journal of monetary Economics 53*(1), 35–53.

Campbell, J. Y. and R. J. Shiller (1988). Stock prices, earnings, and expected dividends. *The Journal of Finance 43*(3), 661–676.

Case, K. E., E. L. Glaeser, and J. A. Parker (2000). Real estate and the macroeconomy. *Brookings Papers on Economic Activity 2000*(2), 119–162.

Chen, Y., P. C. B. Phillips, and S. Shi (2020). Common bubble detection in large dimensional financial systems. *Working paper*.

Diba, B. T. and H. I. Grossman (1988). Explosive rational bubbles in stock prices? *The American Economic Review 78*(3), 520–530.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review 81*(4), 922–930.

Harvey, D. I., S. J. Leybourne, and R. Sollis (2013). Recursive right-tailed unit root tests for an explosive asset price bubble. *Journal of Financial Econometrics 13*(1), 166–187.

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

Harvey, D. I., S. J. Leybourne, and Y. Zu (2020). Sign-based unit root tests for explosive financial bubbles in the presence of deterministically time-varying volatility. *Econometric Theory 36*(1), 122–169.

Hirano, T., M. Inaba, and N. Yanagawa (2015). Asset bubbles and bailouts. *Journal of Monetary Economics 76*, S71–S89.

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

Hu, Y. and L. Oxley (2017). Exuberance, bubbles or froth? some historical results using long run house price data for Amsterdam, Norway and Paris. Technical report, Working Paper, University of Waikato.

Kostakis, A., T. Magdalinos, and M. P. Stamatogiannis (2015). Robust econometric inference for stock return predictability. *The Review of Financial Studies 28*(5), 1506–1553.

Laurent, S. and S. Shi (2020). Volatility estimation and jump detection for drift–diffusion processes. *Journal of Econometrics 217*(2), 259–290.

Lee, J. H. and P. C. B. Phillips (2016). Asset pricing with financial bubble risk. *Journal of Empirical Finance 38*, 590–622.

Mishkin, F. S. (1995). Nonstationarity of regressors and tests on real-interest-rate behavior. *Journal of Business & Economic Statistics 13*(1), 47–51.

Pavlidis, E., A. Yusupova, I. Paya, D. Peel, E. Martínez-García, A. Mack, and V. Grossman (2016). Episodes of exuberance in housing markets: in search of the smoking gun. *The Journal of Real Estate Finance and Economics 53*(4), 419–449.

Pavlidis, E. G., I. Paya, and D. A. Peel (2017). Testing for speculative bubbles using spot and forward prices. *International Economic Review 58*(4), 1191–1226.

Phillips, P. C. B. (2005). Econometric analysis of fisher's equation. *American Journal of Economics and Sociology 64*(1), 125–168.

Phillips, P. C. B. (2016). Modeling speculative bubbles with diverse investor expectations. *Research in Economics 70*(3), 375–387.

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

Phillips, P. C. B. and T. Magdalinos (2009). Econometric inference in the vicinity of unity. *Singapore Management University, CoFie Working Paper 7.*

Phillips, P. C. B. and S. Shi (2019). Detecting financial collapse and ballooning sovereign risk. *Oxford Bulletin of Economics and Statistics 81*(6), 1336–1361.

Phillips, P. C. B. and S. Shi (2020). Real time monitoring of asset markets: Bubbles and crises. In *Handbook of Statistics*, Volume 42, pp. 61–80. Elsevier.

Phillips, P. C. B., S. Shi, and J. Yu (2014). Specification sensitivity in right-tailed unit root testing for explosive behaviour. *Oxford Bulletin of Economics and Statistics 76*(3), 315–333.

Phillips, P. C. B., S. Shi, and J. Yu (2015a). Supplement to two papers on multiple bubbles. *International Economic Review 56*(4).

Phillips, P. C. B., S. Shi, and J. Yu (2015b). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review 56*(4), 1043–1078.

Phillips, P. C. B., S. Shi, and J. Yu (2015c). Testing for multiple bubbles: Limit theory of real-time detectors. *International Economic Review 56*(4), 1079–1134.

Phillips, P. C. B. and S. P. Shi (2018). Financial bubble implosion and reverse regression. *Econometric Theory 34*(4), 705–753.

Phillips, P. C. B. and V. Solo (1992). Asymptotics for linear processes. *The Annals of Statistics 20*(2), 971–1001.

Phillips, P. C. B., Y. Wu, and J. Yu (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review 52*(1), 201–226.

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

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

Rapach, D. E. and C. E. Weber (2004). Are real interest rates really nonstationary? new evidence from tests with good size and power. *Journal of Macroeconomics 26*(3), 409–430.

Rose, A. K. (1988). Is the real interest rate stable? *The Journal of Finance 43*(5), 1095–1112.

Shi, S. (2017). Speculative bubbles or market fundamentals? an investigation of US regional housing markets. *Economic Modelling 66*, 101–111.

Shi, S., S. Hurn, and P. C. B. Phillips (2020). Causal change detection in possibly integrated systems: revisiting the money–income relationship. *Journal of Financial Econometrics 18*(1), 158–180.

Shi, S. and P. C. B. Phillips (2020). Diagnosing housing fever with an econometric thermometer. Technical report, CAMA Working Paper.

Shi, S., A. Rahman, and B. Z. Wang (2020). Australian housing market booms: Fundamentals or speculation? *Economic Record, available at https://doi.org/10.1111/1475-4932.12553*.

