The European Journal of Finance, 2013
Vol. 19, No. 5, 366–391, http://dx.doi.org/10.1080/1351847X.2011.601657

Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask

Petr Geraskin[a]* and Dean Fantazzini[b]

[a]*Higher School of Economics, Faculty of Economics, Moscow, Russia;* [b]*Moscow School of Economics, Moscow State University, Moscow, Russia*

Sornette, Johansen, and Bouchaud (1996), Sornette and Johansen (1997), Johansen, Ledoit, and Sornette (2000) and Sornette (2003a) proposed that, prior to crashes, the mean function of a stock index price time series is characterized by a power law decorated with log-periodic oscillations, leading to a critical point that describes the beginning of the market crash. This article reviews the original log-periodic power law model for financial bubble modeling and discusses early criticism and recent generalizations proposed to answer these remarks. We show how to fit these models with alternative methodologies, together with diagnostic tests and graphical tools, to diagnose financial bubbles in the making in real time. An application of this methodology to the gold bubble which burst in December 2009 is then presented.

Keywords: log-periodic models; LPPL; crash; bubble; anti-bubble; GARCH; forecasting; gold; Google trends

JEL classification: C32, C51, C53, G17

1. Introduction

Detecting a financial bubble and predicting when it will end have become of crucial importance, given the series of financial bubbles that led to the current ‘Second Great Contraction’, using the definition given by Reinhart and Rogoff (2009). As noted by Sornette (2009), Sornette and Woodard (2010), Kaizoji and Sornette (in press), Sornette, Woodard, and Zhou (2009) and Fantazzini (2010a, 2010b), the global financial crisis that had started in 2007 can be considered an example of how the bursting of a bubble can be dealt with by creating new bubbles. This consideration, which is not new in the financial literature (see e.g. Sornette and Woodard 2010 and references therein), was indirectly confirmed by Lou Jiwei, the Chairman of the $298 billion sovereign wealth fund named China Investment Corporation, which was created in 2007 with the goal to manage an important part of the People’s Republic of China’s foreign exchange reserves. On 28 August 2009, Lou told reporters on the sidelines of a forum organized by the Washington-based Brookings Institution and the Chinese ‘Economists 50 Forum’, a Beijing think-tank, that ‘both China and America are addressing bubbles by creating more bubbles and we’re just taking advantage of that. So we can’t lose.’ Moreover, Lou also added that ‘CIC was building a broad investment portfolio that includes products designed to generate both alpha and beta; to hedge against both inflation and deflation; and to provide guaranteed returns in the event of a new crisis’ (see Xin and Zhou Wheatley 2009 for more details). The

*Corresponding author. Email: dean.fantazzini@gmail.com

previous comments clearly point out how important it is to have tools able to detect bubbles in the making.

Unfortunately, there is no consensus in the economic literature on what a bubble is: Gürkaynak (2008) surveyed a large set of econometric tests of asset price bubbles and found that for each paper that finds evidence of bubbles, there is another one that fits the data equally well without allowing for a bubble, so that it is not possible to distinguish bubbles from time-varying fundamentals. A similar situation can also be found in the professional literature: for example, Alan Greenspan stated on 30 August 2002 that '...We, at the Federal Reserve ... recognized that, despite our suspicions, it was very difficult to definitively identify a bubble until after the fact, that is, when its bursting confirmed its existence.' (Greenspan 2002) So, is this a lost cause? Absolutely not.

A model which has quickly gained a lot of attention among financial practitioners and in the physics academic literature due to the many successful predictions is the so-called *log-periodic power law* (LPPL) approach proposed by Sornette, Johansen, and Bouchaud (1996), Sornette and Johansen (1997), Johansen, Ledoit, and Sornette (2000) and Sornette (2003a, 2003b). The Johansen–Ledoit–Sornette (JLS) model assumes the presence of two types of agents in the market: a group of traders with rational expectations and a second group of so-called noise traders, that is, irrational agents with herding behavior. The idea of the JLS model comes from statistical physics and it shares many elements with a model introduced by Ising for explaining ferromagnetism (see e.g. Goldenfeld 1992). According to this model, traders are organized into networks and can have only two states: buy or sell. In addition, their trading actions depend on the decisions of other traders and on external influences. Due to these interactions, agents can form groups with self-similar behavior which can lead the market to a bubble situation, which can be considered a situation of 'order' compared with the 'disorder' of normal market conditions. Another important feature introduced in this model is the positive feedbacks which are generated by the increasing risk and the agents' interactions, so that a bubble can be a self-sustained process.

