DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

Comparing Tests for Identification of Bubbles

Kristoffer Pons Bertelsen

CREATES Research Paper 2019-16

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Comparing Tests for Identification of Bubbles*

Kristoffer Pons Bertelsen†

Abstract

This paper compares the log periodic power law (LPPL) and the supremum augmented Dickey Fuller (supremum ADF) procedures considering bubble detection and time stamping capabilities in a thorough analysis based on simulated data. A generalized formulation of the LPPL procedure is derived and analysed demonstrating performance improvements.

Keywords: Rational bubbles, explosive processes, log periodic power law, critical points theory

JEL: C01, C02, C12, C13, C22, C52, C53, C58, C61, G01

This version: October 9, 2019

*The author thanks Tom Engsted, Esben Høg, Thomas Quistgaard Pedersen, Erik Christian Montes Schütte, and participants at Aarhus University for useful comments and discussions. This work was supported by CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78) funded by the Danish National Research Foundation; and the Center for Scientific Computing, Aarhus (CSCAA).

†CREATES, Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, DK-8210 Aarhus V, Denmark. E-mail: kristoffer@econ.au.dk.

1 Introduction

This paper links two branches of the literature concerning the identification of bubbles in financial markets. On one side there is the mainstream economic literature on asset bubbles currently championed by the Supremum Augmented Dickey-Fuller tests (henceforth referred to as supremum ADF tests) (Phillips et al., 2011, 2015), and on the other side is the completely separated albeit very active literature on asset bubbles that uses the fitting of a Log Periodic Power Law (henceforth referred to as LPPL) as the means of testing for the presence of a bubble (Sornette et al., 1996; Sornette and Johansen, 1998; Johansen et al., 1999, 2000; Sornette and Johansen, 2001). See Brauers et al. (2014) for a recent application. Previous studies have considered the relative performance of various tests for bubbles (Homm and Breitung, 2012), but this is the first paper to introduce the LPPL model to the mainstream economic literature in a similar rigorous setup. We choose the supremum ADF test as the base model from the mainstream economic literature, as it is widely regarded as a state of the art test for bubbles and used by practitioners. Various other tests could also be considered, e.g. the variance bounds tests of Shiller (1981); Leroy and Porter (1981), the two-step test by West (1987), the intrinsic bubbles approaches of Froot and Obstfeld (1991), and with the fractionally integrated approach of Cuñado et al. (2005); Frömmel and Kruse (2012). An overview of the various tests can be found in Gürkaynak (2008).

The mainstream economic literature identifies a bubble when the price process deviates from the fundamentals and was originally introduced by Diba and Grossman (1988a,b) with an adaptation of the Dickey-Fuller test. Under the null hypothesis there is no bubble and the price process contains at most a unit root driven by the dividend process. This is tested against the alternative that the price process has an explosive root that is then assumed to stem from the bubble component. Evans (1991) showed that even though the bubble, and by extension the price, is explosive during bubble expansion, the test would not identify the bubble correctly if it collapses par-

tially one or more times during the test window. The test was improved with the recursive test SADF (Phillips et al., 2011) and a generalized version, GSADF, was introduced in Phillips et al. (2015) with more power in samples with multiple explosive periods. The introduction of the Backward Supremum Augmented Dickey-Fuller test (BSADF) also allowed for time stamping of the bubble periods (Phillips et al., 2015).

One of the drawbacks of the GSADF test is that it assumes a common unit root (cointegration) between the dividend and the price processes which has not been tested. Engsted (2006) and Engsted and Nielsen (2012) developed a framework using a bivariate coexplosive vector autoregression, but while this deals with the non-tested assumption that the price and dividend processes should be cointegrated, it struggles to detect bubbles when allowing for partial collapses, and we will not be using that particular framework for this study. The GSADF test was made robust to autocorrelated innovations by Pedersen and Schütte (2017) by calculating the critical values for the test statistic using the sieve bootstrap (Chang and Park, 2003; Palm et al., 2008). We include this newly proposed test in the study to see if there are significant improvements to the performance over the BSADF test as well as how it compares to the LPPL procedure.

The second branch of the bubble literature is based on the LPPL procedure originating in Sornette et al. (1996), Sornette and Johansen (1998), and Johansen et al. (1999). Under a certain set of assumptions regarding trader interaction the price process can be shown to follow an LPPL in the time preceding the collapse of the bubble. The literature draws on the critical points theory from physics that is also used in the modelling of geophysical events like earthquakes and volcanic eruptions (Sornette and Johansen, 1998). The critical point is the time at which the traders' opinions become highly correlated and sensitive to external signals, and at which point the bubble is most likely to collapse. The LPPL model differs from the conventional economics literature by not including fundamentals. As previously mentioned a bubble is defined as an inflation of the price above its fundamental price, and this

definition can be difficult to implement if one does not have a proxy for the fundamental price or if the proxy is only available at a low frequency. The LPPL model only requires knowledge of the price trajectory to determine whether there is a bubble or not, which means it can be used with e.g. daily observations or index data. For a review of the LPPL literature see for example Geraskin and Fantazzini (2013) and Sornette et al. (2013).

There are also similarities between the two branches of the bubble literature. Abreu and Brunnermeier (2003) discuss how a bubble can exist despite having rational investors in the economy. They argue that the bubble is allowed to persist because the rational traders are unable to coordinate their sell off. Similarly Johansen et al. (2000) explain the collapse of the bubble as the point in time when the traders are able to synchronize their behaviour.

Depending on the nature and dynamics of the bubble it might be reasonable to expect it to lead to a different price process. Hence, we simulate different bubble dynamics to determine the robustness of the three tests. The simulated processes follow mainly those used when presenting the GSADF test in Phillips et al. (2015) with some additions and modifications. This allows us to investigate how often the frameworks erroneously or correctly detect a bubble. These properties share some of the intuition behind the concepts of size and power from statistics.

Up until now the LPPL literature has assumed a zero required rate of return when deriving the LPPL function. This paper extends the model by allowing for a non-zero and possibly time varying required rate of return. The case of a positive constant required rate of return is included in the Monte Carlo study. We also develop and present proofs of the LPPL processes, which can be found in the appendix.

We find that the LPPL function is very flexible and can fit a large variety of processes, despite the imposed parameter restrictions. The root mean squared error requirement turns out to be the driving factor when identifying bubbles. Unfortunately there is no single maximum value of the RMSE that allows the LPPL procedure to

detect existing bubbles and reject non-existing bubbles. The ability of the LPPL procedure to estimate the most likely time of the collapse of the bubble is doubtful, albeit improved by our generalization. We confirm that the supremum ADF tests are oversized when innovations are serially correlated and that the robust supremum ADF test developed by Pedersen and Schütte (2017) is able to reduce the size. Both versions of the supremum ADF test however, are oversized when the innovations follow a GARCH process. There are mixed results indicating that the tests are able to detect a bubble very early in the build up of the bubble.

The paper is organised as follows. Section 2 presents the supremum ADF test including the robust version developed in Pedersen and Schütte (2017). Section 3 presents the LPPL procedure. Section 4 compares the two branches of the bubble literature presenting differences and similarities. Section 5 describes the processes that are used for the Monte Carlo study. Section 6 presents the results from the Monte Carlo study. Section 7 provides a discussion of the robustness of the findings of this paper. And section 8 concludes the paper.

2 The Supremum ADF Tests

The supremum ADF tests presented in Phillips et al. (2011) and Phillips et al. (2015) are widely used in the economics literature to detect bubbles in financial assets. They are based on the idea of testing whether the time series has an explosive root as an alternative to at most a unit root using a Dickey-Fuller framework. The fundamental price of an asset is the sum of the discounted flow of future dividends and perhaps some unobserved fundamentals.

$$P_t^F = \sum_{i=0}^{\infty} \left(\frac{1}{1+R} \right)^i \mathbb{E}_t \left(D_{t+i} + U_{t+i} \right),$$

where R is a constant real discount rate as created by market participants, $\mathbb{E}_t$ is the expectation operator conditional on the information available at time t, D_t is the

dividend received at time t, and U_t contains unobserved fundamentals as originally presented in Diba and Grossman (1988a). The bubble is then modelled as a sub-martingale B_t such that

$$\mathbb{E}_t\left(B_{t+1}\right) = (1+R)\, B_t. \tag{1}$$

Then the resulting observable price process is given by

$$P_t = P_t^F + B_t. \tag{2}$$

Usually the dividend stream is modelled as an $I(1)$ process and the unobserved fundamentals are assumed to be at most $I(1)$. Then since P_t^F, and by extension P_t is a discounted sum of unit root processes they will themselves contain a unit root. If the observed price only contains one unit root, and it is exactly the unit root from the dividends, then this can be removed through cointegration by taking the price-dividend ratio $\frac{P_t}{D_t}$. However, (1) reveals that the bubble component contains an explosive root that will also be carried over to the observed price P_t.

The null hypothesis of the supremum ADF tests is that there is no bubble and that the price-dividend ratio, y_t, follows an AR(1) process with a weak drift.

$$y_t = dT^{-\eta} + (1-\beta)\, y_{t-1} + \epsilon_t, \qquad \epsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2\right),$$

where d is a constant, T is the sample size, and $\eta > \frac{1}{2}$ is a parameter controlling the magnitude of the drift. This functional form was presented in Phillips et al. (2014) to give a realistic model with tractable asymptotic properties. In this paper, as in Phillips et al. (2015), we will consider the case where $d = \eta = 1$ and $\beta = 0$. With this specification it is clear that under the null, the process will be $I(1)$ and conveniently fits the framework of Fuller (1976); Dickey and Fuller (1979).

The tests calculate a Dickey-Fuller test statistic from various sample windows for

increased power. The test windows are characterized by the fractions r_1 and r_2 of the total sample length, T. Let r_w be the length of the test window such that $r_2 = r_1 + r_w$. Then the test window will extend from observation $\lfloor Tr_1 \rfloor$ to observation $\lfloor Tr_2 \rfloor$ and has a length of $\lfloor Tr_w \rfloor$, where $\lfloor \bullet \rfloor$ is the floor function. Finally denote the minimum relative size of the test window by r_0 such that $r_w \geq r_0$ and $\lfloor Tr_w \rfloor \geq \lfloor Tr_0 \rfloor$. Then the regression model identified by r_1 and r_2 can be written as

$$\Delta y_t = \hat{\alpha}_{r_1,r_2} + \hat{\beta}_{r_1,r_2} y_{t-1} + \sum_{i=1}^{k} \hat{\phi}_{i,r_1,r_2} \Delta y_{t-i} + \epsilon_t, \qquad \epsilon_t \overset{iid}{\sim} \mathcal{N}\left(0, \sigma^2_{r_1,r_2}\right). \tag{3}$$

From this regression one can then test whether the process contains a unit root ($\beta_{r_1,r_2} = 0$) or an explosive root ($\beta_{r_1,r_2} > 0$). Bear in mind however, that it has not been shown that there actually is a common unit root (cointegration) between the price and the dividend processes. If the two processes do not share a unit root then even the price-dividend ratio might contain both a unit root and an explosive root from the bubble component, thus rendering the Dickey-Fuller framework invalid. There is also the possibility that the unobserved fundamentals contain a unit root, however in Phillips et al. (2015) the characteristics of the unobserved fundamentals are merely assumed away. If these assumptions would prove not to hold, the Dickey-Fuller framework would be equally invalid. The estimates obtained in (3) depend on the choice of the lag length k, which is selected by BIC over the set $k \in \left[1, \ldots, \lfloor 8\,(T/100)^{1/4} \rfloor\right]$.

The obtained test statistic, $\hat{\beta}^{r_2}_{r_1}$, can then be used to calculate the SADF, GSADF, and BSADF tests statistics. The SADF test is the simplest version and a special case of the GSADF test. It is defined as

$$SADF\,(r_0) = \sup_{r_2 \in [r_0, 1]} \hat{\beta}^{r_2}_0.$$

Note that the starting point of the subsections is fixed at $r_1 = 0$, and the supremum is then taken on the set of test statistics obtained by varying $r_2 \in [r_0, 1]$. This is

rather inflexible which is evident when allowing for partially collapsing bubbles like those presented in Evans (1991). Since the SADF only expands the subsection by increasing r_2 there is a risk that the test mistakes the partially collapsing bubble for a stationary process (Phillips et al., 2015).

The GSADF is then a generalization of the SADF test allowing for more flexibility in the choice of subsections such that the test is also capable of detecting partially collapsing bubbles. The test statistic is defined as

$$GSADF(r_0) = \sup_{r_2 \in [r_0,1] \wedge r_1 \in [0, r_2 - r_0]} \hat{\beta}_{r_1}^{r_2}.$$

From the definition one can see that not only does the subsection vary with r_2 but now also with r_1. This added flexibility ensures that the test is performed on the particular subsection containing only the build up period of the bubble, thus giving the correct significant realisations of the ADF test statistic.

As with the ADF test, the distribution of the supremum ADF test statistics is non-standard and the critical values are found by simulation using iid errors.

2.1 Time Stamping

The concept of varying the subsections of the sample to calculate the GSADF can be altered using the BSADF test to allow for date stamping (Phillips et al., 2015). The idea is that the test procedure moves backwards from $r_2 = 1$ to $r_2 = r_0$, while for each r_2 it calculates a supremum ADF test statistic by varying $r_1 \in [0, r_2 - r_0]$. The BSADF is like the GSADF a function of r_0 but it also depends on r_2. The BSADF is defined as

$$BSADF_{r_2}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \hat{\beta}_{r_1}^{r_2}.$$

Hence, the BSADF method considers a fixed end point, r_2, and works itself backwards from there. It performs multiple ADF tests obtaining the different $\beta_{r_1}^{r_2}$ test statistics for $r_1 \in [0, r_2 - r_0]$. Then the BSADF test statistic is obtained by taking the supremum

on the set of $\beta_{r_1}^{r_2}$ test statistics. Again the distribution of the test statistic is non-standard and the critical values are found by simulation using iid errors.

Then by considering the different BSADF test statistics derived from different values of r_2 one can determine the start and end points of the bubble. Varying $r_2 \in [r_0, 1]$ will give rise to a set of $BSADF_{r_2}(r_0)$ test statistics, and the start of the bubble is estimated as

$$\hat{r}_e = \inf_{r_2 \in [r_0,1]} \left\{ r_2 : BSADF_{r_2}(r_0) > cv_{r_2}^{\alpha} \right\}.$$

The observation marking the beginning of the bubble is then identified as $\lfloor T\hat{r}_e \rfloor$. Similarly the end date of the bubble is found by

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

where δ is a parameter ensuring that the identified bubble period has a certain minimum length. The observation marking the end of the bubble is then identified as $\left\lfloor T\hat{r}_f \right\rfloor$.

One should bear in mind that for the BSADF test, the significance level corresponds to each of the test statistics calculated across the range of r_2. Hence, simply by choosing a sufficiently fine partitioning of the data one can be almost certain to detect a bubble using the BSADF, simply because there is a say 5% probability of erroneously detecting a bubble for each value of r_2. Of course this can be resolved by using some variation of the Bonferroni correction (Dunn, 1958, 1961) or one of the numerous alternative procedures developed in the literature. However, since the GSADF includes all the relevant subsections in one test statistic, the significance level of the critical value will be the significance level for the entire test statistic. Thus, we will use the GSADF test to compare the size and power type properties of the alternative tests, and we will use the BSADF test for the date stamping comparisons. As described in section 3 a similar solution does no exist for the LPPL procedure.

2.2 The Robust Supremum ADF tests

Pedersen and Schütte (2017) find that the supremum ADF tests are significantly oversized when subject to serially correlated innovations. Previous literature has studied how a suitable choice of lag-length can retain the low size of the Dickey-Fuller test given serial correlation without losing too much power (Schwert, 1989; Ng and Perron, 1995, 2001). However, as argued in Pedersen and Schütte (2017) this is not applicable in the supremum ADF frameworks, because the test statistic is the supremum of a range of ADF test statistics. They propose an alternative way to calculate the critical values using the sieve bootstrap (Chang and Park, 2003; Palm et al., 2008) rather than iid errors to make the critical values robust to autocorrelation. See appendix A for the algorithm to calculate the robust critical values.

3 The Logarithmic Periodic Power Law Procedure

The LPPL procedure offers an alternative to the GSADF test. As described in the previous section the GSADF test requires the time series of both prices and fundamentals to construct the price-dividend time series in an attempt to eliminate the unit root. As mentioned this does not necessarily guarantee the removal of the unit root if prices and dividends do not share a unit root, and there might also be a unit root in the unobserved fundamentals, U_t. Furthermore, for almost all asset classes dividend data is collected at a relatively low frequency, and one need long series of data to test for the presence of a bubble. The LPPL procedure only considers the shape of the price trajectory in order to detect a potential build up of a bubble, and since prices are available at a much higher frequency it is possible to perform tests in a real time setting.

The LPPL procedure originates from the widely cited paper Johansen et al. (2000). In Sornette and Johansen (2001) it was defended from criticism put forward by Feigenbaum (2001) who found that the significance of the LPPL shape was fragile when

removing key data points. The model is motivated by interactions between a large number of traders that are allowed to influence one another[1].

3.1 The LPPL Model

Johansen et al. (2000) develop the model by introducing the following assumptions.

Assumption 3.1.

1. The asset pays no dividends.
2. The risk free asset pays zero interest rate.
3. Agents are risk neutral.
4. Markets clear automatically without the need of imposing any conditions.

The agents in the economy are claimed to be rational because they satisfy the rational expectations condition, given the assumptions above

$$\mathbb{E}_t\left(P_{t+1}\right) = P_t. \tag{4}$$

Another implication of the assumptions is that the fundamental value is zero, such that any $P_t > 0$ actually indicates a bubble. This is clearly not realistic and cannot be used as a testable criterion using empirical data.