Sun, Y. and P. C. B. Phillips (2004). Understanding the Fisher equation. *Journal of Applied Econometrics 19*(7), 869–886.

Wang, X. and J. Yu (2019). Bubble testing under deterministic trends. Technical report, Singapore Management University Economics & Statistics Working Papers.

White, H. (2000). A reality check for data snooping. *Econometrica 68*(5), 1097–1126.

Wöckl, I. (2019). Bubble detection in financial markets-a survey of theoretical bubble models and empirical bubble detection tests. *Available at SSRN 3460430*.

Yang, B., W. Long, L. Peng, and Z. Cai (2020). Testing the predictability of US housing price index returns based on an IVX-AR model. *Journal of the American Statistical Association 115*, 1598–1619.

Appendix A: IVX Estimation

The IVX method was proposed by Phillips and Magdalinos (2009), extended to allow for heteroskedastic errors in Kostakis et al. (2015) and subsequently with Cochrane-Orcutt type parametric adjustments for serial correlation in Yang et al. (2020). In what follows we outline the IVX-AR method of Yang et al. (2020).

Define the transformed dependent and explanatory variables as $y_{\phi,t} = y_t - \sum_{j=1}^{q} \phi_j y_{t-j}$ and $X_{\phi,t} = X_t - \sum_{j=1}^{q} \phi_j X_{t-j}$, where $\phi = (\phi_1, \ldots, \phi_q)$ and $\{\phi_i\}$ are the coefficients in (20). Rewrite the regression equation (17) as

$$y_{\phi,t} = \alpha_\phi + \beta X_{\phi,t-1} + \eta_t$$

where $\alpha_\phi = \alpha \left(1 - \sum_{j=1}^{q} \phi_j\right)$. The IVX estimation method uses instrumental variables $z_{\phi,t}$ constructed from the predictors $X_{\phi,t}$ as follows:

$$z_{\phi,t} = \sum_{j=1}^{t} \rho_z^{t-j} \Delta X_{\phi,j} \text{ with } \rho_z = I_k + \frac{C_z}{T^\beta} \text{ for some } C_z \text{ and } \beta \in (0,1),$$

where $t = 2, \ldots, T$, $z_{\phi,1} = 0$, k is the dimension of X_t, and I_k is the $k \times k$ identity matrix. When the X_t are unit root or near unit root processes, the instrumental variables $z_{\phi,t}$ are less persistent than X_t because the autoregressive coefficient matrix ρ_z generates mildly integrated time series from $\Delta X_{\phi,t}$. The IVX parameters are set to $C_z = -I_k$ and $\beta = 0.95$, folliwing the recommendation of Kostakis et al. (2015).

Let $\underline{Y}_\phi = \left(\underline{y}_{\phi,1}, \ldots, \underline{y}_{\phi,T}\right)'_{T\times 1}$, $\underline{X}_\phi = \left(\underline{x}_{\phi,1}, \ldots, \underline{x}_{\phi,T}\right)'_{T\times k}$, and $Z_\phi = \left(z_{\phi,1}, \ldots, z_{\phi,T}\right)'_{T\times k}$, where $\underline{y}_{\phi,t} = y_{\phi,t} - \frac{1}{T}\sum_{t=1}^{T} y_{\phi,t}$ and $\underline{X}_{\phi,t} = X_{\phi,t} - \frac{1}{T}\sum_{t=1}^{T} X_{\phi,t}$ denote the demeaned series of $y_{\phi,t}$ and $X_{\phi,t}$. The IVX-AR estimation method is implemented as follows.

Step 1: Compute the IVX estimator with a given $\phi \in \left[\underline{\phi}, \overline{\phi}\right]$,

$$\hat{\beta}^{(\phi)} = \underline{Y}'_\phi Z_\phi \left(\underline{X}'_\phi Z_\phi\right)^{-1}$$

and obtain the residuals $\hat{\eta}_{\phi,t} = \underline{y}_{\phi,t} - \hat{\beta}^{(\phi)} \underline{X}_{\phi,t}$.

Step 2: Find the optimal moving average coefficients ϕ^* by minimizing the sum of squared residuals

$$\phi^* = \arg\min_{\phi} \left(\sum_{t=1}^{T} \hat{\eta}_{\phi,t}^2 \right).$$

Step 3: Compute the IVX-AR estimator $\hat{\beta}_{IVX}$ with ϕ^* as

$$\hat{\beta}_{IVX} = \underline{Y}'_{\phi^*} Z_{\phi^*} \left(\underline{X}'_{\phi^*} Z_{\phi^*} \right)^{-1},$$

and the corresponding fitted intercept $\hat{\alpha}_{IVX} = \bar{y} - \hat{\beta}_{IVX}\bar{X}$ with $\bar{y}$ and $\bar{X}$ being the sample means of y_t and X_t.

The asymptotic and finite sample properties of the IVX-AR method are explored in Yang et al. (2020). This method has all the advantages of the usual IVX method, including endogenous instrumentation and valid asymptotic inference, for use in the reduced form regression (17) for the determination of fundamentals via predictive regression. These advantages are coupled with anticipated improvements in the method's finite sample properties over standard IVX arising from the AR transform of the system in the presence of serially correlated errors, which help to capture in the equation's fitted values the essential characteristics of innovations in fundamentally driven prices. Construction of the fundamentals time series then proceeds as in (22) with the residuals capturing the bubble component. Simulations show that this method works well in practice for testing the presence of bubbles in asset prices.