Many examples of calibrations of financial bubbles with LPPLs were reported by Sornette (2003a), who suggested that the LPPL model provides a good starting point to detect bubbles and forecast their most probable end. Johansen and Sornette (2004) identified the most extreme cumulative losses (i.e. drawdowns) in a large set of financial assets and showed that they belong to a probability density distribution, which is distinct from the distribution of the 99% of the smaller drawdowns which represent the normal market regime. Moreover, they showed that for two-thirds of these extreme drawdowns, the market prices followed a super-exponential behavior prior to their occurrences, as confirmed by a calibration of an LPPL model. These particular drawdowns (or outliers) are called 'dragon kings' in Sornette (2009). Interestingly, this approach allowed to diagnose bubbles *ex ante*, as shown in a series of real-life tests (see Zhou and Sornette 2003, 2006, 2008, 2009, Sornette and Zhou 2006, Sornette, Woodard, and Zhou 2008). Furthermore, it is currently being used at the Financial Crisis Observatory, which is a scientific platform set up at the ETH – Zurich, aimed at 'testing and quantifying rigorously the hypothesis that financial markets exhibit a degree of inefficiency and a potential for predictability, especially during regimes when bubbles develop', (Financial Crisis Observatory website).

The goal of this article is to present an easy-to-use and self-contained guide for bubble modeling and detecting with LPPLs, which contains all the sufficient steps to derive the main models in this growing and interesting field of the literature, and discuss the important aspects for practitioners and researchers.

The rest of the paper is organized as follows. Section 2 reviews the original JLS model with the main steps required for its derivation. Section 3 discusses the early criticism to this approach and recent generalizations proposed to answer these remarks. Section 4 discusses how to fit LPPL models, by presenting three estimation methodologies: the original two-step nonlinear optimization proposed by Johansen, Ledoit, and Sornette (2000), the genetic algorithm (GA) approach proposed by Jacobsson (2009) and the two-step/three-step maximum-likelihood (ML) approach proposed by Fantazzini (2010a). Section 5 is devoted to the diagnosis of bubbles in the making by using a set of different techniques. We describe diagnostic tests based on the LPPL fitting residuals and diagnostic tests based on rational expectation models with stochastic mean-reverting termination times, as well as graphical tools useful for capturing bubble development and for understanding whether a crash is in sight or not. Section 6 presents a detailed empirical application devoted to the burst of the gold bubble in December 2009, while Section 7 briefly concludes.

2. The original LPPL model

Johansen, Ledoit, and Sornette (2000) considered an ideal market with no dividends and where interest rates, risk aversion and market liquidity constraints are ignored. Therefore, the fundamental value for an asset is $p(t) = 0$, so any positive value of $p(t)$ represents a bubble. In general, $p(t)$ can be viewed as the price in excess of the fundamental value of an asset. In this framework, there are two types of agents: first, a group of rational agents who are identical in their preferences and characteristics, and so they can be substituted with a single representative agent, and second, a group of irrational agents whose herding behavior leads to the development of a financial bubble. When this tendency develops till a certain critical value, a large proportion of agents will then assume the same short position, thus causing a crash. A financial crash is not a certain event in this model, but it is characterized by a probability distribution: as a consequence, it is rational for financial agents to continue investing, because the risk for the crash to happen is compensated by the positive return generated by the financial bubble and there exists a small probability for the bubble to disappear smoothly, without the occurrence of a crash.

The key variable to model the price behavior before a crash is the crash hazard rate $h(t)$, that is, the probability per unit of time that the crash will take place, given that it has not yet occurred. The hazard rate $h(t)$ quantifies the probability that a great number of agents will assume the same sell position simultaneously, a position that the market will not be able to satisfy unless the prices decrease substantially. We remark that a strong collective answer (as it is the case for a crash) is not necessarily the consequence of one elaborated internal mechanism of global coordination in this model, but it can appear starting from imitative local micro-interactions, which are then transmitted by the market resulting in a macroscopic effect. In this regard, Johansen, Ledoit, and Sornette (2000) first discussed a macroscopic 'mean field' approach and then turned to a more microscopic approach.

2.1 *Macroscopic modeling*

According to the mean field theory from Statistical Mechanics (Stanley 1971, Goldenfeld 1992), a simple way for describing an imitative process is to assume that the hazard rate $h(t)$ can be described by the following equation:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = Ch^{\delta}, \tag{1}$$

where $C > 0$ is a constant and $\delta > 1$ represents the average number of interactions among traders minus one. Thus, it follows that an amplification of interactions increases the hazard rate. If we integrate Equation (1), we have

$$h(t) = \left(\frac{h_0}{t_c - t} \right)^{\alpha}, \quad \alpha = \frac{1}{\delta - 1}, \tag{2}$$

where t_c is the critical time determined by the initial conditions at some origin of time. It can be shown that the condition $\delta > 1$ (and consequently $\alpha > 0$) is crucial to obtain a growth of $h(t)$ as $t \to t_c$ and, therefore, a critical point in finite time. Moreover, the condition that $\alpha < 1$ is required for the price not to diverge at t_c. Rewriting these condition for δ, we have that $2 < \delta < \infty$, that is, an agent should be connected at least with two agents.