Given that there is a bubble that has not collapsed yet, then the cumulative distribution function of a crash as a function of time is given by Q_t, with the corresponding probability density function $q_t = \frac{\mathrm{d}Q_t}{\mathrm{d}t}$ and hazard rate $h_t = \frac{q_t}{1-Q_t}$. We then assume that the price increases at some deterministic rate μ_t in the absence of the collapse of the bubble, whereas the price drops by a fixed fraction $\kappa \in (0, 1)$ if the bubble collapses.

[1] In the appendix we develop and present proofs for the theorems and corollaries below. A proof of the LPPL process is included in Bree and Joseph (2013) but relies on the use of Wolfram's Mathematica online integrator.

The price process can then be written as

$$\mathrm{d}P_t = \mu_t P_t \mathrm{d}t - \kappa P_t \mathrm{d}j, \tag{5}$$

where $j = 0$ before a crash and $j = 1$ during a crash and is associated with the hazard rate h_t such that $\mathbb{E}_t(\mathrm{d}j) = h_t \mathrm{d}t$. Then given risk neutral agents and zero interest rate the price process must satisfy the martingale condition

$$\mathbb{E}_t(\mathrm{d}P_t) = \mu_t P_t \mathrm{d}t - \kappa P_t \mathbb{E}_t(\mathrm{d}j) = 0, \tag{6}$$

which implies that

$$\mu_t = \kappa h_t.$$

The restriction $\mathbb{E}_t(\mathrm{d}P_t) = 0$ will be relaxed in section 3.3.

Hence, under the assumptions given above the investors require a zero return on their investment. But since there is a positive probability of a crash there must be a positive growth in the price exactly equal to the expected price drop from a crash. Plugging this into (5) we get prior to a crash (i.e. $\mathrm{d}j = 0$) that $\mathrm{d}P_t = \kappa h_t P_t \mathrm{d}t$, which is an ordinary differential equation with the solution

$$\log P_t = \log P_{t_0} + \kappa \int_{t_0}^{t} h_{t'} \mathrm{d}t'. \tag{7}$$

The traders are then modelled to rely partly on their connections and partly on idiosyncratic signals when forming their opinion on the price level. In particular the traders can be in either of two states $\{-1, +1\}$. The states can be understood as "buy" and "sell", such that the state of trader i is given by

$$s_i = \mathrm{sign}\left(K_t \sum_{j \in N(i)} s_j + \sigma \epsilon_i \right), \tag{8}$$

where $N(i)$ contains the indices of the traders that are connected with trader i, K is positive, and ϵ_i is a Gaussian random variable representing idiosyncratic noise or knowledge. A more general formulation of this model with q states and two types of interactions (positive for equal states and negative for different states) is known from statistical mechanics as the Potts model (Wu, 1982). The average state across all traders is then given by $M = \frac{1}{I}\sum_{i=1}^{I} s_i$, which has the expectation $\mathbb{E}(M) = 0$. To understand the dynamic features of the trader interactions we define the susceptibility of the system. First we augment (8) with a global influence term G such that

$$s_i = \text{sign}\left(K_t \sum_{j \in N(i)} s_j + \sigma\epsilon_i + G\right), \tag{9}$$

and then the susceptibility of the system is defined as

$$\chi = \left.\frac{\mathrm{d}\mathbb{E}(M)}{\mathrm{d}G}\right|_{G=0}.$$

The susceptibility of the system is the average sensitivity to a change in the global influence term. Similarly, the influence from one agent being in a particular state on another agent will be proportional to the susceptibility of the system. Thus, Johansen et al. (2000) argue that the susceptibility measures the agents' tendency to agree on e.g. a sell state, and that the process driving the susceptibility of the system is similar to that driving the hazard rate. It is important to define the structure of the trader connections in order to describe the susceptibility process. Johansen et al. (2000) use a hierarchical diamond lattice because it has already been solved for the Ising model (Derrida et al., 1983).

[Insert Figure 1 about here]

The hierarchical diamond lattice is constructed by starting with two connected traders as in figure 1a. Then two additional traders are introduced by replacing the single connection with a diamond as in figure 1b. Each of these connections is then

replaced by diamonds as in figure 1c. Then after p iterations there will be $\frac{2}{3}(2+4^p)$ traders and 4^p connections. In particular in figure 1c after $p=2$ iterations there are 12 traders and 16 connections. For this network of traders there exists a K_t in (9) and some critical value K_c such that $\chi < \infty$ when $K_t < K_c$, and $\chi \to \infty$ for $K_t \to K_c$. Then the first order approximation of the susceptibility is given by (see Derrida et al. (1983); Johansen et al. (2000))

$$\begin{aligned}\chi &\approx \mathcal{R}\left(A_0 (K_c - K_t)^{-\gamma} + A_1 (K_c - K_t)^{-\gamma+i\omega}\right) \\ &= A_0' (K_c - K_t)^{-\gamma} + A_1' (K_c - K_t)^{-\gamma} \cos\left(\omega \log (K_c - K_t) + \psi'\right),\end{aligned}$$

where $\mathcal{R}$ returns the real part of the argument and so A_0', A_1', ω and ψ' are real numbers.

Then as mentioned above we assume that the hazard rate follows a process similar to that of the susceptibility. We assume that K_t is sufficiently smooth and slow-moving, such that with t_c being the first time when $K_{t_c} = K_c$, we can write the approximation $K_c - K_t \propto t_c - t$. Then rewriting the approximated susceptibility the hazard rate is given by

$$h_t = B' (t_c - t)^{\beta-1} + C' (t_c - t)^{\beta-1} \cos\left(\omega \log (t_c - t) + \psi\right). \tag{10}$$

Then inserting h_t from (10) into (7) we get the following theorem.

Theorem 3.1. *Let assumptions 3.1 be satisfied and let the traders be connected by means of a hierarchical diamond lattice and influenced by their neighbours as given by (8). Then the price follows*

$$\log P_t = \log P_{t_c} + B (t_c - t)^{\beta} + C (t_c - t)^{\beta} \cos\left(\omega \log (t_c - t) + \phi\right), \tag{11}$$

where P_{t_c} is the price of the asset at the critical time, t_c, and B, C, ω, ϕ, and β are constants. See the proof in appendix B for their relation to the parameters in (10).

Proof. See appendix B. □

And then this can be rewritten such that the non-linear parameter ϕ is transformed into a linear parameter which will make the estimation procedure much more robust.

Theorem 3.2. *Let assumptions 3.1 be satisfied and let the traders be connected by means of a hierarchical diamond lattice and influenced by their neighbours as given by (8). Then the price follows*

$$\begin{aligned}\log P_t = \log P_{t_c} + B\left(t_c - t\right)^{\beta} &+ C_1\left(t_c - t\right)^{\beta}\cos\left(\omega\log\left(t_c - t\right)\right)\\ &+ C_2\left(t_c - t\right)^{\beta}\sin\left(\omega\log\left(t_c - t\right)\right),\end{aligned} \tag{12}$$

where $C_1 = C\cos\phi$ *and* $C_2 = -C\sin\phi$.

Proof. See appendix C. □

Since the hazard rate is a probability it must be positive. That gives us the following corollary.

Corollary 3.2.1. *The hazard rate is positive if*

$$B < \sqrt{\left(C_1^2 + C_2^2\right)\frac{\omega^2 + \beta^2}{\beta^2}}. \tag{13}$$

Proof. See appendix D. □

Corollary 3.2.2. *The non-oscillating part of the hazard rate is increasing if*

$$B\beta\left(\beta - 1\right) > 0. \tag{14}$$

Proof. see appendix E. □

3.2 Calibration of the LPPL Model

The model is calibrated in a two-step approach. Since $\log P_{t_c}$ is calibrated as a whole rather than by its components we denote $A = \log P_{t_c}$ for coherence, such that the model reads

$$\begin{aligned}\log P_t = A + B\,(t_c - t)^{\beta} &+ C_1\,(t_c - t)\cos\left(\omega \log\left(t_c - t\right)\right) \\ &+ C_2\,(t_c - t)^{\beta}\sin\left(\omega \log\left(t_c - t\right)\right). \end{aligned} \tag{15}$$

The linear parameters, A, B, C_1, and C_2 in (15), can be slaved to the estimates of the non-linear parameters, t_c, β, and ω, by rewriting the model into a least squares loss function (Filimonov and Sornette, 2013).

$$\begin{aligned} F(t_c,\beta,\omega,A,B,C_1,C_2) = \sum_{t=1}^{T} \Big[&\log P_t - A - B\,(t_c - t)^{\beta} \\ &- C_1\,(t_c - t)^{\beta}\cos\left(\omega \log\left(t_c - t\right)\right) \\ &- C_2\,(t_c - t)^{\beta}\sin\left(\omega \log\left(t_c - t\right)\right)\Big]^2 \\ = \sum_{t=1}^{T} &\left[\log P_t - A - Bf_t - C_1 g_t - C_2 h_t\right]^2, \end{aligned}$$

where, $y_i = \log P(t_i)$, $f_i = (t_c - t_i)^{\beta}$, $g_i = f_i \cos\left(\omega \log\left(t_c - t_i\right)\right)$, and $h_i = f_i \sin\left(\omega \log\left(t_c - t_i\right)\right)$. Hence, for given parameter values of t_c, β, and ω the loss function, F, is uniquely minimized by the least squares estimates of A, B, C_1, and C_2

$$\begin{pmatrix} N & \sum f_i & \sum g_i & \sum h_i \\ \sum f_i & \sum f_i^2 & \sum f_i g_i & \sum f_i h_i \\ \sum g_i & \sum f_i g_i & \sum g_i^2 & \sum g_i h_i \\ \sum h_i & \sum f_i h_i & \sum g_i h_i & \sum h_i^2 \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{B} \\ \hat{C}_1 \\ \hat{C}_2 \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum y_i f_i \\ \sum y_i g_i \\ \sum y_i h_i \end{pmatrix}. \tag{16}$$

The current LPPL literature does not provide any assumptions or results regarding any stochastic error term used for the least squares estimation. Hence, we do not

know the statistical properties of the obtained estimates. Interestingly Feigenbaum (2001) obtains standard errors on the least squares estimates. they do no, however, explain how they are obtained.

Given the method for obtaining the estimates of A, B, C_1, C_2, the non-linear parameters β, ω, and t_c are chosen to minimize the resulting mean squared error by running a simple Nelder-Mead search algorithm (Filimonov and Sornette, 2013).

The estimates are then saved to determine whether the fit indicates the presence of a bubble or not by checking if the estimates are economically meaningful (Johansen et al., 2000; Filimonov and Sornette, 2013). We require that $\beta \in (0.0, 0.8)$ to ensure that the hazard rate is increasing and $P_t \to P_{t_c}$ as $t \to t_c$. Rejecting low estimates of ω avoids the slow oscillations to just fit the trend, and rejecting too high estimates of ω avoids just fitting the noise. The two conditions (13) and (14) ensure that the hazard rate is positive and increasing respectively. For the purposes of this paper we impose the restriction $\omega \in (2, 13)$ as well which is similar to what is done in the literature. Furthermore, the root mean squared error (RMSE) is used in some studies as a goodness of fit criterion (Johansen et al., 1999, 2000; Graf and Meister, 2003; Brauers et al., 2014), and as will be evident in the Monte Carlo results in section 6 the choice of maximum RMSE will have a large impact on whether the LPPL procedure detects a bubble or not. As a technicality we also require that the matrix in (16) is non-singular and well conditioned. See section 7 and appendix G for a discussion on the bounds on β and ω.

[Insert Figure 2 about here]

Two examples of fitting the log periodic power law can be seen in the figures 2a and 2b showing an approved and rejected fit, respectively. The time series in figure 2a is a realization of the process described in section 5.4 and the times series in figure 2b is a realization of the process described in section 5.1. The plots show that the LPPL has quite some flexibility in fitting various processes, and it is clear

that the parameter requirements play an important role in detecting bubbles. The fit shown in figure 2a has all parameter values within the bounds, whereas the fit in figure 2b violates $\hat{\beta} \in (0.0, 0.8)$.

Similarly to the GSADF the sample is divided into subsections identified by r_0, r_1, and r_2, which are defined as in section 2. The LPPL is then calibrated based on each of the subsections and a bubble is detected for the entire sample if there is a fit satisfying the restrictions in at least one of the subsections. Note that there is no nominal size for the LPPL procedure given the lack of a test statistic. Hence, its objective is simply to detect as few bubbles, under the null. There is no measure of power either, however in this regard the objective is the same as for the GSADF; detect all of the bubbles.

3.3 The Generalized LPPL Framework

The model above is based on the assumptions 3.1 which result in a required rate of return of zero. In the following we present the consequences of a non-zero required rate of return. A non-zero required rate of return could be the result of allowing for a non-zero risk free rate or by allowing the agents to be risk averse. The generalized LPPL (GLPPL) model is obtained by rewriting the martingale condition in (6) into

$$\mathbb{E}_t(\mathrm{d}P_t) = \mu_t P_t \mathrm{d}t - \kappa P_t \mathbb{E}_t(\mathrm{d}j) = \gamma_t, \tag{17}$$

where γ_t the required rate of return that is now allowed to be non-zero and time varying.

Theorem 3.3. *Let the required rate of return be given by γ_t such that it is allowed to be non-zero and time varying. Then the logarithm of the price follows*

$$\log P_t = \log P_{t_c} + B\,(t_c - t)^{\beta} + C\,(t_c - t)^{\beta} \cos\left(\omega \log\left(t_c - t\right) + \phi\right) - \int_t^{t_c} \gamma_{t'} \mathrm{d}t'. \tag{18}$$

Proof. See appendix F. □

By allowing the required rate of return to be non-zero the integral of the required rate of return from the current time, t, to the critical point, t_c, is subtracted from the right-hand side in (12). The shape of this integral as a function of t depends on the shape of γ_t. In order to obtain further insights we can restrict γ_t to be constant.

Corollary 3.3.1. *Let the required rate of return be given by γ. Then the logarithm of the price follows*

$$\log P_t = \log P_{t_c} + B\left(t_c - t\right)^{\beta} + C\left(t_c - t\right)^{\beta} \cos\left(\omega \log\left(t_c - t\right) + \phi\right) - \gamma\left(t_c - t\right), \tag{19}$$

Proof. See appendix F. □

By allowing the required rate of return to be non-zero but restricting it to a constant reveals that a linear trend is added to $\log P_t$. The hazard function is the same as for the zero required return case and so are the conditions in lemma 3.2.1 and]3.2.2. By taking the exponential of either (18) or (19) it is clear that γ_t and γ will indeed enter as the continuously compounded return of P_t, which is intuitive as γ_t and γ were defined as the required return of the asset in (17).

3.4 Calibrating the Generalized LPPL Model

The generalized LPPL model is calibrated with a two-step approach like the standard LPPL model. We rewrite the model as in (15) such that

$$\log P_t = A + B\left(t_c - t\right)^{\beta} + \gamma t + C_1\left(t_c - t\right)\cos\left(\omega \log\left(t_c - t\right)\right) \tag{20}$$

$$+ C_2\left(t_c - t\right)^{\beta} \sin\left(\omega \log\left(t_c - t\right)\right), \tag{21}$$

where $A = \log P_{t_c} - \gamma t_c$. The linear coefficients A, B, C_1, C_2, and γ can then be slaved to the linear parameters t_c, β, and ω. Note how the extension of the model only adds linearly to the complexity of the model. Hence, we have the same number of non-

linear parameters as before. We rewrite this into a least squares loss function.

$$F(t_c, \beta, \omega, A, B, C_1, C_2, \gamma) = \sum_{t=1}^{T} \left[\log P_t - A - Bf_t - C_1 g_t - C_2 h_t - \gamma t\right]^2,$$

where $f_t = (t_c - t)^\beta$, $g_t = f_t \cos(\omega \log(t_c - t))$, and $h_t = f_t \sin(\omega \log(t_c - t))$.

As for the standard LPPL procedure the concepts of size and power are undefined. However, the interpretation is the same as for the standard LPPL procedure described in setion 3.1.

3.5 Time Stamping

Similarly to the method described in section 2.1 we can consider the various subsections of the sample in order to time stamp the bubbles identified with the LPPL and GLPPL frameworks. The main idea is very similar to Phillips et al. (2015), and deploys a backward looking recursive framework. The end of the subsection varies over $r_2 \in (r_0, 1)$, and for each value of r_2 the start of the subsection varies over $r_1 \in (0, r_2 - r_0)$. Then the start of the bubble is given by the first value of r_2 for which one of the subsections is found to contain a bubble, and the end of the bubble is found by the subsequent value of r_2 where no bubble is found in any of the subsections. See section 6.2 for a detailed description of the actual requirements used for the study.

4 Similarities and Differences

Until now the economic asset bubbles literature and the LPPL literature have been completely separated. The supremum ADF tests are based on economic theory by the notion that the bubble can be identified by the deviation of the observed price from the latent fundamental price. The link to economic theory for the LPPL test is not as clear. The literature does however provide some motivation based on herding behaviour of the traders, which is an established topic in the mainstream behavioural economics literature (Scharfstein and Stein, 1990; Froot et al., 1992). The motivation

for the structure of the herding behaviour, however, comes from physics rather than economics (Johansen et al., 2000).

The main driver of the build up and subsequent collapse of the bubble in the LPPL framework is that the traders are not synchronized in their actions initially and at some point, namely the critical point, they manage to coordinate a sell off. This is motivated by a certain structure of trader interaction and the critical points theory. A somewhat similar idea has been explored in the economics literature (Abreu and Brunnermeier, 2003). Here it is argued that an economy with rational traders can sustain a bubble because the traders are unable to synchronize their actions, and that the bubble will eventually collapse once the traders get coordinated. This can potentially be viewed as a theoretical link between the two sets of the literature.

The vast majority of the current LPPL literature has almost exclusively been evaluating the performance of the model using in-sample empirical data. Often the collapse of a bubble is defined as a remarkably large draw-down in asset prices and then the preceding time series is analysed using the LPPL framework (Johansen et al., 2000; Johansen and Sornette, 2010; Brauers et al., 2014). One exception is Graf and Meister (2003) which find that the predictive performance of the LPPL model can be very poor. This is because while the LPPL does indeed fit the time series well in the time preceding the collapse of a bubble it fits quiet periods without bubbles too. As a consequence it is difficult to know whether a good fit actually indicates the presence of a bubble or not. This paper evaluates the LPPL test against the supremum ADF tests more thoroughly than previous studies focusing on simulated data to create a controlled environment for the study.

The LPPL framework also faces challenges similar to those of the BSADF test. As discussed in section 2.1, the BSADF test will be sure to identify a bubble if just given a sufficiently fine partitioned time series. One could easily imagine that a subsection of the time series could provide an approved fit for the LPPL test simply by chance and then, by extension, one could also expect that the finer the time series is partitioned,

the higher would be the probability to detect a bubble in just one subsection. The main difference is that as an alternative to the BSADF test we have the GSADF test for which the significance level applies to the entire time series and not just the subsections, whereas we have no such correction for the LPPL procedure. There is also a distribution for the BSADF test statistic and so one could potentially correct for multiple testing as well.

A major difference between the supremum ADF and the LPPL is that the supremum ADF tests have a test statistic with a distribution under both the null and the alternative (Phillips et al., 2015). The LPPL does not even have a test statistic. This allowed Phillips et al. (2015) to do a study of the size and power of the supremum ADF tests, whereas that approach is not readily applicable to the LPPL. Hence, this study relies on the calculation of pseudo power and size measures simply obtained as bubble detection rates in scenarios containing and not containing bubbles.

The bubbles in the mainstream economic literature is said to be rational because the expectation of the size of the following period's bubble component in (1) is consistent with rational expectations (Campbell et al., 1997). The LPPL model, however, does also turn out to be consistent with rational expectations. Given the assumptions introduced in section 3.1 the fundamental value of the asset is zero, and the price in (2) can simply be reduced to $P_t = B_t$. Then substituting this into (4) and realizing that $R = 0$ in (1) because of the assumptions made we get

$$\mathbb{E}_t(B_{t+1}) = (1 + R)B_t = B_t$$

and it turns out that (1) and (4) are actually equivalent given the assumptions. Hence, the framework of the standard LPPL model is actually consistent with rational expectations. It is, however, unclear if the equivalence is intact after defining the network of agents in the LPPL model. Here the agents are allowed to influence the state of a subset of the agents simply by being in a certain state themselves. This

does not seem consistent with rationality but rather linked to behavioural finance. It provides an additional link to the theoretical model of Abreu and Brunnermeier (2003) as both models rely on a subgroup of the agents agents that is in a "positive" or "buy" state and thus not prepared to sell and let the bubble collapse. If the agents are assumed to feature rational expectations they are assumed to know the model when forming their expectations (Muth, 1961). However, that is inconsistent with the agents being influenced by a limited number of connected agents. Furthermore, no information regarding connections between agents is used for the forming of expectations, so it is not clear how they could be regarded as rational after all.

Both frameworks can be viewed as methods to search for structural breaks in the processes, i.e. periods with or without explosive roots and fits to deterministic functional forms. This could in principle refer to any number of economic concepts, like time varying risk premia, but it so happens in this literature that these shifts are considered "bubbles".

5 Simulated Bubble Dynamics

We choose various processes to simulate the dynamics of a potential bubble component. They have been chosen such that they cover a broad range of dynamics to determine how the LPPL model compares with the mainstream economic literature.

5.1 Random Walk without Bubble

To estimate the probability of a type I error we follow the procedure in Phillips et al. (2015) by simulating a unit root process without bubbles. This will represent the null hypothesis of no bubbles. The dynamics of the process are as follows.

$$P_t = T^{-1} + P_{t-1} + \epsilon_t, \quad \epsilon_t \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right),$$

where the innovations are iid and the parameters are set to $\sigma_\epsilon = 6.79$ and $P_0 = 100$ for the simulations. The parameter values are the same as those used in Phillips et al. (2015).

However, as argued in Pedersen and Schütte (2017) the GSADF does not perform well when innovations are serially correlated as the test statistics become significantly oversized. Thus, we also simulate processes with serially correlated innovations.

$$P_t = T^{-1} + P_{t-1} + \nu_t$$
$$\nu_t = \phi_1 \nu_{t-1} + \epsilon_t + \gamma_1 \epsilon_{t-1} + \gamma_2 \epsilon_{t-2} + \gamma_3 \epsilon_{t-3}, \qquad \epsilon_t \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right).$$

We consider two types of serial correlation. First we consider the case where the innovations follow an AR(1) model such that $\phi_1 = 0.8$ and $\gamma_1 = \gamma_2 = \gamma_3 = 0$. In the second case the innovations follow and MA(3) such that $\gamma_1 = \gamma_2 = \gamma_3 = 0.8$ and $\phi_1 = 0$. The remaining parameters are as in the iid case above. The values of the AR(1) and MA(3) coefficients are the same as those used in Pedersen and Schütte (2017) and are based on an extensive study of evidence of AR(1) and MA(3) serial correlation in housing markets of various countries.

We also simulate the random walk process with GARCH(1,1) innovations as an additional robustness measure.

$$P_t = T^{-1} + P_{t-1} + \epsilon_t$$
$$\epsilon_t = \nu_t \sqrt{h_t}, \quad \nu_t \overset{iid}{\sim} N(0,1)$$
$$h_t = \omega + \alpha \epsilon_{t-1}^2 + \beta h_{t-1},$$

with parameters set to $\omega = 10$, $\alpha = 0.3$, $\beta = 0.6$, and $P_0 = 100$.

A visual representation of the four processes described above can be seen in figures 3, 4, 5, and 6. Based on the visuals alone we judge that the time series could

resemble financial price processes to an extent that justifies including them in the study. Additional realisations of the processes are similar to those reported.

[Insert Figure 3 about here]

[Insert Figure 4 about here]

[Insert Figure 5 about here]

[Insert Figure 6 about here]

5.2 Partially Collapsing Bubbles

Even though the bubble is not permitted to collapse completely and then restart (Diba and Grossman, 1988b), the bubble may collapse partially and then initiate a new build up. If this is the case the testing methods originally proposed by Diba and Grossman (1988a,b) will not be able to detect the bubble consistently (Evans, 1991). However the GSADF test can successfully detect partially collapsing bubbles because of its multi-window framework (Phillips et al., 2015).

For the simulation of partially collapsing bubbles the fundamentals follow a random walk given by

$$D_t = \mu + D_{t-1} + \epsilon_{D_t}, \qquad \epsilon_{D_t} \overset{iid}{\sim} N\left(0, \sigma_D^2\right),$$

which implies that the fundamental price is given by

$$F_t = \frac{\mu\rho}{(1-\rho)^2} + \frac{\rho}{1-\rho} D_t.$$

The bubble component follows a threshold function such that the bubble grows at rate $1 + r$ when $B_t \leq \alpha$ and grows at an accelerated pace with the added possibility of

a collapse when $B_t > \alpha$. The bubble process is then given by

$$B_t = \begin{cases} (1+r)\, B_{t-1}\epsilon_{B_t}, & \text{if } B_{t-1} \leq \alpha \\ \left[\delta + \pi^{-1}\,(1+r)\,\theta_t\left(B_{t-1} - (1+r)^{-1}\,\delta\right)\right], & \text{if } B_{t-1} > \alpha. \end{cases}$$

where θ_t is a Bernoulli process taking value 1 with probability π and value 0 with probability $1-\pi$, where $0 < \pi \leq 1$. The bubble collapses with probability $1-\pi$ when $\theta = 0$. ρ is a discount factor and is related to the expected growth in the bubble component by $\rho^{-1} = 1 + r > 1$. After a collapse the bubble component is equal to δ where $0 < \delta < (1+r)\alpha$. ϵ_{B_t} is a positive random variable with $E_t\left[\epsilon_{B_{t+1}}\right] = 1$. Like Evans (1991), we choose $\epsilon_{B_t} = \exp(y_t\sigma_B - \sigma_B^2/2)$ where $y_t \overset{iid}{\sim} N(0,1)$. For consistency with (1) it must hold that $R > r$ for the slowly growing regime and $\pi^{-1}\,(1+r) > 1 + R$ for the fast growing regime. Furthermore, we use the scaling parameter κ such that the price process is given by $P_t = F_t + \eta B_t$.

We choose the parameter values as in Phillips et al. (2015) for comparability such that $\mu = 0.0024$, $\rho = 0.985$, $\alpha = 1$, $\pi = 0.85$, $\delta = 0.5$, $\sigma_D = 0.0316$, $\sigma_B = 0.05$, and $\eta = 20$.

A visual representation of the process can be found in figure 7. The partially collapsing bubbles appear to be rather steep during their build-up period. This does not mimic empirical prices particularly well. However, it has been used and referred to repeatedly in previous studies.

[Insert Figure 7 about here]

5.3 Random Walk with Explosive Periods

As an alternative to the bubble dynamics proposed by Evans (1991) we also simulate an explosive AR(1) process. We configure the process with either one or two bubble

periods. The process containing only a single bubble is given by

$$P_t = \begin{cases} P_{t-1} + \epsilon_t, & t = 1, ..., \tau_e - 1 \\ \delta_T P_{t-1} + \epsilon_t, & t = \tau_e, ..., \tau_f \\ P_{\tau_e} + \epsilon_t, & t = \tau_f + 1 \\ P_{t-1} + \epsilon_t, & t = \tau_f + 2, ..., T, \end{cases}$$

following the set-up in Phillips et al. (2015) and Pedersen and Schütte (2017). Similarly we let $\delta_T = 1 + cT^{-\alpha}$, and $\epsilon_t \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right)$. This yields a process that initially follows a random walk up until time τ_e, where it changes into an explosive process up until time τ_f, where it collapses to $P\tau_e$ and continues as a random walk. The thresholds defining the bubble period are constructed as $\tau_e = \lfloor r_e T \rfloor$ and $\tau_f = \lfloor r_f T \rfloor$. Again following Phillips et al. (2015) and Pedersen and Schütte (2017) we set the parameter values to $P_0 = 100$, $\sigma_\epsilon = 6.79$, $c = 1$, $\alpha = 0.8$, $r_e = 0.3$, and $r_f = 0.75$.

The two-period process is given by

$$P_t = \begin{cases} P_{t-1} + \epsilon_t, & t = 1, ..., \tau_{e_1} - 1 \\ \delta_T P_{t-1} + \epsilon_t, & t = \tau_{e_1}, ..., \tau_{f_1} \\ P_{\tau_{e_1}} + \epsilon_t, & t = \tau_{f_1} + 1 \\ P_{t-1} + \epsilon_t, & t = \tau_{f_1} + 2, ..., \tau_{e_2} - 1 \\ \delta_T P_{t-1} + \epsilon_t, & t = \tau_{e_2}, ..., \tau_{f_2} \\ P_{\tau_{e_2}} + \epsilon_t, & t = \tau_{f_2} + 1 \\ P_{t-1} + \epsilon_t, & t = \tau_{f_2} + 2, ..., T. \end{cases}$$

This creates a series with an explosive period ranging from $\tau_{e_1} = \lfloor r_{e_1} T \rfloor$ to $\tau_{f_1} = \lfloor r_{f_1} T \rfloor$ and a second explosive period ranging from $\tau_{e_2} = \lfloor r_{e_2} T \rfloor$ to $\tau_{f_2} = \lfloor r_{f_2} T \rfloor$. The explosive root $\delta_T = 1 + cT^{-\alpha}$ is defined as in the single-bubble random walk. We also choose the parameter values in line with those in Phillips et al. (2015) and Pedersen and

Schütte (2017), such that $P_0 = 100$, $c = 1$, $\alpha = 0.8$, $r_{e_1} = 0.1$, $r_{f_1} = 0.4$, $r_{e_2} = 0.6$, $r_{f_2} = 0.9$, and $\sigma_\epsilon = 6.79$.

Like the simulations under the null in section 5.1 we also simulate the single bubble processes with serially correlated innovations.

$$P_t = \begin{cases} P_{t-1} + \nu_t, & t = 1, ..., \tau_e - 1 \\ \delta_T P_{t-1} + \nu_t, & t = \tau_e, ..., \tau_f \\ P_{\tau_e} + \nu_t, & t = \tau_f + 1 \\ P_{t-1} + \nu_t, & t = \tau_f + 2, ..., T \end{cases}$$

$$\nu_t = \phi_1 \nu_{t-1} + \epsilon_t + \gamma_1 \epsilon_{t-1} + \gamma_2 \epsilon_{t-2} + \gamma_3 \epsilon_{t-3}, \qquad \epsilon_t \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right),$$

with $\sigma_\epsilon = 6.79$. As in section 5.1, the innovations follow either an AR(1) with $\phi_1 = 0.8$ and $\gamma_1 = \gamma_2 = \gamma_3 = 0$, or an MA(3) with $\gamma_1 = \gamma_2 = \gamma_3 = 0.8$ and $\phi_1 = 0$.

A visual representation of the four processes described above can be found in figures 8, 9, 10, and 11. These time series seem to resemble empirical prices rather well. Only the collapse of the bubbles might be a bit compact as it happens over a single time increment. However, since none of the tests use the shape of the collapse of the bubble to detect the bubbles, this should change little in terms of performance. We also discuss this issue in section 6.1.[2]

[Insert Figure 8 about here]

[Insert Figure 9 about here]

[Insert Figure 10 about here]

[Insert Figure 11 about here]

[2]See also section 7 for a general discussion on the validity of the simulated processes.

5.4 Random Walk with LPPL Bubble

We also create a bubble following a log periodic power law to compare the two procedures in an environment representing the LPPL framework. It is based on a random walk with a predetermined explosive period driven by the log periodic power law, so the price process will be a sum of what could roughly be described as a fundamental price and a bubble component.

$$P_t = F_t + B_t,$$

where F_t is a random walk

$$F_t = T^{-1} + F_{t-1} + \epsilon_t, \qquad \epsilon_t \overset{iid}{\sim} N\left(0, \sigma_\epsilon^2\right),$$

and B_t is the log periodic power law during the bubble period and zero otherwise.

$$B_t = \begin{cases} 0, & t = 1, ..., \tau_e - 1 \\ \exp\Big(A + B\left(t_c - t\right)^\beta + C_1\left(t_c - t\right)^\beta \cos\left(\omega\left(t_c - t\right)\right) & \\ \qquad + C_2\left(t_c - t\right)^\beta \sin\left(\omega\left(t_c - t\right)\right)\Big), & t = \tau_e, ..., \tau_f \\ 0, & t = \tau_f + 2, ..., T. \end{cases}$$

The parameters are chosen such that for the random walk $\sigma_\epsilon = 6.79$ and $F_0 = 100$, and for the log periodic power law $A = 8.5$, $B = -2$, $C_1 = C_2 = -0.5$, $\beta = 0.1$, $\omega = 6$, $r_e = 0.30$, $r_f = 0.75$, and $t_c = \tau_f + 1$. The values are chosen such that the resulting series provide bubbles that are similar in size to those formed by the random walk with explosive periods in section 5.3 as is illustrated in figures 12 and 13.

We also simulate the process with a positive required rate of return such that B_t

is given by

$$B_t = \begin{cases} 0, & t = 1, ..., \tau_e - 1 \\ \exp\Big(A + B\left(t_c - t\right)^{\beta} + C_1\left(t_c - t\right)^{\beta}\cos\left(\omega\left(t_c - t\right)\right) & \\ \qquad + C_2\left(t_c - t\right)^{\beta}\sin\left(\omega\left(t_c - t\right)\right) + \gamma t\Big), & t = \tau_e, ..., \tau_f \\ 0, & t = \tau_f + 2, ..., T, \end{cases}$$

where $\gamma = 0.01$.

A visual representation of the processes described above can be found in figures 12 and 13. This time series has some merit in mimicking empirical prices because of the exponential shape rather than a linear trend. Again the collapse of the bubble is rather steep, however as argued in section 6.1 this has little effect on the performance of the tests.

[Insert Figure 12 about here]

[Insert Figure 13 about here]

6 Monte Carlo

We perform a Monte Carlo study on each of the discussed bubble dynamics with 1,000 replications for time series with 100, 200, and 400 data points. For each time series we test for the presence of bubbles for various sub-periods. For all tests the minimum length of each sub-period is given by $\left\lfloor T\left(0.01 + 1.8T^{-\frac{1}{2}}\right)\right\rfloor$ as in Phillips et al. (2015), where T is the total number of observations. For the time series with 100, 200, and 400 data points the rule yields a minimum length of 19, 27, and 40 observations respectively.

Section 6.1 covers bubble detection rates and section 6.2 covers the abilities of the frameworks to time stamp the bubbles.

6.1 Bubble Detection

The bubble detection part of the simulation study is divided into two categories. First we estimate the probability of detecting a bubble in the different processes without bubbles. The standard random walk process will be a pure representation of the null hypothesis, and the processes with serially correlated and GARCH innovations will give some indication as to the robustness of the tests under alternative formulations of the null. The second category consists of the remaining processes that all include bubbles and represent different formulations of the alternative hypothesis.

We need to deal with a technicality regarding the robust GSADF. When the simulated bubbles collapse over a single time increment the error at that particular point in time will be extremely large and decrease the performance of the test if they are drawn in the bootstrap reshuffle. Since these large errors are not seen empirically we will replace all errors of a magnitude larger than 10 standard deviations with an iid random variable with the same standard deviation as the sampled errors. We do not consider this a severe manipulation, since draw-downs of this magnitude are not seen empirically and so this modification will only make the performance of the test reflect reality more closely. This does not affect the results presented here.

We use 199 replications to obtain the quantiles in the GSADF and robust GSADF tests [3].