Another important feature of this approach is the possibility of self-fulfilling crisis, which is a concept recently proposed to explain the recession in the 1990s in seven countries (Argentina, Indonesia, Hong Kong, Malaysia, Mexico, South Korea and Thailand) (see Krugman 1998, Sornette 2003a). It is suggested that the loss of investors' confidence caused a self-fulfilling process in these countries and thus led to severe recessions. This feedback process can be modeled by using the previous mean field approach:

$$\frac{\mathrm{d}h}{\mathrm{d}t} = Dp^{\mu}, \quad \mu > 0, \tag{3}$$

where D is a positive constant. The underlying idea is that the lack of confidence quantified by the hazard rate increases when the market price departs from its fundamental value. Therefore, the price has to increase to compensate the increasing risk.

2.2 *Microscopic modeling*

Johansen, Ledoit, and Sornette (2000) and Sornette (2003a) assumed that the group of irrational agents are connected into a network. Each agent is indexed by a integer number $i = 1, \ldots, I$ and $N(i)$ represents the number of agents who are directly connected to agent i in the network. Johansen, Ledoit, and Sornette (2000) assumed that each agent can have only two possible states s_i: 'buy' ($s_i = +1$) or 'sell' ($s_i = -1$). Johansen, Ledoit, and Sornette (2000) supposed that the state of agent i is determined by the following Markov process:

$$s_i = \operatorname{sign}\left(K \sum_{k \in N(i)} s_j + \sigma \varepsilon_i \right), \tag{4}$$

where the sign function sign(x) is equal to +1 if $x > 0$ and to -1 if $x < 0$, K is a positive constant and ε_i is an i.i.d. standard normal random variable. In this model, K governs the tendency of imitation among traders, while σ governs their idiosyncratic behavior. If K increases, the order in the network increases as well, while the reverse is true when σ increases. If order wins, the agents will imitate their close neighbors and their imitation will spread all over the network, thus causing a crash.[1] More specifically and in analogy with the Ising model, there exists a critical point K_c that determines the separation between the different regimes: when $K < K_c$, the disorder reigns and the sensibility to a small global influence is low. When the imitation force K grows approaching K_c, a hierarchy of groups of agents acting collectively and with the same position is formed. As a consequence, the market becomes extremely sensitive to small global disturbances.

Finally, for a larger imitation force so that $K > K_c$, the tendency of imitation is so intense that there exists a strong predominance of one state/position among agents.

A physical quantity that represents the degree of a system sensitivity to an external perturbation (or general global influence) is the so-called *susceptibility* of the system. This quantity describes the probability that a large group of agents will have the same state, given the existent external influences in the network. Let us assume the existence of a term G which measures the global influence and add it to Equation (4):

$$s_i = \text{sign}\left(K \sum_{k \in N(i)} s_j + \sigma \varepsilon_i + G\right). \tag{5}$$

If we define the average state of the market as $M = (1/I) \sum_{i=1}^{I} s_i$, for $G = 0$, we have $E[M] = 0$ by symmetry. For $G > 0$, we have $M > 0$, while for $G < 0$, $M < 0$. Thus, it follows that $E[M] \times G \geq 0$. The susceptibility of the system is then defined as $\chi = \mathrm{d}E[M]/\mathrm{d}G|_{G=0}$. In general, the susceptibility has three possible interpretations: first, it measures the sensitivity of M to a small change in the global influence. Secondly, it is (a constant times) the variance of M around its zero expectation, caused by idiosyncratic shocks ε_i. Finally, if we consider two agents and we force one to be in a certain state, the impact that our intervention will have on the second agent will be proportional to the susceptibility.

2.3 ***Price dynamics and derivation of the JLS model***

As anticipated previously, the rational agent considered by Johansen, Ledoit, and Sornette (2000) is risk neutral and has rational expectations. Thus, the asset price $p(t)$ follows a martingale process, that is, $E_t[p(t')] = p(t)$, $\forall t' > t$, where $E_t[\cdot]$ represents the conditional expectation, given all information available up to time t. In the case of market equilibrium, the previous equality is a necessary condition for no arbitrage.