[Insert Table 1 about here]

[Insert Table 2 about here]

Table 2 reports the Monte Carlo results for the time series with 100, 200, and 400 data points. Looking at the first four columns in table 2 we consider the processes without bubbles. We immediately confirm the finding in Pedersen and Schütte (2017) that the GSADF test is oversized for the processes with serially correlated innovations presented in the second and thirds columns. The solution proposed by Pedersen

[3] Thus ensuring that $\alpha(B+1)$ is an integer for B being the number of bootstrap replications and $\alpha = 0.05$ being the significance level (MacKinnon, 2002).

and Schütte (2017) and presented in section 2.2 reduces the size and generally brings the size back to the 5% level. Interestingly both the GSADF and the GSADF^B appear undersized under the null in [1]. For low T we observe that the GSADF and the GSADF^B are oversized for GARCH innovations and the problem becomes more severe as the sample size increases, reaching 0.32 for $T = 400$. This contradicts Phillips et al. (2015) who state that GARCH innovations do not appear to be an issue for the GSADF test. The finding is however in line with Harvey et al. (2016) where a wild bootstrap implementation is used to correct for the size.

We find that the two LPPL procedures LPPL$_{0.01}$ and GLPPL$_{0.01}$ outperform the GSADF in almost all of the cases. Just for $T = 400$ in [2] with AR(1) innovations does GLPPL$_{0.01}$ the GSADF with detection rates of 0.06 and 0.08, respectively[4]. The LPPL$_{0.01}$ and GLPPL$_{0.01}$ also have elevated detection rates for MA(3) innovations, but just not as severely as the GSADF. However, for $T = 400$ they are quite close when correcting for the respective objectives of the procedures. LPPL$_{0.1}$, LPPL$_{0.05}$, GLPPL$_{0.1}$, and GLPPL$_{0.05}$ all detect bubbles in almost all of the simulations.

Turning to columns [5] through [11] in table 2 we consider the processes with bubbles. The LPPL procedures LPPL$_{0.1}$ and LPPL$_{0.05}$ deliver very high detection rates for all processes regardless of the required rate of return. However, this is neither surprising nor impressive since their detection rates were almost as high for the processes without bubbles. The picture is more nuanced for the LPPL$_{0.01}$ and the GLPPL$_{0.01}$ as they do show a strict increase in the detection rates for the processes in [6], [7], and [8] compared to their non-bubble counterparts in [1], [2], and [3], respectively. They perform decent in detecting the partially collapsing bubbles in [5] for $T = 100$, as well as in [7] for $T = 200$ and $T = 400$, and in [8] for all T. We find a clear indication of the LPPL procedure being able to detect bubbles in those processes that are most often used for validating the tests in the mainstream economic literature. They do perform surprisingly poorly for the LPPL processes [9]

[4]Again, note that the objective for GLPPL is to get as close to zero as possible, whereas the GSADF tries to get as close to its theoretical size of 0.05. Thus the GSADF is close to its objective in this case.

and [10], except in [10] with $T = 400$. They struggle to detect bubbles for the random walk with two explosive periods in [11]. We find that the proposed generalization of the LPPL procedure performs slightly better than the standard LPPL procedure. For $T = 100$ and $T = 200$ the $\text{GLPPL}_{0.01}$ improves on $\text{LPPL}_{0.01}$ for seven of the eleven processes.

The GSADF and GSADF^B show higher detection rates compared to both LPPL procedures requiring $RMSE < 0.01$ for all processes. The GSADF and GSADF^B show decent performance in detecting the partially collapsing bubbles in [5] for $T = 100$ but improves for $T = 200$ and $T = 400$. In general, however, there is no uniform pattern in the detection rates across the different sample sizes. The GSADF and GSADF^B are also much better at detecting the LPPL bubbles finding almost all of them. Like the LPPL procedures they struggle to detect the two bubbles in [11] but not nearly as much. This supports the findings in Phillips et al. (2015) who also find high detection rates for the partially collapsing bubbles in [5] and low detection rates for the random walk with two bubbles in [11]. There are no large differences between the GSADF and the GSADF^B although the GSADF^B does lose some power in its attempt to bring down the size. The next question is how many of these bubble detections can actually be attributed to the detection of the actual bubble. We will illuminate that by investigating the ability of the various tests to time stamp the bubbles.

6.2 Time Stamping

Simulations for investigating each of the methods' ability to correctly time stamp the bubble period are only done for the five random walk processes and not the partially collapsing bubble process, because the starting and ending points must be fixed across the simulations. For the random walk processes, however, the bubble period is characterized by the start and end points given by $\lfloor r_e T \rfloor$ and $\lfloor r_f T \rfloor$ respectively where $r_e = 0.25$ and $r_f = 0.75$, such that the bubble has a total duration of $\lfloor r_d T \rfloor = \lfloor \left(r_f - r_e \right) T \rfloor$.

Then the following three requirements must all be met for the method to correctly identify the start and end of the bubble.

Condition 1. For at least 80% of the values of $r_2 \in (r_e + 0.5r_d, r_f)$ the method must identify a bubble in at least one of the relevant subsections formed by varying $r_1 \in (0, r_2 - r_0)$.

Condition 2. For at least 8 of the $\lfloor 0.5r_dT + 11 \rfloor$ values of $r_2 \in (r_e - 10, r_e + 0.5r_d)$ the method must be unable to detect a bubble in all of the relevant subsections formed by varying $r_1 \in (0, r_2 - r_0)$.

Condition 3. For at least 8 of the 11 values of $r_2 \in (r_f, r_f + 10)$ the method must be unable to detect a bubble in all of the relevant subsections formed by varying $r_1 \in (0, r_2 - r_0)$.

[Insert Figure 14 about here]

These requirements provide some flexibility in time stamping the beginning and the end of the bubble period, as the time stamping is approved in the study as long as they fall on $[r_e - 10, r_e + 0.5r_d]$ and $\left[r_f, r_f + 10\right]$ respectively. This flexibility is allowed because the methods cannot time stamp with perfect accuracy. Previous results showed close to zero success rate for stricter requirements. Thus, the simulations illustrate the relative performance of the procedures but not their absolute performance. Written as above this holds for the random walk processes containing a single bubble. For the random walk process with two bubble periods simply replace the shares r_x with r_{x_1} and r_{x_2} for $x = e$, f, and d.

[Insert Table 3 about here]

[Insert Table 4 about here]

Table 3 shows the share of the total amount of simulations that the test is able to time stamp the bubble correctly. Table 4 shows the adjusted shares of correctly time

stamped bubbles. The shares are adjusted by only including the simulations in which the relevant test actually detect a bubble rather than include all of the simulations. Notice that [11]* denotes the ability to detect either one of the two bubbles and [11]** denotes the ability to detect both bubbles simultaneously.

The BSADF generally performs better than the BSADF^B with few and small exceptions in processes [8] and [10]. There appears to be no impact of serially correlated innovations. The time stamping succes rate declines for the standard LPPL bubble in [9] and the two bubbles in [11]** for increasing sample sizes. Interestingly, introducing the positive required rate of return in the LPPL process [10] helps the BSADF and the BSADF^B procedures in time stamping the bubbles.

With such a large number of simulations there is a chance that some of the bubbles will be of a small magnitude relative to the underlying random walk, making the bubble impossible to detect for both the tests and the naked eye. To give a more precise measure of the time stamping performance we introduce the adjusted time stamping success rate by considering only those simulations in which a bubble is actually detected. This works fine for large bubble detection rates, but the rates introduce a lot of uncertainty for low detection rates as they rely on relatively few realizations. These adjusted rates show the rate at which the procedure time stamps the bubbles correctly given that it has already detected a bubble. They are by construction at least as high as the standard rates. We do not observe any changes in the relative performance of the BSADF and the BSADF^B, which is not too surprising given their similar detection rates. The adjusted rates are generally inflated compared to the standard rates, indicating that whenever the GSADF procedures detect a bubble it is mainly due to identifying the actual bubble. The time stamping success rates increase more for the bubble with a random walk and iid innovations in [6] compared to those with serially correlated innovations in [7] and [8]. There are no gains for the LPPL bubble with zero required rate of return in [9] with very low time stamping success rates compared to the other processes. In contrast we observe

substantial gains for the process with two bubbles in [11].

The standard time stamping success rates are quite low in general for all six of the LPPL procedures (and zero for the LPPL procedure requiring $RMSE \leq 0.01$ in particular). The $\text{LPPL}_{0.1}$ procedure does, however, compare with the BSADF procedures when it comes to the LPPL bubble with zero required rate of return in [9] for $T = 200$ and even outperforming for $T = 400$, and it shows performance comparable to the BSADF procedures for the two bubble process in [11] for $T = 100$. The $\text{GLPPL}_{0.1}$ does show somewhat higher time stamping rates for the bubble with a standard random walk in [6] as well as the LPPL bubble with positive required rate of return in [10] for $T = 400$. We also observe increasing success rates for the two single bubble random walks serially correlated innovations as T increases, but they are still low.

The adjusted success rates for the LPPL tests are similarly low and do not increase as was observed for BSADF procedures. This is due to the overall high detection rate of bubbles and indicates that the LPPL tests do not necessarily identify the actual bubble. According to these results it is very likely that the LPPL procedures spuriously detect a bubble in some random subsamble of the time series. The NaN entries occur when the procedure has failed to identify even a single bubble.

[Insert Table 5 about here]

[Insert Table 6 about here]

Table 5 shows the average fraction of the sample on which a bubble was first and last detected. These numbers only include the simulations where a bubble was detected in at least one of the subsections. Hence, some of the table entries can rely on quite few realizations for some of the procedures. In particular, NaN entries occur where no bubbles were detected in any of the simulations. To shed some light on the influence of outliers table 6 shows the median fraction of the sample on which a bubble is first detected. Recall from section 5 that for all processes the bubble begins at $r_e = 0.3$ and ends at $r_f = 0.75$.

The $\text{LPPL}_{0.1}$ and $\text{LPPL}_{0.05}$ generally underestimates r_e for all T. For $T = 100$ the $\text{GLPPL}_{0.1}$ and $\text{GLPPL}_{0.05}$ detect the bubbles with some delay, but for $T = 200$ and $T = 400$ $\text{GLPPL}_{0.05}$ comes quite close to estimating r_e. The early detections of bubbles is also a symptom of the fact that the procedures are prone to detect bubbles under the null. The GSADF and GSADF^B generally estimates r_e well in the processes [6], [9], and [10] and especially so for large T. They do, however, tend to underestimate r_e when exposed to serially correlated innovations in [7] and [8]. The performance is very poor for $\text{LPPL}_{0.01}$ and $\text{GLPPL}_{0.01}$ due to the small number of detected bubbles.

When it comes to estimating r_f, $\text{LPPL}_{0.1}$ and $\text{LPPL}_{0.05}$ come quite close for $T = 100$ but their performance deteriorates for larger T. $\text{GLPPL}_{0.1}$ performs slightly better, but only $\text{GLPPL}_{0.05}$ is able to show consistent performance for all T. Both the GSADF and GSADF^B show strong performance for the processes in [6], [9], and [10], and they estimate r_f with quite some delay for the serially correlated innovations in [7] and [8].

[Insert Table 7 about here]

[Insert Figure 15 about here]

[Insert Figure 17 about here]

Table 7 and figures 15 and 17 show the ability of the LPPL and GLPPL procedures to predict the most likely time of collapse. The correct value is $r_f = 0.75$. It is clear that our generalization of the LPPL procedure allowing for a positive required rate of return improves the estimate of the critical point. From table 7 we see that the center of the distribution is moved closer to the true value and from figures 15 and 17 we see that the distribution is much narrower for our GLPPL procedure.

7 Robustness

The Monte Carlo study has been run for various specifications of the price processes such as different innovation volatility and different magnitudes of explosiveness, and we have obtained similar results. Additionally, the simulations have been run with 100, 200, and 400 data points to determine the effect of the length of the time series or the frequency of the observations. Increasing the number of observations further proved computationally infeasible.

We also considered completely different formulations of the simulated processes, including more elaborate models where dividends follow a random walk and fundamentals are computed as in section 5.2. We also considered having the processes follow a log unit root to ensure positivity. Instead, simulations with negative prices were rerun. However, all these models yielded results similar to those reported here.

There are many different versions of the parameter restrictions for the LPPL procedure. See appendix G for an overview. We have made a choice on the restrictions that we consider in line with the existing literature, but in trying alternative restrictions we found little changes in the results regarding bubble detection.

We have studied the results using different formulations of the time stamping conditions presented in section 6.2, and when moving from 80% to 100% in condition 1, the shares of correct time stamps drop drastically. Also, when the methods are required to detect the bubble earlier than halfway through the expansive period the shares of correct time stamps drop quickly as well.

8 Conclusions

We find that the LPPL procedure has some interesting features since it does not rely on availability of dividend data and cointegration between prices and dividends. This allows for using higher frequency data since only the price process is needed. However, the strengths and weaknesses of the LPPL procedures varies greatly over the

different formulations. The $\text{LPPL}_{0.01}$ and $\text{GLPPL}_{0.01}$ procedures are better at not detecting a bubble under the null, but struggle to detect the bubbles when under the alternative. The $\text{LPPL}_{0.1}$, $\text{LPPL}_{0.05}$, $\text{GLPPL}_{0.1}$, and $\text{GLPPL}_{0.05}$ procedures are able to detect the bubbles under the alternative but also find bubbles under the null. All formulations struggle in time stamping the bubbles, but as shown in tables 5 and 6 the $\text{GLPPL}_{0.05}$ does show strong performance in finding the start and end points of the bubbles. These findings are, however, not necessarily in violation of the vast amount of empirical studies performed using the LPPL procedure that find a good fit on ex-post identified bubbles (Sornette et al., 1996; Johansen et al., 1999, 2000; Johansen and Sornette, 2010; Brauers et al., 2014). These studies all consider data samples in which a bubble is assumed to exist. Taken to the extreme, for example, it should not be hard to get a good fit during an explosive period if the model fits equally well regardless of the presence of a bubble. From the above analysis it seems that the LPPL procedure has too much flexibility in fitting the time series for bubble detection. Even for tighter bounds on the parameters the bubble detection shares under the null are still very close to 1 for the $\text{LPPL}_{0.1}$, $\text{LPPL}_{0.05}$, $\text{GLPPL}_{0.1}$, and $\text{GLPPL}_{0.05}$ procedures. We find that in general our suggested generalization of the LPPL procedure shows stronger performance than the standard formulation, especially for bubble detection and estimation of the most likely time of collapse. We confirm the findings in Pedersen and Schütte (2017) that the GSADF tests are oversized for serially correlated innovations and our findings contrast Phillips et al. (2015) in documenting oversized test statistics for GARCH innovation for both the standard GSADF and GSADF^{B} tests.

We find that the proposed time stamping method for the LPPL procedures works and is comparable to the BSADF and BSADF^{B} in some cases. However, this is limited to the $\text{GLPPL}_{0.1}$ and $\text{LPPL}_{0.1}$ which perform badly in the bubble detection study. The time stamping results support the conclusion that the LPPL procedures can detect bubbles in any subsection and not just the subsection containing the actual bubble.

There is no increase when looking at the adjusted share of correctly time stamped bubbles for the LPPL procedures but there is for the BSADF and BSADF^B procedures.

Finally there is the obvious challenge with the LPPL procedure that there is no actual test statistic with a distribution. As mentioned one can use the RMSE but it is not easily generalized for different processes and from application to application. Consider tuning the maximum RMSE such that the detection rate is close to 5% under the null. Earlier results showed that the detection rates under the alternative hypothesis would be of similar magnitude. Hence, there is very little to be gained from trying to fine-tune the required RMSE and it would also be very hard to do reliably on empirical data. While the LPPL framework relies on the choice of maximum RMSE the supremum ADF tests also relies on the choice of the significance level. We point out that this choice is also subjective and arbitrary to some extend.

References

Abreu, D. and Brunnermeier, M. K. (2003). Bubbles and Crashes. *Econometrica*, 71(1):173–204.

Brauers, M., Thomas, M., and Zietz, J. (2014). Are There Rational Bubbles in REITs? New Evidence from a Complex Systems Approach. *The Journal of Real Estate Finance and Economics*, 49(2):165–184.

Bree, D. S. and Joseph, N. L. (2013). Fitting the Log Periodic Power Law to financial crashes: a critical analysis.

Campbell, J. Y., Lo, A. W., and MacKinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton University Press, 1 edition.

Chang, Y. and Park, J. Y. (2003). A Sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis*, 24(4):379–400.

Cuñado, J., Gil-Alana, L. A., and de Gracia, F. P. (2005). A test for rational bubbles in the NASDAQ stock index: A fractionally integrated approach. *Journal of Banking and Finance*, 29(10):2633–2654.

Derrida, B., De Seze, L., and Itzykson, C. (1983). Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics*, 33(3):559–569.

Diba, B. T. and Grossman, H. L. (1988a). Explosive Rational Bubbles in Stock Prices? *American Economic Review*, 78(3):520–530.

Diba, B. T. and Grossman, H. L. (1988b). The Theory of Rational Bubbles in Stock Prices. *The Economic Journal*, 98(392):746–754.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366):427–431.

Dunn, O. J. (1958). Estimation of the Means of Dependent Variables. *The Annals of Mathematical Statistics*, 29(4):1095–1111.

Dunn, O. J. (1961). Multiple Comparisons among Means. *Journal of the American Statistical Association*, 56(293):52–64.

Engsted, T. (2006). Explosive bubbles in the cointegrated VAR model. *Finance Research Letters*, 3(2):154–162.

Engsted, T. and Nielsen, B. (2012). Testing for rational bubbles in a coexplosive vector autoregression. *Econometrics Journal*, 15(2):226–254.

Evans, G. W. (1991). Pitfalls in Testing for Explosive Bubbles in Asset Prices. *The American Economic Review*, 81(4):922–930.

Feigenbaum, J. A. (2001). More on a statistical analysis of log-periodic precursors to financial crashes. *Quantitative Finance*, 1(5):527–532.

Filimonov, V. and Sornette, D. (2013). A stable and robust calibration scheme of the log-periodic power law model. *Physica A: Statistical Mechanics and its Applications*, 392(17):3698–3707.

Frömmel, M. and Kruse, R. (2012). Testing for a rational bubble under long memory. *Quantitative Finance*, 12(11):1723–1732.

Froot, K. A. and Obstfeld, M. (1991). Intrinsic Bubbles: The Case of Stock Prices. *American Economic Review*, 81(5):1189–1214.

Froot, K. A., Scharfstein, D. S., and Stein, J. C. (1992). Herd on the Street : Informational Inefficiencies in a Market with Short-Term Speculation. *The Journal of Finance*, 47(4):1461–1484.

Fuller, W. A. (1976). *Introduction to Statistical Time Series*. John Wiley & Sons.

Geraskin, P. and Fantazzini, D. (2013). Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask. *The European Journal of Finance*, 19(5):366–391.

Graf, H. C. and Meister, C. (2003). Predicting critical crashes? A new restriction for the free variables. *Physica A: Statistical Mechanics and its Applications*, 320:539–547.

Gürkaynak, R. S. (2008). Econometric tests of asset price bubbles: Taking stock. *Journal of Economic Surveys*, 22(1):166–186.

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

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

Johansen, A., Ledoit, O., and Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance*, 3:219–255.

Johansen, A. and Sornette, D. (2010). Shocks , Crashes and Bubbles in Financial Markets. *Brussels Economics Review*, 53(2):201–254.

Johansen, A., Sornette, D., and Ledoit, O. (1999). Predicting Financial Crashes Using Discrete Scale Invariance. *Journal of Risk*, 1(4):5–32.

Leroy, S. F. and Porter, R. D. (1981). The Present-Value Relation : Tests Based on Implied Variance Bounds. *Econometrica*, 49(3):555–574.

MacKinnon, J. G. (2002). Bootstrap Inference in Econometrics. *The Canadian Journal of Economics*, 35(4):615–645.

Muth, J. F. (1961). Rational Expectations and the Theory of Price Movements. *Econometrica*, 29(3):315–335.

Ng, S. and Perron, P. (1995). Unit Root Tests ARMA Models with Data Dependent Methods for the Selection of the Truncation Lag. *Journal of the American Statistical Association*, 90(429):268–281.

Ng, S. and Perron, P. (2001). Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power. *Econometrica*, 69(6):1519–1554.

Palm, F. C., Smeekes, S., and Urbain, J.-P. (2008). Bootstrap Unit Root Tests : Comparison and Extensions. *Journal of Time Series Analysis*, 29(2):371–401.

Pedersen, T. Q. and Schütte, E. C. M. (2017). Testing for Explosive Bubbles in the Presence of Autocorrelated Innovations. *CREATES Research Paper*, pages 1–34.

Phillips, P. C. B., Shi, S., and Yu, J. (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., Shi, S., and Yu, J. (2015). 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., Wu, Y., and Yu, J. (2011). Explosive Behavior In The 1990S Nasdaq: When Did Exuberance Escalate Asset Values? *International Economic Review*, 52(1):201–226.

Scharfstein, D. S. and Stein, J. C. (1990). Herd Behavior and Investment. *The American Economic Review*, 80(3):465–479.

Schwert, G. W. (1989). Tests for Unit Roots: A Monte Carlo Investigation. *Journal of Business & Economic Statistics*, 7(2):5–17.

Shiller, R. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *The American Economic Review*, 71:421–436.

Sornette, D. and Johansen, A. (1998). A hierarchical model of financial crashes. *Physica A: Statistical Mechanics and its Applications*, 261(3):581–598.

Sornette, D. and Johansen, A. (2001). Significance of log-periodic precursors to financial crashes. *Quantitative Finance*, 1(4):452–471.

Sornette, D., Johansen, A., and Bouchaud, J.-P. (1996). Stock Market Crashes, Precursors and Replicas. *Journal de Physique I*, 6(1):167–175.

Sornette, D., Woodard, R., Yan, W., and Zhou, W. X. (2013). Clarifications to questions and criticisms on the Johansen-Ledoit-Sornette financial bubble model. *Physica A: Statistical Mechanics and its Applications*, 392(19):4417–4428.

West, K. D. (1987). A Specification Test for Speculative Bubbles. *Quarterly Journal of Economics*, 102(3):553–580.

Wosnitza, J. H. and Leker, J. (2014). Why credit risk markets are predestined for exhibiting log-periodic power law structures. *Physica A: Statistical Mechanics and its Applications*, 393:427–449.

Wu, F. Y. (1982). The Potts Model. *Reviews of Modern Physics*, 54(1):235–268.

Tables

Table 1: Overview of simulated processes used for the Monte Carlo study.

[1]	Random walk with iid innovations
[2]	Random walk with AR(1) innovations
[3]	Random walk with MA(3) innovations
[4]	Random walk with GARCH innovations
[5]	Partially collapsing bubbles
[6]	Random walk with bubble and iid innovations
[7]	Random walk with bubble and AR(1) innovations
[8]	Random walk with bubble and MA(3) innovations
[9]	Bubble following log-periodic-power-law
[10]	Bubble following log-periodic-power-law and positive required return
[11]	Random walk with two bubbles and iid innovations

Table 2: Bubble detection rates

	Processes without Bubbles				Processes with Bubbles						
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
$T = 100$											
$LPPL_{0.1}$	1.0000	0.9990	0.9960	1.0000	1.0000	0.9830	0.9620	0.9710	0.9600	0.9930	0.9760
$LPPL_{0.05}$	1.0000	0.9990	0.9960	1.0000	1.0000	0.8790	0.8570	0.9060	0.8040	0.9040	0.8070
$LPPL_{0.01}$	0.0020	0.0840	0.2510	0.0020	0.2950	0.0490	0.1870	0.3690	0.0010	0.0120	0.0000
$GLPPL_{0.1}$	0.9440	0.9200	0.8820	0.9610	0.9740	0.9630	0.8240	0.7960	0.9550	0.9900	0.9200
$GLPPL_{0.05}$	0.9440	0.9200	0.8820	0.9610	0.9740	0.8730	0.6570	0.6990	0.8110	0.9300	0.7630
$GLPPL_{0.01}$	0.0020	0.0530	0.1710	0.0040	0.3100	0.0500	0.1410	0.2610	0.0080	0.1010	0.0040
GSADF	0.0320	0.6420	0.4120	0.1530	0.5010	0.7380	0.8380	0.9180	1.0000	1.0000	0.3930
$GSADF^B$	0.0380	0.0560	0.0620	0.1300	0.4640	0.6180	0.7840	0.8690	0.9590	1.0000	0.2460
$T = 200$											
$LPPL_{0.1}$	1.0000	1.0000	0.9970	1.0000	1.0000	0.9890	0.9890	0.9930	0.9790	0.9920	0.9780
$LPPL_{0.05}$	1.0000	1.0000	0.9970	1.0000	1.0000	0.8840	0.9090	0.9610	0.8630	0.9320	0.7790
$LPPL_{0.01}$	0.0020	0.0600	0.3050	0.0070	0.1090	0.0950	0.2520	0.4950	0.0020	0.0610	0.0000
$GLPPL_{0.1}$	1.0000	1.0000	0.9940	1.0000	1.0000	0.9900	0.9670	0.9790	0.9920	1.0000	0.9840
$GLPPL_{0.05}$	1.0000	1.0000	0.9940	1.0000	1.0000	0.9170	0.8870	0.9280	0.9200	0.9760	0.8380
$GLPPL_{0.01}$	0.0010	0.0490	0.2740	0.0070	0.1930	0.0990	0.2500	0.4640	0.0040	0.0660	0.0000
GSADF	0.0430	0.7290	0.4490	0.2220	0.7890	0.7070	0.9010	0.9260	1.0000	1.0000	0.4160
$GSADF^B$	0.0370	0.0660	0.0530	0.1790	0.7420	0.5590	0.8470	0.8850	0.9710	1.0000	0.2210
$T = 400$											
$LPPL_{0.1}$	1.0000	0.9910	0.9790	1.0000	1.0000	0.9970	0.9930	0.9980	0.9940	0.9980	0.9820
$LPPL_{0.05}$	1.0000	0.9910	0.9790	1.0000	1.0000	0.9040	0.9360	0.9760	0.8920	0.9740	0.7970
$LPPL_{0.01}$	0.0020	0.0390	0.2700	0.0180	0.0630	0.1900	0.3330	0.5310	0.0000	0.2900	0.0040
$GLPPL_{0.1}$	1.0000	0.9910	0.9790	1.0000	1.0000	1.0000	0.9930	0.9950	0.9980	1.0000	0.9910
$GLPPL_{0.05}$	1.0000	0.9910	0.9790	1.0000	1.0000	0.9560	0.9540	0.9730	0.9290	0.9980	0.8820
$GLPPL_{0.01}$	0.0020	0.0600	0.2710	0.0190	0.1090	0.1680	0.3310	0.5590	0.0040	0.3100	0.0120
GSADF	0.0560	0.7230	0.4250	0.3220	0.9390	0.7620	0.8830	0.9250	1.0000	1.0000	0.4480
$GSADF^B$	0.0480	0.0500	0.0520	0.2330	0.9040	0.6650	0.8480	0.8820	0.9960	1.0000	0.2460

Table 3: Time stamping success rate

	[6]	[7]	[8]	[9]	[10]	[11]*	[11]**
$T = 100$							
$\text{LPPL}_{0.1}$	0.2500	0.0390	0.0350	0.1580	0.1580	0.3050	0.0390
$\text{LPPL}_{0.05}$	0.0350	0.0050	0.0100	0.0070	0.0040	0.0960	0.0030
$\text{LPPL}_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\text{GLPPL}_{0.1}$	0.0400	0.0020	0.0020	0.0060	0.0280	0.0570	0.0000
$\text{GLPPL}_{0.05}$	0.0130	0.0000	0.0000	0.0000	0.0000	0.0250	0.0000
$\text{GLPPL}_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BSADF	0.6460	0.4710	0.4520	0.3130	0.9990	0.3580	0.0980
BSADF^B	0.5420	0.3920	0.4360	0.2700	0.9990	0.3360	0.0390
$T = 200$							
$\text{LPPL}_{0.1}$	0.2060	0.0450	0.0520	0.1250	0.0330	0.1970	0.0080
$\text{LPPL}_{0.05}$	0.0200	0.0060	0.0090	0.0000	0.0000	0.0260	0.0000
$\text{LPPL}_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\text{GLPPL}_{0.1}$	0.0990	0.0100	0.0070	0.0230	0.0300	0.0690	0.0020
$\text{GLPPL}_{0.05}$	0.0090	0.0000	0.0010	0.0000	0.0020	0.0270	0.0000
$\text{GLPPL}_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BSADF	0.5690	0.5470	0.4810	0.1590	0.9890	0.3420	0.0670
BSADF^B	0.4430	0.5160	0.5060	0.0680	0.9950	0.2860	0.0280
$T = 400$							
$\text{LPPL}_{0.1}$	0.1760	0.1000	0.0950	0.1070	0.1140	0.1300	0.0040
$\text{LPPL}_{0.05}$	0.0220	0.0210	0.0240	0.0020	0.0010	0.0080	0.0000
$\text{LPPL}_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$\text{GLPPL}_{0.1}$	0.1680	0.0390	0.0230	0.0350	0.2060	0.1110	0.0100
$\text{GLPPL}_{0.05}$	0.0110	0.0020	0.0010	0.0000	0.0020	0.0260	0.0010
$\text{GLPPL}_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BSADF	0.6170	0.6110	0.5670	0.0300	0.9770	0.3450	0.0660
BSADF^B	0.4650	0.5810	0.6060	0.0070	0.9940	0.3030	0.0230

Table 4: Adjusted time stamping success ratios

	[6]	[7]	[8]	[9]	[10]	[11]*	[11]**
$T = 100$							
$LPPL_{0.1}$	0.2543	0.0405	0.0360	0.1646	0.1591	0.3125	0.0400
$LPPL_{0.05}$	0.0398	0.0058	0.0110	0.0087	0.0044	0.1190	0.0037
$LPPL_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	NaN	NaN
$GLPPL_{0.1}$	0.0415	0.0024	0.0025	0.0063	0.0283	0.0620	0.0000
$GLPPL_{0.05}$	0.0149	0.0000	0.0000	0.0000	0.0000	0.0328	0.0000
$GLPPL_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BSADF	0.8374	0.5621	0.4902	0.3130	0.9990	0.6183	0.2317
$BSADF^B$	0.7767	0.4949	0.5017	0.2805	0.9990	0.6626	0.0992
$T = 200$							
$LPPL_{0.1}$	0.2083	0.0455	0.0524	0.1277	0.0333	0.2014	0.0082
$LPPL_{0.05}$	0.0226	0.0066	0.0094	0.0000	0.0000	0.0334	0.0000
$LPPL_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	NaN	NaN
$GLPPL_{0.1}$	0.1000	0.0103	0.0072	0.0232	0.0300	0.0701	0.0020
$GLPPL_{0.05}$	0.0098	0.0000	0.0011	0.0000	0.0020	0.0322	0.0000
$GLPPL_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	NaN	NaN
BSADF	0.7751	0.6071	0.5194	0.1590	0.9890	0.6178	0.1810
$BSADF^B$	0.7048	0.6080	0.5718	0.0700	0.9950	0.6606	0.0673
$T = 400$							
$LPPL_{0.1}$	0.1765	0.1007	0.0952	0.1076	0.1142	0.1324	0.0041
$LPPL_{0.05}$	0.0243	0.0224	0.0246	0.0022	0.0010	0.0100	0.0000
$LPPL_{0.01}$	0.0000	0.0000	0.0000	NaN	0.0000	0.0000	0.0000
$GLPPL_{0.1}$	0.1680	0.0393	0.0231	0.0351	0.2060	0.1120	0.0101
$GLPPL_{0.05}$	0.0115	0.0021	0.0010	0.0000	0.0020	0.0295	0.0011
$GLPPL_{0.01}$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
BSADF	0.7848	0.6920	0.6130	0.0300	0.9770	0.6004	0.1789
$BSADF^B$	0.6617	0.6851	0.6871	0.0071	0.9940	0.6829	0.0513

Table 5: Mean start and end of bubble

	[6]		[7]		[8]		[9]		[10]	
	r_e	r_f	r_e	r_f	r_e	r_f	r_e	r_f	r_e	r_f
$T = 100$										
$\text{LPPL}_{0.1}$	0.2900	0.8036	0.3375	0.7479	0.3439	0.7276	0.2965	0.7796	0.2917	0.7951
$\text{LPPL}_{0.05}$	0.3423	0.7361	0.3986	0.6971	0.3845	0.6846	0.3234	0.6894	0.3484	0.7265
$\text{LPPL}_{0.01}$	0.7014	0.7169	0.6139	0.6502	0.5803	0.6660	0.9700	0.9700	0.6783	0.6825
$\text{GLPPL}_{0.1}$	0.4072	0.7420	0.4634	0.6744	0.4708	0.6645	0.4613	0.7560	0.4155	0.7683
$\text{GLPPL}_{0.05}$	0.4322	0.7136	0.4883	0.6445	0.4763	0.6359	0.4858	0.7282	0.4609	0.7451
$\text{GLPPL}_{0.01}$	0.6802	0.6986	0.6145	0.6415	0.5960	0.6485	0.7350	0.7375	0.7399	0.7428
GSADF	0.4065	0.7642	0.2958	0.8455	0.2717	0.8712	0.4753	0.7676	0.4362	0.7662
GSADF^B	0.4202	0.7553	0.2967	0.7994	0.2695	0.8285	0.4592	0.7635	0.4213	0.7603
$T = 200$										
$\text{LPPL}_{0.1}$	0.2106	0.8742	0.2529	0.8287	0.2453	0.8156	0.2132	0.8074	0.2172	0.7674
$\text{LPPL}_{0.05}$	0.3133	0.7938	0.3622	0.7732	0.3017	0.7637	0.3051	0.7021	0.3275	0.6920
$\text{LPPL}_{0.01}$	0.6879	0.7096	0.6296	0.6813	0.5693	0.6798	0.8025	0.8025	0.6436	0.6484
$\text{GLPPL}_{0.1}$	0.3052	0.8230	0.3516	0.7736	0.3451	0.7473	0.3149	0.8219	0.2998	0.8254
$\text{GLPPL}_{0.05}$	0.3628	0.7738	0.4273	0.7287	0.3887	0.7171	0.3915	0.7652	0.3750	0.7653
$\text{GLPPL}_{0.01}$	0.6787	0.7073	0.6244	0.6734	0.5836	0.6763	0.6763	0.7250	0.6631	0.6725
GSADF	0.3510	0.7811	0.2090	0.8931	0.1874	0.9360	0.3579	0.7921	0.3378	0.7774
GSADF^B	0.3949	0.7601	0.2245	0.8362	0.1987	0.8915	0.3756	0.7771	0.3457	0.7640
$T = 400$										
$\text{LPPL}_{0.1}$	0.1627	0.8946	0.1945	0.8769	0.1898	0.8773	0.1684	0.8619	0.1739	0.8009
$\text{LPPL}_{0.05}$	0.3128	0.8132	0.3380	0.8042	0.2664	0.8310	0.3112	0.7484	0.3353	0.7312
$\text{LPPL}_{0.01}$	0.6718	0.7162	0.6221	0.6864	0.5605	0.7027	NaN	NaN	0.6361	0.6518
$\text{GLPPL}_{0.1}$	0.2238	0.8763	0.2531	0.8559	0.2590	0.8412	0.2297	0.8678	0.2302	0.8734
$\text{GLPPL}_{0.05}$	0.3381	0.8132	0.3711	0.8051	0.3333	0.8052	0.3561	0.8052	0.3374	0.8151
$\text{GLPPL}_{0.01}$	0.6713	0.7081	0.6105	0.6833	0.5693	0.6973	0.7256	0.7275	0.6461	0.6685
GSADF	0.2858	0.8048	0.1523	0.9376	0.1377	0.9616	0.2710	0.8307	0.2473	0.8264
GSADF^B	0.3528	0.8048	0.1689	0.9376	0.1483	0.9616	0.3309	0.7966	0.2880	0.7746