Considering that there exists a non-zero probability for the crash to happen, we can define a jump process j which is equal to zero before crash and one after the occurrence of the crash at time t_c. Since t_c is unknown, it is described by a stochastic variable with a probability density function $q(t)$, a cumulative distribution function $Q(t)$ and a hazard rate given by $h(t) = q(t)/[1 - Q(t)]$, which is the probability per unit of time of the crash taking place in the next instant, given that it has not yet occurred. Assuming for simplicity that the price falls during a crash by a fixed percentage $k \in (0, 1)$, the asset price dynamics is given by

$$\begin{aligned}
\mathrm{d}p &= \mu(t)p(t)\mathrm{d}t - kp(t)\mathrm{d}j \\
\Rightarrow E[\mathrm{d}p] &= \mu(t)p(t)\mathrm{d}t - kp(t)[P(\mathrm{d}j = 0) \times (\mathrm{d}j = 0) + P(\mathrm{d}j = 1) \times (\mathrm{d}j = 1)] \\
&= \mu(t)p(t)\mathrm{d}t - kp(t)[0 + h(t)\mathrm{d}t] = \mu(t)p(t)\mathrm{d}t - kp(t)h(t)\mathrm{d}t.
\end{aligned} \tag{6}$$

The no-arbitrage condition and rational expectations together imply that $E[\mathrm{d}p] = 0$, so that $\mu(t)p(t)\mathrm{d}t - kp(t)h(t)\mathrm{d}t = 0$, which yields $\mu(t) = kh(t)$. Substituting the last equality into Equation (6), we obtain the differential equation defining the price dynamics before the occurrence of the crash given by $\mathrm{d}(\ln p(t)) = kh(t)$, whose solution is

$$\ln\left[\frac{p(t)}{p(t_0)}\right] = \kappa \int_{t_0}^{t} h(t')\,\mathrm{d}t'. \tag{7}$$

The idea is that the higher the probability of the crash is, the faster the price should grow to compensate investors for the increased risk of a crash in the market (see also Blanchard 1979). At this point, Johansen, Ledoit, and Sornette (2000) employed the result that a system of variables close to a critical point can be described by a *power law* and the susceptibility of the system diverges as follows:

$$\chi \approx A(K_c - K)^{-\gamma}, \tag{8}$$

where A is a positive constant and $\gamma > 0$ is called the critical exponent of the susceptibility (equal to $7/4$ for the two-dimensional Ising model). Unfortunately, the two-dimensional Ising model considers only investors interconnected in an uniform way, while in real markets, some agents can be more connected than others. Modern financial markets are constituted by a collection of interacting investors, which differ substantially in size, going from the individual investors until the large pension funds. Furthermore, all investors in the world are organized inside a network (family, friends, work, etc.), within which they locally influence each other. A more appropriate representation for the current structure of financial markets is given by a *hierarchical diamond lattice,* which was used by Johansen, Ledoit, and Sornette (2000) to develop a model of rational imitation. This structure can be described as follows: first, consider two agents linked to each other, so that we have one link and two agents. Secondly, substitute this link with four new links forming a diamond: the two original agents are now situated in the two diametrically opposite vertices, whereas the two other vertices are occupied by two new traders. Thirdly, for each one of these four links, substitute them with four new links, forming a diamond in the same way. If we repeat this operation an arbitrary number of times, we will get a hierarchical diamond lattice. As a result, after n iterations, there will be $N = (2/3) * (2 + 4^n)$ agents and $L = 4^n$ links among them. For example, the last generated agents will have only two links and the initial agents will have 2^n neighbors, while the others will have an intermediate number of neighbors in between. A version of this model was solved by Derrida, De Seze, and Itzykson (1983). The basic properties are similar to those of the rational imitation model using the bi-dimensional network. The only crucial difference is that the critical exponent γ of the susceptibility in Equation (8) can be a complex number. Therefore, the general solution is given by

$$\begin{aligned} \chi &\approx \mathrm{Re}[A_0(K_c - K)^{-\gamma} + A_1(K_c - K)^{-\gamma + i\omega} + \cdots] \\ &\approx A'_0(K_c - K)^{-\gamma} + A'_1(K_c - K)^{-\gamma}\cos[\omega \ln(K_c - K) + \psi] + \cdots, \end{aligned} \tag{9}$$

where A_0, A_1 and ω are real numbers and $\mathrm{Re}[\cdot]$ represents the real part of a complex number. The power law in Equation (9) is now corrected by oscillations called 'log-periodic', because they are periodic in the logarithm of the variable $(K_c - K)$, and $\omega/2$ is their log frequency. These oscillations are accelerating since their frequency explodes as it reaches the critical time. Considering this mechanism, Johansen, Ledoit, and Sornette (2000) assumed that the crash hazard rates behave in a similar way to the susceptibility in the neighborhood of a critical point. Therefore, using Equation (9) and considering a hierarchical lattice for the financial market, the hazard rate has the following behavior:

$$h(t) \approx B_0(t_c - t)^{-\alpha} + B_1(t_c - t)^{-\alpha}\cos[\omega \ln(t_c - t) + \psi']. \tag{10}$$