Table 6: Median start and end of bubble

	[6]		[7]		[8]		[9]		[10]	
	r_e	r_f	r_e	r_f	r_e	r_f	r_e	r_f	r_e	r_f
$T = 100$										
$\text{LPPL}_{0.1}$	0.2700	0.7500	0.3000	0.7200	0.3100	0.7200	0.2600	0.7400	0.2700	0.7300
$\text{LPPL}_{0.05}$	0.2800	0.7400	0.3600	0.7000	0.3400	0.6900	0.2600	0.6900	0.2900	0.7000
$\text{LPPL}_{0.01}$	0.7000	0.7200	0.6200	0.6600	0.6100	0.6800	0.9700	0.9700	0.6800	0.6800
$\text{GLPPL}_{0.1}$	0.3900	0.7400	0.4500	0.6900	0.4600	0.6800	0.4100	0.7400	0.3900	0.7500
$\text{GLPPL}_{0.05}$	0.4100	0.7300	0.4800	0.6700	0.4700	0.6600	0.4600	0.7400	0.4300	0.7500
$\text{GLPPL}_{0.01}$	0.6900	0.7200	0.6200	0.6600	0.6100	0.6600	0.7400	0.7400	0.7500	0.7500
GSADF	0.3900	0.7500	0.2800	0.8600	0.2500	0.9400	0.5400	0.7500	0.4900	0.7500
GSADF^B	0.4200	0.7500	0.2800	0.7200	0.2400	0.7200	0.5400	0.7500	0.4650	0.7500
$T = 200$										
$\text{LPPL}_{0.1}$	0.1800	0.9550	0.2150	0.9050	0.2150	0.8500	0.1800	0.7500	0.1750	0.7400
$\text{LPPL}_{0.05}$	0.2175	0.8850	0.3250	0.7350	0.2700	0.7350	0.2100	0.7150	0.2250	0.7200
$\text{LPPL}_{0.01}$	0.7100	0.7300	0.6350	0.7000	0.5850	0.6950	0.8025	0.8025	0.6350	0.6400
$\text{GLPPL}_{0.1}$	0.2750	0.7500	0.3200	0.7350	0.3150	0.7350	0.2650	0.7500	0.2700	0.7500
$\text{GLPPL}_{0.05}$	0.3500	0.7500	0.4100	0.7200	0.3550	0.7200	0.3450	0.7350	0.3600	0.7250
$\text{GLPPL}_{0.01}$	0.6950	0.7300	0.6350	0.6925	0.6050	0.6900	0.6700	0.6850	0.6450	0.6650
GSADF	0.3600	0.7500	0.1850	0.9400	0.1650	0.9850	0.3800	0.7500	0.3750	0.7500
GSADF^B	0.4000	0.7500	0.2000	0.7350	0.1750	0.9450	0.4200	0.7500	0.3850	0.7500
$T = 400$										
$\text{LPPL}_{0.1}$	0.1275	0.9575	0.1600	0.9350	0.1675	0.9325	0.1275	0.9325	0.1300	0.8525
$\text{LPPL}_{0.05}$	0.2350	0.8925	0.2875	0.8725	0.2250	0.8962	0.2387	0.8525	0.2775	0.7125
$\text{LPPL}_{0.01}$	0.6837	0.7350	0.6375	0.7100	0.5700	0.7100	NaN	NaN	0.6713	0.6825
$\text{GLPPL}_{0.1}$	0.1875	0.9300	0.2225	0.9150	0.2275	0.8700	0.1900	0.9187	0.1950	0.9250
$\text{GLPPL}_{0.05}$	0.3050	0.7500	0.3425	0.7425	0.3100	0.7425	0.3250	0.7500	0.3362	0.7375
$\text{GLPPL}_{0.01}$	0.6850	0.7250	0.6225	0.7025	0.5700	0.7050	0.7325	0.7362	0.6675	0.6800
GSADF	0.2900	0.7500	0.1275	0.9750	0.1175	0.9950	0.2400	0.7550	0.2500	0.7550
GSADF^B	0.3625	0.7500	0.1450	0.9750	0.1300	0.9950	0.3150	0.7550	0.3325	0.7550

Table 7: Mean and median of critical point estimates.

	[6]	[7]	[8]	[9]	[10]
			$T = 100$		
LPPL					
Mean	0.5931	0.6346	0.6587	0.6780	0.6307
Median	0.5893	0.6279	0.6562	0.6802	0.6245
GLPPL					
Mean	0.6401	0.6408	0.6399	0.7174	0.6929
Median	0.6451	0.6326	0.6402	0.7415	0.7181
			$T = 200$		
LPPL					
Mean	0.5679	0.5886	0.6033	0.6227	0.6027
Median	0.5620	0.5850	0.5986	0.6247	0.6004
GLPPL					
Mean	0.6231	0.6083	0.6061	0.6632	0.6512
Median	0.6150	0.6089	0.6100	0.6741	0.6585
			$T = 400$		
LPPL					
Mean	0.5539	0.5662	0.5731	0.6088	0.5946
Median	0.5499	0.5591	0.5678	0.6116	0.5975
GLPPL					
Mean	0.5993	0.5917	0.5945	0.6286	0.6072
Median	0.5950	0.5875	0.5913	0.6389	0.6088

Figures

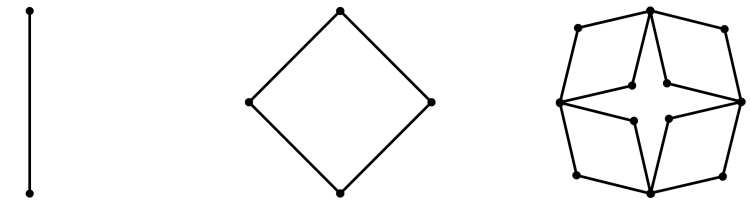
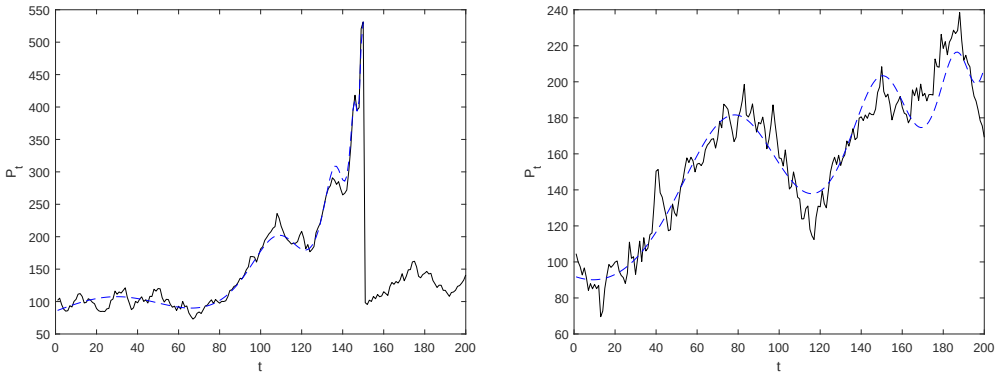


Figure 1: Construction of hierarchical diamond lattice.



(a) Solid line represents simulated price and the dashed line shows corresponding fit of the LPPL. Parameter estimates are $\hat{\beta} = 0.372$, $\hat{\omega} = 5.89$, $\hat{t}_c = 151$, $\hat{A} = 6.61$, $\hat{B} = -0.366$, $\hat{C}_1 = -0.0425$, $\hat{C}_2 = -0.0184$ **(b)** Solid line represents simulated price and the dashed line shows corresponding fit of the LPPL. Parameter estimates are $\hat{\beta} = 0.854$, $\hat{\omega} = 9.10$, $\hat{t}_c = 223$, $\hat{A} = 5.46$, $\hat{B} = -0.0068$, $\hat{C}_1 = 0.0006$, $\hat{C}_2 = 0.0031$

Figure 2: Fitting the log-periodic-power-law to simulated prices.

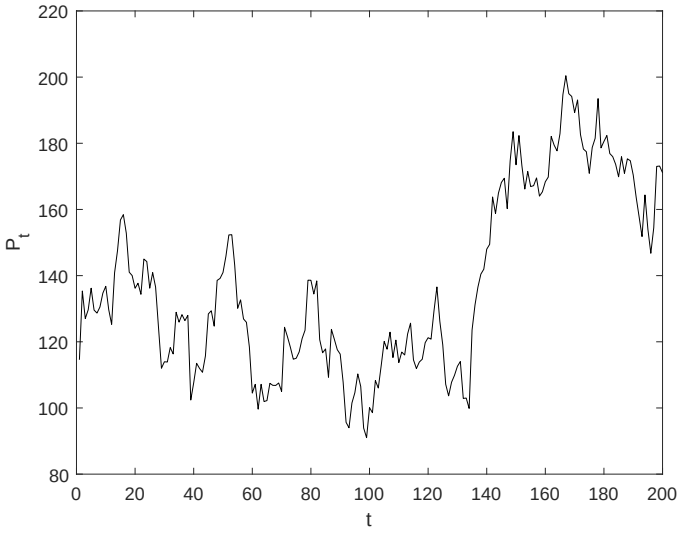


Figure 3: Random walk

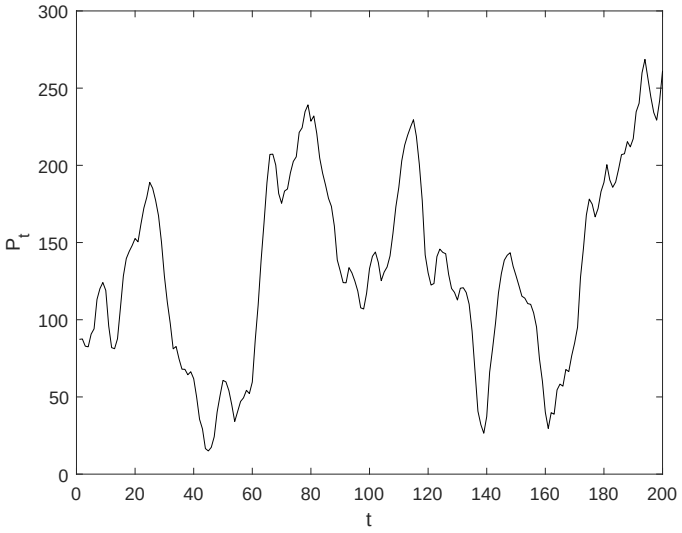


Figure 4: Random walk with MA(3) innovations

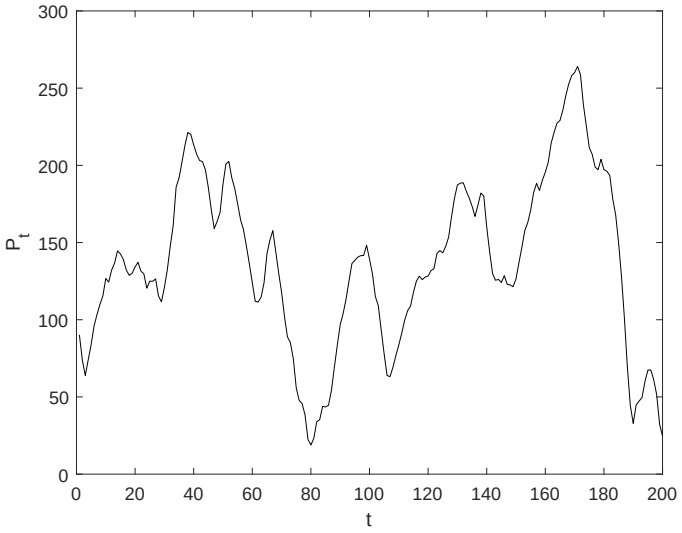


Figure 5: Random walk with AR(1) innovations

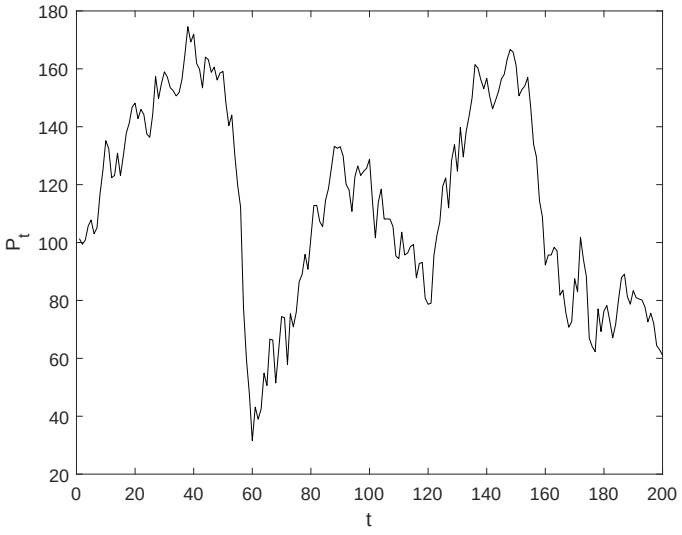


Figure 6: Random walk with GARCH innovations

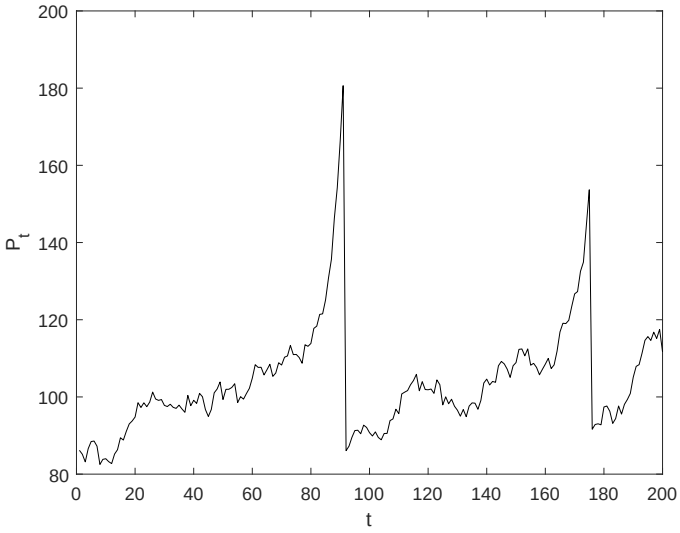


Figure 7: Partially collapsing bubble

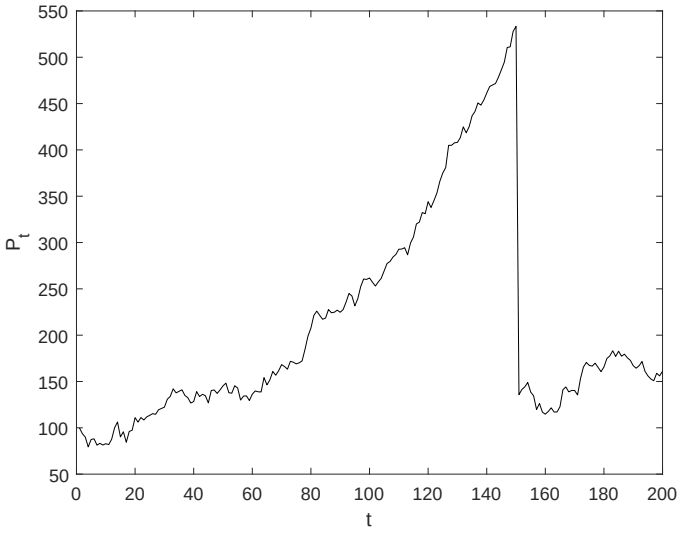


Figure 8: Random walk with one bubble

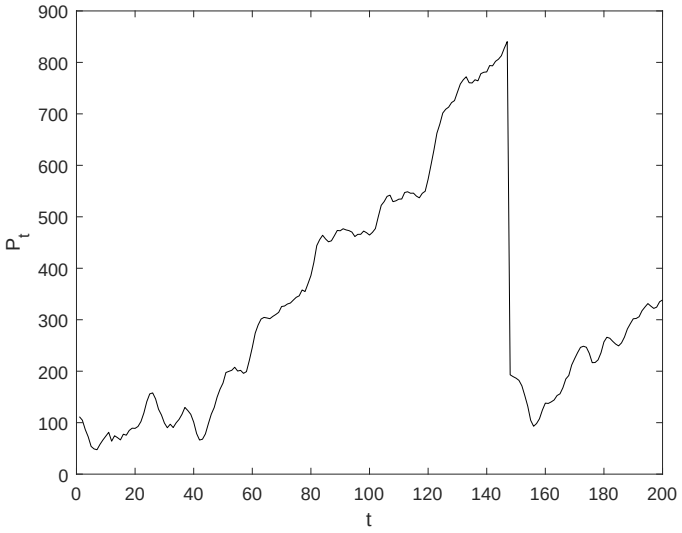


Figure 9: RW with one bubble and MA(3) innovations

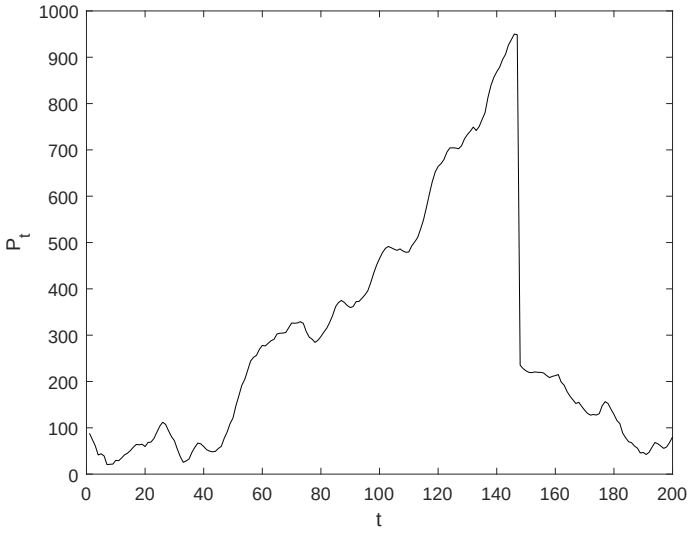


Figure 10: RW with one bubble and AR(1) innovations

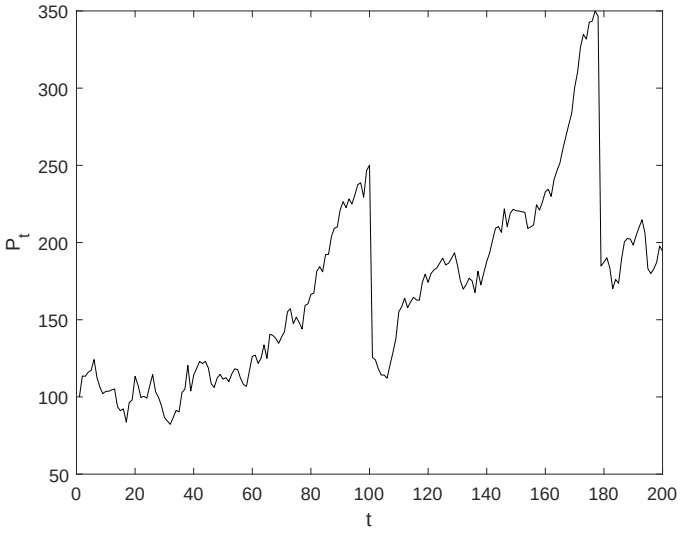


Figure 11: Random walk with two bubbles

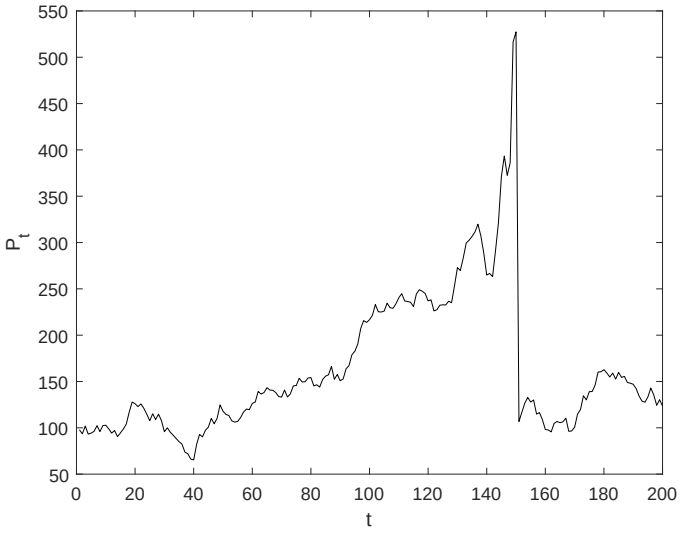


Figure 12: Bubble following LPPL

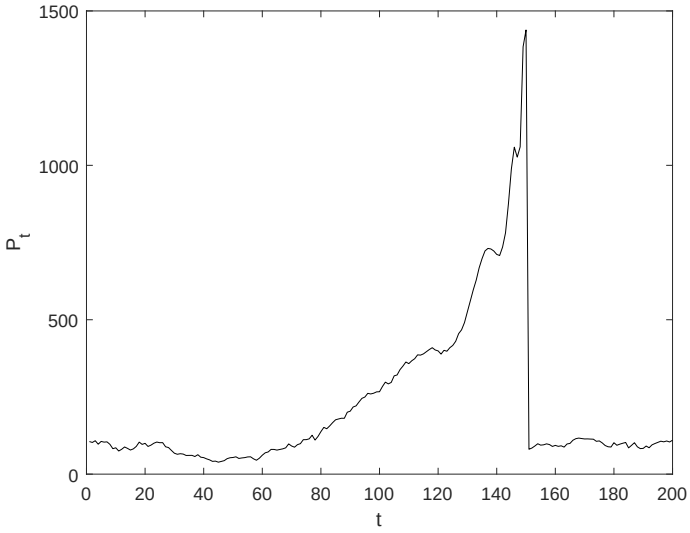


Figure 13: Bubble following LPPL with positive required return

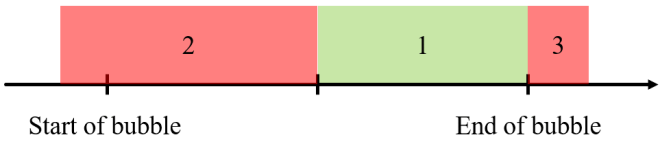
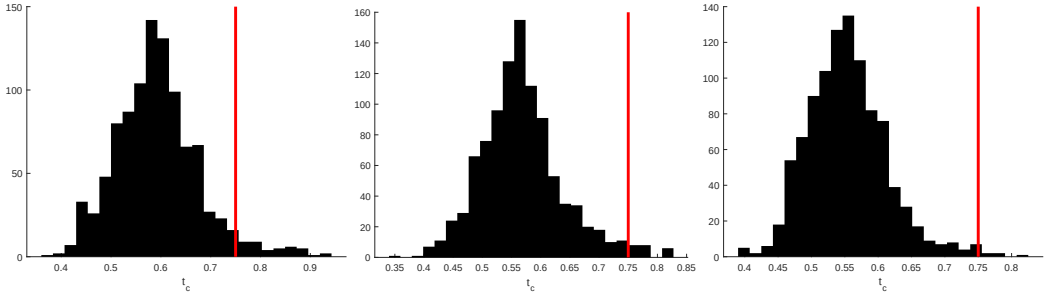
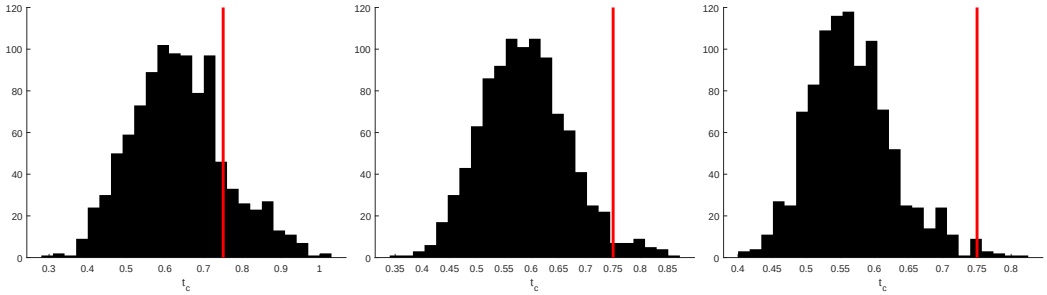


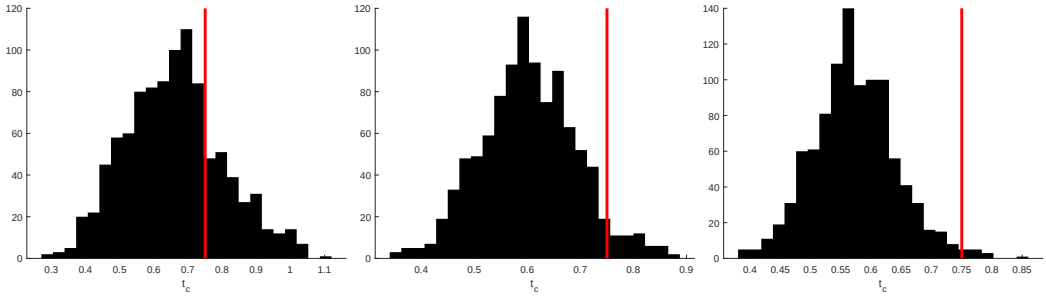
Figure 14: Graphical illustration of the three conditions for successful time stamping.



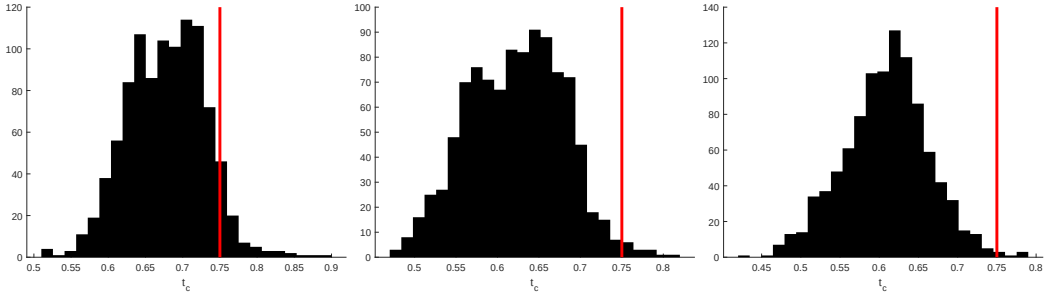
(a) Process [6] with $T = 100$. **(b)** Process [6] with $T = 200$. **(c)** Process [6] with $T = 400$.



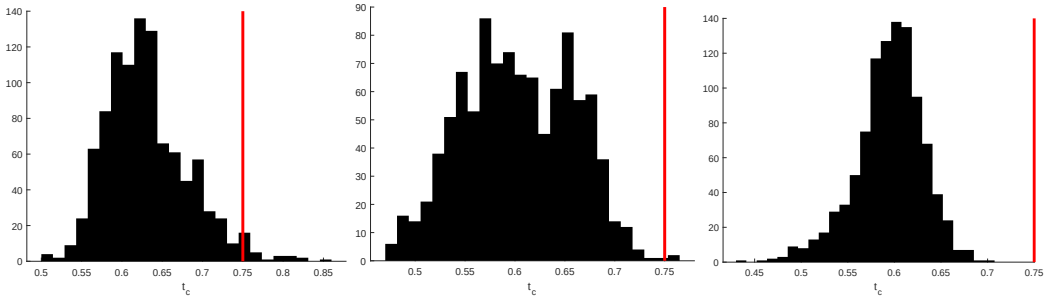
(d) Process [7] with $T = 100$. **(e)** Process [7] with $T = 200$. **(f)** Process [7] with $T = 400$.



(g) Process [8] with $T = 100$. **(h)** Process [8] with $T = 200$. **(i)** Process [8] with $T = 400$.

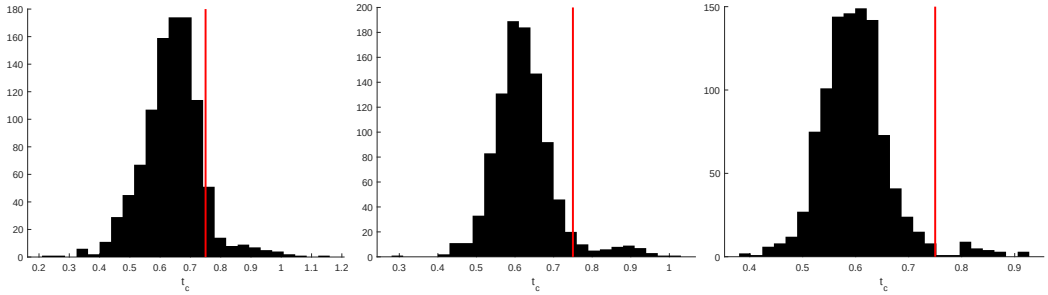


(j) Process [9] with $T = 100$. **(k)** Process [9] with $T = 200$. **(l)** Process [9] with $T = 400$.

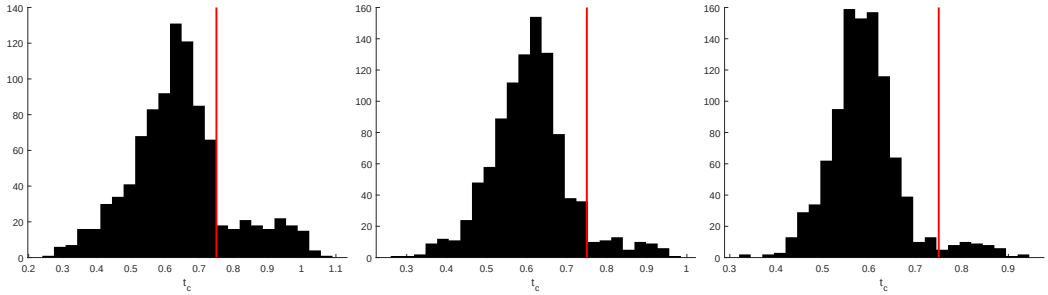


(m) Process [10] with $T = 100$. **(n)** Process [10] with $T = 200$. **(o)** Process [10] with $T = 400$.

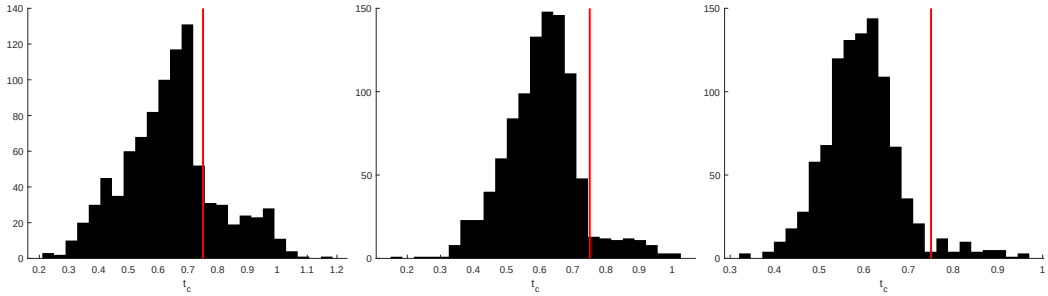
Figure 15: Histogram of estimates of the critical point t_c for the LPPPL procedure with the vertical line indicating the true value.



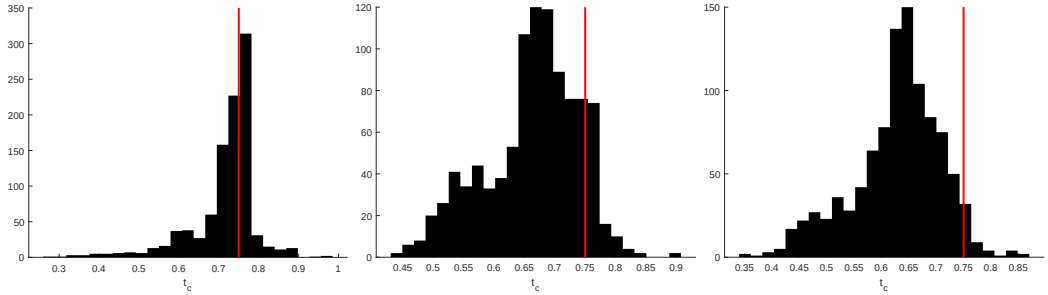
(a) Process [6] with $T = 100$. **(b)** Process [6] with $T = 200$. **(c)** Process [6] with $T = 400$.



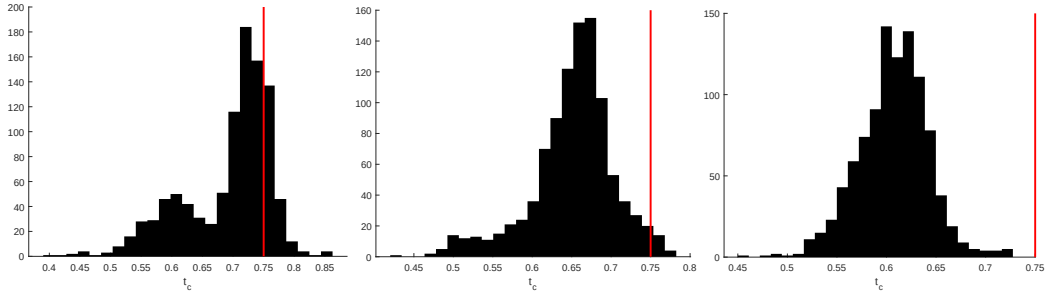
(d) Process [7] with $T = 100$. **(e)** Process [7] with $T = 200$. **(f)** Process [7] with $T = 400$.



(g) Process [8] with $T = 100$. **(h)** Process [8] with $T = 200$. **(i)** Process [8] with $T = 400$.



(j) Process [9] with $T = 100$. **(k)** Process [9] with $T = 200$. **(l)** Process [9] with $T = 400$.



(m) Process [10] with $T = 100$. **(n)** Process [10] with $T = 200$. **(o)** Process [10] with $T = 400$.

Figure 16: Histogram of estimates of the critical point t_c for the GLPPL procedure with the vertical line indicating the true value.

Figure 17

A Algorithm for Robust Critical Values

Following Pedersen and Schütte (2017) the algorithm for calculating the robust critical values using the sieve bootstrap is as follows:

Step 1. Obtain the $\hat{\phi}_{i,T}$ estimates and residuals $\hat{\epsilon}_T$ from the ADF regression

$$\hat{\epsilon}_{t,T} = \Delta y_t - \hat{\alpha} - \hat{\beta} y_{t-1} - \sum_{i=1}^{k} \hat{\phi}_{i,T} \Delta y_{t-i}, \qquad t = k+1, ...T.$$

Step 2. Draw bootstrap errors, ϵ_t^*, randomly with replacement from

$$\hat{\epsilon}_{t,T} - (T-k)^{-1} \sum_{t=1+k}^{T} \hat{\epsilon}_{t,T}$$

Step 3. Construct u_t^* from ϵ_t^* by

$$u_t^* = \sum_{i=1}^{k} \hat{\phi}_{i,T} u_{t-i}^* + \epsilon_t^*$$

Step 4. Construct y_t^* as

$$y_t^* = y_{t-1}^* + u_t^*, \qquad t = 1, ..., T, \quad y_0^* = 0.$$

Step 5. Calculate the $GSADF(r_0)$ or $BSADF_{r_2}(r_0)$ test statistics as in section 2.