This behavior of the hazard rate shows that the risk of a crash per unit of time, given that it has not yet occurred, increases drastically when the interactions among investors become sufficiently strong. However, this acceleration is interrupted and superimposed with an accelerating sequence of phases where the risk decreases, which is represented by the log-periodic oscillations. Applying Equation (10) to Equation (7), we get the following evolution for the asset price before a crash:

$$\ln[p(t)] \approx \ln[p(c)] - \frac{\kappa}{\beta}\{B_0(t_\mathrm{c} - t)^\beta + B_1(t_\mathrm{c} - t)^\beta \cos[\omega \ln(t_\mathrm{c} - t) + \phi]\}, \tag{11}$$

which can be rewritten in a more suitable form for fitting a financial time series as follows:

$$\ln[p(t)] \approx A + B(t_\mathrm{c} - t)^\beta\{1 + C\cos[\omega \ln(t_\mathrm{c} - t) + \phi]\}, \tag{12}$$

where $A > 0$ is the value of $[\ln p(t_\mathrm{c})]$ at the critical time, $B < 0$ is the increase in $[\ln p(t)]$ over the time unit before the crash if C were to be close to zero, $C \neq 0$ is the proportional magnitude of the oscillations around the exponential growth, $0 < \beta < 1$ should be positive to ensure a finite price at the critical time t_c of the bubble and quantifies the power law acceleration of prices, and ω is the frequency of the oscillations during the bubble, while $0 < \phi < 2\pi$ is a phase parameter. Expression (12), which is known as the LPPL, is the fundamental equation that describes the temporal growth of prices before a crash and it has been proposed in different forms in various papers (e.g. Sornette 2003a, Lin, Ren, and Sornette 2009 and references therein). We remark that A, B, C and ϕ are just units distributions of betas and omegas, as described in Sornette and Johansen (2001) and Johansen (2003), and do not carry any structural information.

3. Criticism and recent generalizations

3.1 *Criticism*

The most important and detailed criticism against the LPPL approach was put forward by Chang and Feigenbaum (2006), who tested the mechanism underlying the LPPL using Bayesian methods applied to the time series of returns (see also Laloux et al. 1999 for additional criticism and the reply given by Johansen 2002). By comparing marginal likelihoods, they showed that a null hypothesis model without log-periodical structure outperforms the JLS model. And if the JLS model was true, they found that parameter estimates obtained by curve fitting have a small posterior probability. As a consequence, they suggested to abandon the class of models in which the LPPL structure is revealed through the expected return trajectory. These problems are due to the fact that the JLS model considers a deterministic time-varying drift decorated by a non-stationary stochastic random walk component: the latter component has a variance which increases over time, so that the deterministic trajectory moves away from the observable price path and model estimation with prices is no more consistent. Therefore, Chang and Feigenbaum (2006) considered the time series of returns instead of prices and resorted to Bayesian methods to simplify the analysis of a complicated time-series model like the JLS model (see Bernardo and Smith 1994 or Koop 2003 for an introduction to Bayesian theory). The benchmark model in Chang and Feigenbaum (2006) is represented by the Black-Scholes model, whose logarithmic returns are given by

$$r_i \sim N(\mu(t_i - t_{i-1}), \sigma^2(t_i - t_{i-1})), \tag{13}$$

where $r_i = q_i - q_{i-1}$ and q_i is the log of the price. The drift μ is drawn from the prior distribution $N(\mu_r, \sigma_r)$, while the variance σ^2 is specified in terms of its inverse $\tau = 1/\sigma^2$, known as the precision, which is higher the more precisely the random variable is known. The precision is drawn from the prior distribution $\tau \sim \Gamma(\alpha_\tau, \beta_\tau)$. The alternative hypothesis model proposed by Chang and Feigenbaum (2006) is the LPPL model with a constant drift μ in the mean function (which was not included in the original JLS model):

$$\begin{aligned} r_i &\sim N(\mu(t_i - t_{i-1}) + \Delta H_{i,i-1}, \sigma^2(t_i - t_{i-1})), \\ \text{where} \quad \Delta H_{i,i-1} &= B(t_\text{c} - t_{i-1})^\beta \left[1 + \frac{C}{\sqrt{1+(\omega/\beta)^2}} \cos(\omega \ln(t_\text{c} - t_{i-1}) + \phi)\right] \\ &\quad - B(t_\text{c} - t_i)^\beta \left[1 + \frac{C}{\sqrt{1+(\omega/\beta)^2}} \cos(\omega \ln(t_\text{c} - t_i) + \phi)\right]. \end{aligned} \tag{14}$$

The LPPL model is characterized by the parameter vector $\xi = (A, B, C, \beta, \omega, \phi, t_\text{c})$, and these parameters are drawn independently from the following prior distributions:

$$\begin{aligned} &A \sim N(\mu_A, \sigma_A), \quad B \sim \Gamma(\alpha_B, \beta_B), \quad C \sim U(0,1), \quad \beta \sim B(\alpha_\beta, \beta_\beta), \\ &\omega \sim \Gamma(\alpha_\omega, \beta_\omega), \quad \phi \sim U(0, 2\pi), \quad t_\text{c} - t_N \sim \Gamma(\alpha_{t_\text{c}}, \beta_{t_\text{c}}), \end{aligned}$$

where Γ, B and U denote the gamma distribution, beta distribution and uniform distribution, respectively. Given the independence among prior distributions, the prior density for this model is simply given by the product of all marginal priors, while the probability data density for q_i is

$$f(q_i|q_{i-1}, \theta_\text{LPPL}; \text{LPPL}) = \sqrt{\frac{\tau}{2\pi(t_i - t_{i-1})}} \exp\left[-\frac{\tau(q_i - q_{i-1} - \mu(t_i - t_{i-1}) - \Delta H_{i,i-1})^2}{2(t_i - t_{i-1})}\right]$$

so that the likelihood function for the observed data Q is given by

$$f(Q|\theta_\text{LPPL}; \text{LPPL}) = \prod_{i=1}^{N} f(q_i|q_{i-1}, \theta_\text{LPPL}; \text{LPPL}).$$

Finally, the log marginal likelihood necessary for the computation of the Bayes factor is given by

$$\mathcal{L} = \ln\left(\int_\Theta f(Q|\theta_\text{LPPL}; \text{LPPL})\varphi(\theta_\text{LPPL}; \text{LPPL})\, \mathrm{d}\theta_\text{LPPL}\right),$$

which can be computed with Monte-Carlo methods and a large number of sampling values.

By using relatively diffuse priors with large variances in order to encompass the true values of the parameters, Chang and Feigenbaum (2006) found that the marginal likelihoods remain basically the same, whether they consider the LPPL specification in the mean function or only the drift term μ. This result remains robust to a change of prior distributions and they showed that the null hypothesis outperforms the JLS model in terms of marginal likelihood with different sets of priors.

Apart from the problem with weakly informative prior densities (see e.g. Bauwens, Lubrano, and Richard 2000) for a discussion, Lin, Ren, and Sornette (2009) pointed out that the Bayes approach to hypothesis testing assumes that some kind of ergodicity on a single data sample applies and that this sample has to be of sufficiently large size (which is not always the case). Clearly, this has to be tested and is far from being trivial. Furthermore, it is known that LPPL models can have likelihoods with several local maxima (see Jacobsson 2009) for a recent review, and the Bayes approach aims to solve this problem by integration, that is, by smoothing. However, for small to medium sample sizes, the smoothing in the marginal likelihoods can be harmful, particularly in the case of poor priors, and can decrease the number of local maxima at the price of a loss of efficiency. This may explain why the null hypothesis model with no log-periodic components showed a better result than the LPPL model.

3.2 *The generalized LPPL model with mean-reverting residuals*

The work by Chang and Feigenbaum (2006) represented the most important challenge to the original JLS model, and this is why it prompted a response by Sornette and his co-authors in 2009. Lin, Ren, and Sornette (2009) proposed a generalization of the original model which wants to make the process consistent with direct price calibration. As we have reported in the previous sections, the original JLS model has a random walk component with increasing variance, which makes direct estimation with prices inconsistent, as well as causes the lack of power of Bayesian methods, as shown by Lin, Ren, and Sornette (2009). Instead, the 'volatility-confined LPPL model' proposed by Lin, Ren, and Sornette (2009) combines a mean-reverting volatility process together with a stochastic conditional return which represents the continuous reassessments of investors' beliefs for future returns. As a consequence, the daily logarithmic returns are no longer described by a deterministic drift decorated by a Gaussian-distributed white noise, and the expected returns become stochastic.

Using the standard framework of rational expectations, Lin, Ren, and Sornette (2009) assumed that the price dynamics during a bubble is governed by the following process:

$$\frac{\mathrm{d}I}{I} = \mu(t)\mathrm{d}t + \sigma_Y \mathrm{d}Y + \sigma_W \mathrm{d}W - \kappa \mathrm{d}j,$$

$$\mathrm{d}Y = -\alpha Y \mathrm{d}t + \mathrm{d}W,$$

where I is the stock price index or the price of a generic asset, W is the standard Wiener process, $\mu(t)$ is a time-varying drift characteristic of a bubble regime and j is equal to zero before the crash and one afterwards, while κ represents the percentage by which the asset price falls during a crash. When $0 < \alpha < 1$, Y denotes an Ornstein–Uhlenbeck process, so that dY and Y are both stationary, and the volatility remains bounded till the crash. This property guarantees that direct estimation with prices is consistent. We remark that if $\alpha = 0$, we retrieve the original JLS model. The corresponding model in discrete time is given by

$$\ln I_{i+1} - \ln I_i = \mu_i + \sigma_Y(Y_{i+1} - Y_i) + \sigma_W \varepsilon_i - \kappa \Delta j_i, \tag{15}$$

$$Y_{i+1} = (1-\alpha)Y_i + \varepsilon_i, \quad \varepsilon_i \sim N(0,1). \tag{16}$$