Step 6. Repeat steps 2 to 5 M^* times and obtain the q-quantile of the ordered $GSADF(r_0)$ or $BSADF_{r_2}(r_0)$ test statistics.

Step 7. Calculate the actual $GSADF^B(r_0)$ or $BSADF^B_{r_2}(r_0)$ test statistics using the orig-

inal time series.

$$GSADF^{B}(r_0) = \sup_{r_2 \in [r_0,1] \wedge r_1 \in [0, r_2 - r_0]} \hat{\beta}_{r_1}^{r_2}$$

$$BSADF_{r_2}^{B}(r_0) = \sup_{r_1 \in [0, r_2 - r_0]} \hat{\beta}_{r_1}^{r_2}.$$

Reject the null if the test statistic exceeds the bootstrap critical value.

As for the GSADF the lag length k in Step 1 is chosen by BIC over a set of potential $k \in \left[1, \ldots, \lfloor 8\,(T/100)^{1/4} \rfloor\right]$.

B Proof of Theorem 3.1

Proof. We assert that when the hazard rate is given by

$$h_t = B'\,(t_c - t)^{\beta-1} + C'\,(t_c - t)^{\beta-1} \cos\left(\omega \log\left(t_c - t\right) + \phi'\right) \tag{22}$$

Then the indefinite integral of the hazard rate is given by

$$\begin{aligned}\int h_t \mathrm{d}t = -\frac{B'}{\beta}\,(t_c - t)^{\beta} - \frac{C'}{\omega^2 + \beta^2}\,(t_c - t)^{\beta}\,(\omega \sin\left(\omega \log\left(t_c - t\right) + \phi'\right) \\ + \beta \cos\left(\omega \log\left(t_c - t\right) + \phi'\right)) + W,\end{aligned} \tag{23}$$

where W is the constant of integration.

Proof. Differentiate (23) with respect to t.

$$
\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\left[-\frac{C'}{\omega^2+\beta^2}(t_c-t)^\beta\left(\omega\sin\left(\omega\log(t_c-t)+\phi'\right)+\beta\cos\left(\omega\log(t_c-t)+\phi'\right)\right)\right.\\
&\left.W-\frac{B'}{\beta}(t_c-t)^\beta\right]\\
=&\frac{\beta C'}{\beta^2+\omega^2}(t_c-t)^{\beta-1}\left(\omega\sin\left(\omega\log(t_c-t)+\phi'\right)+\beta\cos\left(\omega\log(t_c-t)+\phi'\right)\right)\\
&-\frac{C'}{\beta^2+\omega^2}(t_c-t)^\beta\left(\beta\sin\left(\omega\log(t_c-t)+\phi'\right)-\omega\cos\left(\omega\log(t_c-t)+\phi'\right)\right)\frac{\omega}{t_c-t}\\
&+B'(t_c-t)^{\beta-1}\\
=&\frac{C'}{\beta^2+\omega^2}(t_c-t)^{\beta-1}\left(\beta\omega\sin\left(\omega\log(t_c-t)+\phi'\right)+\beta^2\cos\left(\omega\log(t_c-t)+\phi'\right)\right)\\
&+B'(t_c-t)^{\beta-1}\\
&-\frac{C'}{\beta^2+\omega^2}(t_c-t)^{\beta-1}\left(\beta\omega\sin\left(\omega\log(t_c-t)+\phi'\right)-\omega^2\cos\left(\omega\log(t_c-t)+\phi'\right)\right)\\
=&B'(t_c-t)^{\beta-1}+\frac{C'}{\beta^2+\omega^2}(t_c-t)^{\beta-1}\cos\left(\omega\log(t_c-t)+\phi'\right)\left(\beta^2+\omega^2\right)\\
=&B'(t_c-t)^{\beta-1}+C'(t_c-t)^{\beta-1}\cos\left(\omega\log(t_c-t)+\phi'\right),
\end{aligned}
$$

which is exactly the hazard rate in (22). □

Then the price can be found by substituting (23) into (7).

$$
\begin{aligned}
\log P_t =&\kappa\int_{t_0}^{t}h_{t'}\,\mathrm{d}t'\\
=&\kappa\left[-\frac{C'}{\omega^2+\beta^2}(t_c-t')^\beta\left(\omega\sin\left(\omega\log(t_c-t')+\phi'\right)+\beta\cos\left(\omega\log(t_c-t')+\phi'\right)\right)\right.\\
&\left.-\frac{B'}{\beta}(t_c-t')^\beta\right]_{t_0}^{t}+\log P_{t_0},
\end{aligned}
$$

and expanding

$$\begin{aligned}\log P_t = &\log P_{t_0} - \kappa\frac{B'}{\beta}\left(t_c - t\right)^\beta + \kappa\frac{B'}{\beta}\left(t_c - t_0\right)^\beta \\ &+ \frac{\kappa C'}{\omega^2+\beta^2}\left(t_c - t_0\right)^\beta\left(\omega\sin\left(\omega\log\left(t_c - t_0\right)+\phi'\right) + \beta\cos\left(\omega\log\left(t_c - t_0\right)+\phi'\right)\right) \\ &- \frac{\kappa C'}{\omega^2+\beta^2}\left(t_c - t\right)^\beta\left(\omega\sin\left(\omega\log\left(t_c - t\right)+\phi'\right) + \beta\cos\left(\omega\log\left(t_c - t\right)+\phi'\right)\right). \end{aligned} \tag{24}$$

Then realizing that

$$\begin{aligned}\log P_{t_c} = &\frac{\kappa C'}{\omega^2+\beta^2}\left(t_c - t_0\right)^\beta\left(\omega\sin\left(\omega\log\left(t_c - t_0\right)+\phi'\right) + \beta\cos\left(\omega\log\left(t_c - t_0\right)+\phi'\right)\right) \\ &+ \log P_{t_0} + \kappa\frac{B'}{\beta}\left(t_c - t_0\right)^\beta,\end{aligned}$$

we can write $\log P_t$ as

$$\begin{aligned}\log P_t = &\log P_{t_c} - \kappa\frac{B'}{\beta}\left(t_c - t\right)^\beta \\ &- \frac{\kappa C'}{\omega^2+\beta^2}\left(t_c - t\right)^\beta\left(\omega\sin\left(\omega\log\left(t_c - t\right)+\phi'\right) + \beta\cos\left(\omega\log\left(t_c - t\right)+\phi'\right)\right),\end{aligned}$$

which due to the Harmonic Addition Theorem can be reduced to

$$\begin{aligned}\log P_t &= \log P_{t_c} - \kappa\frac{B'}{\beta}\left(t_c - t\right)^\beta - \frac{\kappa C'}{\omega^2+\beta^2}\left(t_c - t\right)^\beta\cos\left(\omega\log\left(t_c - t\right)+\phi\right)\sqrt{\beta^2+\omega^2} \\ &= \log P_{t_c} + B\left(t_c - t\right)^\beta + C\left(t_c - t\right)^\beta\cos\left(\omega\log\left(t_c - t\right)+\phi\right),\end{aligned}$$

where $B = -\kappa\frac{B'}{\beta}$ and $C = -\frac{\kappa C'}{\sqrt{\omega^2+\beta^2}}$. Then this is exactly equal to the process in (11). □

C Proof of Theorem 3.2

Proof. Rewrite (11) such that

$$\log P_t = \log P_{t_c} + B\left(t_c - t\right)^{\beta} + C\left(t_c - t\right)^{\beta} \cos\left(\omega \log\left(t_c - t\right)\right) \cos\phi - C\left(t_c - t\right)^{\beta} \sin\left(\omega \log\left(t_c - t\right)\right) \sin\phi, \tag{25}$$

Insert $C_1 = C\cos\phi$ and $C_2 = -C\sin\phi$. □

D Proof of Corollary 3.2.1

Proof. Given the hazard rate in (10) then

$$B'\left(t_c - t\right)^{\beta-1} + C'\left(t_c - t\right)^{\beta-1} \cos\left(\omega \log\left(t_c - t\right) + \psi\right) > 0$$

$$B' > C'.$$

Then since $B' = -\frac{B\beta}{\kappa}$ and $C' = -\frac{C\sqrt{\omega^2+\beta^2}}{\kappa}$ (See appendix B) and since $C = \sqrt{C_1^2 + C_2^2}$ then

$$-\frac{B\beta}{\kappa} > -\frac{C\sqrt{\omega^2+\beta^2}}{\kappa}$$

$$B < \sqrt{\left(C_1^2 + C_2^2\right)\frac{\omega^2+\beta^2}{\beta^2}}.$$

□

E Proof of Corollary 3.2.2

Proof. The oscillating part is bounded and alternates between being increasing and decreasing. Hence, it does not make sense to set restrictions for the entire hazard rate to be increasing at all points for $t \to t_c$. Rather we require that the non-oscillating part of the hazard rate is increasing for $0 < t < t_c$. Only differentiating the non-oscillating

part with respect to t yields

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} B' (t_c - t)^{\beta-1} &> 0 \\ -B' (\beta - 1) (t_c - t)^{\beta-2} &> 0 \\ \frac{1}{\kappa} B\beta (\beta - 1) &> 0 \\ B\beta (\beta - 1) &> 0 \end{aligned} \tag{26}$$

From the first to the second line we differentiate the non-oscillating part with respect to t. When moving to the third line we insert $B' = -\frac{B\beta}{\kappa}$ and realize that $(t_c - t)^{\beta-2} > 0$ for $0 < t < t_c$. When moving to the fourth line we use that $\kappa > 0$. (26) is then a sufficient condition for the hazard rate to be increasing (disregarding the oscillating part) as $t \to t_c$. □

F Proof of Theorem 3.3 and Corollary 3.3.1

Proof. Rather than the martingale condition in (6), assuming a required rate of return of zero, we consider a more general required rate of return of γ_t such that the martingale condition reads

$$\mathbb{E}_t (\mathrm{d}P_t) = \mu_t P_t \mathrm{d}t - \kappa P_t \mathbb{E}_t (\mathrm{d}j) = \gamma_t,$$

which reduces to

$$\mu_t = \kappa h_t + \gamma_t.$$

Inserting this into (5) assuming that no crash has materialized, $\mathrm{d}j = 0$ yields

$$\mathrm{d}P_t = (\kappa h_t + \gamma_t) P_t \mathrm{d}t,$$

which is an ordinary differential equation with the solution

$$\begin{aligned}\log P_t &= \log P_{t_0} + \int_{t_0}^{t} \kappa h_{t'} + \gamma_{t'} \mathrm{d}t' \\ &= \log P_{t_0} + \kappa \int_{t_0}^{t} h_{t'} \mathrm{d}t' + \int_{t_0}^{t} \gamma_{t'} \mathrm{d}t'.\end{aligned}$$

Then using the result in (24) while splitting the integral of γ at t_c.

$$\begin{aligned}\log P_t = &\log P_{t_0} - \kappa \frac{B'}{\beta} (t_c - t)^{\beta} + \kappa \frac{B'}{\beta} (t_c - t_0)^{\beta} \\ &- \frac{\kappa C'}{\omega^2 + \beta^2} (t_c - t)^{\beta} \left(\omega \sin\left(\omega \log (t_c - t) + \phi'\right) + \beta \cos\left(\omega \log (t_c - t) + \phi'\right)\right) \\ &+ \frac{\kappa C'}{\omega^2 + \beta^2} (t_c - t_0)^{\beta} \left(\omega \sin\left(\omega \log (t_c - t_0) + \phi'\right) + \beta \cos\left(\omega \log (t_c - t_0) + \phi'\right)\right) \\ &+ \int_{t_0}^{t_c} \gamma_{t'} \mathrm{d}t' - \int_{t}^{t_c} \gamma_{t'} \mathrm{d}t'.\end{aligned}$$

Then a time t_c the price is given by

$$\begin{aligned}\log P_{t_c} = &\log P_{t_0} + \kappa \frac{B'}{\beta} (t_c - t_0)^{\beta} + \int_{t_0}^{t_c} \gamma'_t \mathrm{d}t' \\ &+ \frac{\kappa C'}{\omega^2 + \beta^2} (t_c - t_0)^{\beta} \left(\omega \sin\left(\omega \log (t_c - t_0) + \phi'\right) + \beta \cos\left(\omega \log (t_c - t_0) + \phi'\right)\right),\end{aligned}$$

such that the process reduces to

$$\begin{aligned}\log P_t = &\log P_{t_c} - \kappa \frac{B'}{\beta} (t_c - t)^{\beta} - \int_{t}^{t_c} \gamma_{t'} \mathrm{d}t' \\ &- \frac{\kappa C'}{\omega^2 + \beta^2} (t_c - t)^{\beta} \left(\omega \sin\left(\omega \log (t_c - t) + \phi'\right) + \beta \cos\left(\omega \log (t_c - t) + \phi'\right)\right).\end{aligned}$$

And finally by the Harmonic Addition Theorem

$$\log P_t = \log P_{t_c} + B (t_c - t)^{\beta} + C (t_c - t)^{\beta} \cos\left(\omega \log (t_c - t) + \phi\right) - \int_{t}^{t_c} \gamma_{t'} \mathrm{d}t',$$

where $B = -\kappa \frac{B'}{\beta}$ and $C = -\frac{\kappa C'}{\sqrt{\omega^2 + \beta^2}}$.

If we assume a constant but positive required rate of return such that $\gamma_t = \gamma$, then

we can write

$$\log P_t = \log P_{t_c} + B\,(t_c - t)^{\beta} + C\,(t_c - t)^{\beta} \cos\left(\omega \log\left(t_c - t\right) + \phi\right) - \gamma\,(t_c - t)\,,$$

□

G Parameter Restrictions in the LPPL Literature

There are many different versions of the parameter restrictions for the LPPL procedure. To name a few Johansen et al. (1999) find that $\beta \in (0.19, 0.6)$ and $\omega \in (5.2, 8.2)$, Johansen et al. (2000) find that $\beta \approx 0.6$ and $\omega \in (5, 10)$, Feigenbaum (2001) finds that $\beta \in (0.42, 0.74)$ and $\omega \in (6.0, 9.6)$, Graf and Meister (2003) find $\beta \in (0, 1)$ and $\omega \in (7, 13)$, Filimonov and Sornette (2013) find that $\beta \in (0.1, 0.9)$ and $\omega \in (6, 13)$, and Wosnitza and Leker (2014) find that $\beta \in (0, 1)$ and $\omega \in (2, 4)$. These values show quite some variation, and it leaves quite a bit of freedom in choosing the parameter restrictions. We have made a choice on the restrictions that we consider in line with the existing literature, but in trying alternative restrictions we found little changes in the results regarding bubble detection. Choosing tighter bounds for β and ω only decreased the share of simulations in which bubbles were detected from almost unity to not much less than 0.95. The broader bands that has been used for this study capture more of the findings from the current literature and is found to improve time stamping performance for the LPPL test.

Research Papers 2019

2018-37: Niels S. Grønborg, Asger Lunde, Kasper V. Olesen and Harry Vander Elst: Realizing Correlations Across Asset Classes

2018-38: Riccardo Borghi, Eric Hillebrand, Jakob Mikkelsen and Giovanni Urga: The dynamics of factor loadings in the cross-section of returns

2019-01: Andrea Gatto and Francesco Busato: Defining, measuring and ranking energy vulnerability

2019-02: Federico Carlini and Paolo Santucci de Magistris: Resuscitating the co-fractional model of Granger (1986)

2019-03: Martin M. Andreasen and Mads Dang: Estimating the Price Markup in the New Keynesian Model

2019-04: Daniel Borup, Bent Jesper Christensen and Yunus Emre Ergemen: Assessing predictive accuracy in panel data models with long-range dependence

2019-05: Antoine A. Djogbenou, James G. MacKinnon and Morten Ørregaard Nielsen: Asymptotic Theory and Wild Bootstrap Inference with Clustered Errors

2019-06: Vanessa Berenguer-Rico, Søren Johansen and Bent Nielsen: The analysis of marked and weighted empirical processes of estimated residuals

2019-07: Søren Kjærgaard, Yunus Emre Ergemen, Kallestrup-Lamb, Jim Oeppen and Rune Lindahl-Jacobsen: Forecasting Causes of Death using Compositional Data Analysis: the Case of Cancer Deaths

2019-08: Søren Kjærgaard, Yunus Emre Ergemen, Marie-Pier Bergeron Boucher, Jim Oeppen and Malene Kallestrup-Lamb: Longevity forecasting by socio-economic groups using compositional data analysis

2019-09: Debopam Bhattacharya, Pascaline Dupas and Shin Kanaya: Demand and Welfare Analysis in Discrete Choice Models with Social Interactions

2019-10: Martin Møller Andreasen, Kasper Jørgensen and Andrew Meldrum: Bond Risk Premiums at the Zero Lower Bound

2019-11: Martin Møller Andrasen: Explaining Bond Return Predictability in an Estimated New Keynesian Model

2019-12: Vanessa Berenguer-Rico, Søren Johansen and Bent Nielsen: Uniform Consistency of Marked and Weighted Empirical Distributions of Residuals

2019-13: Daniel Borup and Erik Christian Montes Schütte: In search of a job: Forecasting employment growth using Google Trends

2019-14 Kim Christensen, Charlotte Christiansen and Anders M. Posselt: The Economic Value of VIX ETPs

2019-15 Vanessa Berenguer-Rico, Søren Johansen and Bent Nielsen: Models where the Least Trimmed Squares and Least Median of Squares estimators are maximum likelihood

2019-16 Kristoffer Pons Bertelsen: Comparing Tests for Identification of Bubbles