Using the theory of the stochastic discount factor (SDF), complete markets and no arbitrage, Lin, Ren, and Sornette (2009) showed that the asset log returns follow this process:

$$\ln I_{i+1} = \ln I_i + \Delta H_{i,i-1} - \alpha(\ln I_i - H_i) + u_i, \tag{17}$$

where $\Delta H_{i,i-1}$ is given by expression (14) and u_i is a Gaussian white noise, while the conditional probability distribution for the logarithmic returns is given by

$$r_{i+1} = \ln I_{i+1} - \ln I_i \sim N(\Delta H_{i+1,i} - \alpha(\ln I_i - H_i), \sigma_u^2(t_{i+1} - t_i)). \tag{18}$$

Differently from the original JLS model, the additional term $-\alpha(\ln I_i - H_i)$ ensures that the log price fluctuates around the LPPL trajectory H_t, thus guaranteeing the consistency of direct estimation with prices.

Lin, Ren, and Sornette (2009) remarked that the previous model based on rational expectation separates rather artificially the noise traders and the rational investors. Moreover, even though the rational investors cannot make profit on average, rational agents endowed with different preferences may, in principle, arbitrage the risk-neutral agents. Therefore, assuming that rational investors have homogeneous preferences is rather restrictive. Nevertheless, Lin, Ren, and Sornette (2009) showed that the previous results can be obtained by using a complete different approach, which considers the theory of the so-called *behavioral SDF*, where the price movements follow the dynamics of the *market sentiment* (see Shefrin 2005) for a textbook treatment of the behavioral approach to asset pricing. We refer to Lin, Ren, and Sornette (2009) for more details about this alternative approach.

3.3 *Other generalizations: the log-periodic AR(1)–GARCH(1,1) model*

While the original LPPL specification can model the long-range dynamics of price movements, nevertheless it is unable to consider the short-term market dynamics, thus showing residual terms which can be strongly autocorrelated and heteroskedastic. As a consequence, Gazola et al. (2008) proposed the following AR(1)–GARCH(1,1) log-periodic model:

$$\begin{aligned} I_i &= A + B(t_c - t_i)^\beta + C(t_c - t_i)^\beta \cos[w \ln(t_c - t_i) + \phi] + u_i, \\ u_i &= \rho u_{i-1} + \eta_i, \\ \eta_i &= \sigma_i \varepsilon_i, \quad \varepsilon_i \sim N(0,1), \\ \sigma_i^2 &= \alpha_0 + \alpha_1 \eta_{i-1}^2 + \alpha_2 \sigma_{i-1}^2, \end{aligned} \tag{19}$$

where ε_i is a standard white noise term satisfying $E[\varepsilon_i] = 0$ and $E[\varepsilon_i^2] = 1$, whereas the conditional variance σ_i^2 follows a GARCH(1,1) process. Under the normality assumption for the error term ε_i, the ML estimator for the parameter vector $\mathbf{\Pi} = [A, B, C, t_c, \beta, w, \phi, \rho, \alpha_0, \alpha_1, \alpha_2]$ is obtained through the numerical maximization of the log-likelihood:

$$\ln L(\Theta) = -\frac{1}{2}(N-1)\ln(2\pi) - \frac{1}{2}\sum_{i=2}^{N} \ln \sigma_i^2 - \frac{1}{2}\sum_{i=2}^{N} \frac{\eta_i^2}{\sigma_i^2}. \tag{20}$$

In order to improve the optimization procedure, each parameter of the log-periodic model (19) denoted by θ and defined in a restricted interval denoted by $[a, b]$ can be re-parameterized according to the following monotonic transformation:

$$\theta = b\frac{\exp(\tilde{\theta})}{1+\exp(\tilde{\theta})} + a\left(1 - \frac{\exp(\tilde{\theta})}{1+\exp(\tilde{\theta})}\right). \tag{21}$$

This monotonic transformation turns the original estimation problem over a restricted space of solutions into an unrestricted problem, which eases estimation particularly when poor starting values are chosen. In this case, the delta method can be used to compute the standard errors of the estimate. We remind that the delta method is used to compute an estimator for the variance of functions of estimators and the corresponding confidence bands. Let $\hat{V}[\tilde{\theta}]$ be the estimated variance–covariance matrix of $\tilde{\theta}$, then, by using the delta method, a variance–covariance matrix for a general nonlinear transformation $g(\tilde{\theta})$ is given by (see Hayashi 2000 for more details)

$$\hat{V}[g(\tilde{\theta})] = \frac{\partial g(\tilde{\theta})}{\partial \tilde{\theta}'} \hat{V}[\tilde{\theta}] \frac{\partial g(\tilde{\theta})}{\partial \tilde{\theta}}' .$$

Gazola et al. (2008) used a two-step procedure to choose the starting values for the numerical maximization of Equation (20):

(1) the starting values for the set of parameters $\Phi = [A, B, C, t_c, \beta, w, \phi]$ are retrieved from the estimation of the original LPPL model (12);
(2) the starting values for the set of parameters $[\rho, \alpha_0, \alpha_1, \alpha_2]$ of the short-term stochastic component u_i are obtained by estimating an AR(1)–GARCH(1,1) model on the residuals $\hat{u}_i$ from the original LPPL model (12).

4. How to fit LPPL models?

Estimating LPPL models, in general, has never been easy due to the frequent presence of many local minima of the cost function where the minimization algorithm can get trapped. However, some recent developments have considerably simplified the estimation process.

4.1 *The original two-step nonlinear optimization*

Johansen, Ledoit, and Sornette (2000) noted that noisy data, relatively small samples and a large number of parameters make the estimation of LPPL models rather difficult. Therefore, they proposed to reduce the number of free parameters by slaving the three linear parameters and computing them from the estimated nonlinear parameters.

More specifically, if we rewrite the original LPPL model as follows:

$$y_i = A + B(t_c - t_i)^\beta + C(t_c - t_i)^\beta \cos(\omega \ln(t_c - t_i) + \phi) \tag{22}$$

or more compactly as

$$y_i = A + Bf_i + Cg_i,$$

where

$$y_i = \ln I_i \text{ or } I_i, \quad f_i = (t_c - t_i)^\beta$$
$$g_i = (t_c - t_i)^\beta \cos(\omega \ln(t_c - t_i) + \phi),$$

then it is straightforward to see that the linear parameters A, B and C can be obtained analytically by using ordinary least squares:

$$\begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} y_i f_i \\ \sum_{i=1}^{N} y_i g_i \end{pmatrix} = \begin{pmatrix} N & \sum_{i=1}^{N} f_i & \sum_{i=1}^{N} g_i \\ \sum_{i=1}^{N} f_i & \sum_{i=1}^{N} f_i^2 & \sum_{i=1}^{N} f_i g_i \\ \sum_{i=1}^{N} g_i & \sum_{i=1}^{N} f_i g_i & \sum_{i=1}^{N} g_i^2 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix}. \tag{23}$$

We can write the previous system compactly by using matrix notation:

$$\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X})\mathbf{b}, \quad \text{where } \mathbf{X} = \begin{pmatrix} 1 & f_1 & g_1 \\ \vdots & \vdots & \vdots \\ 1 & f_N & g_N \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} A \\ B \\ C \end{pmatrix} \tag{24}$$

so that

$$\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})\mathbf{X}'\mathbf{y}, \tag{25}$$

and we have only four free parameters to estimate (see also Jacobsson 2009 for a similar derivation). We remark that this simplification can also be seen as an example of concentrated ML.

The estimation procedure consists of two steps:

(1) Use the so-called *Taboo search* (Cvijović and Klinowski 1995) to find 10 candidate solutions, where only the cases with $B < 0$, $0 < \beta < 1$ and $t_c > t_i$ (if a bubble) are considered (see also Sornette and Johansen 2001). However, alternative grid searches can also be considered. Recently, Lin, Ren, and Sornette (2009) have imposed stronger restrictions, by considering $0.1 < \beta < 0.9$ and $6 \leq \omega \leq 15$ so that the log-periodic oscillations are neither too fast (to avoid fitting noise) nor too slow (otherwise they would provide a contribution to the trend), and $|C| < 1$ to ensure that the hazard rate $h(t)$ remains always positive.[2]

(2) Each of these 10 solutions is then used as the starting value in a Levenberg–Marquardt nonlinear least squares algorithm. The solution with the minimum sum of squares between the fitted model and the observations is taken as the final solution.

4.2 *Genetic algorithms*

The GA is an algorithm inspired by Darwin's 'survival of the fittest idea', and its theory was developed by John Holland in 1975. The GA is a computer simulation that aims to mimic the natural selection in biological systems, which is governed by four phases: a *selection mechanism*, a *breeding mechanism*, a *mutation mechanism* and a *culling mechanism*. The GA does not require the computation of any gradient or curvature and it does not need the cost function to be smooth or continuous.

The use of GA to estimate LPPL models has been proposed by Jacobsson (2009) following the GA methodology proposed by Gulsten, Smith, and Tate (1995). Similarly Johansen, Ledoit, and Sornette (2000), Jacobsson (2009) reduced the number of free parameters to four, by 'slaving' the three linear parameters A, B and C, which are computed by using Equation (25). Her procedure consists of four steps:

(1) *Selection mechanism: Generating the initial population.* Each member of the 'financial' population is represented by a vector of the four nonlinear coefficients t_c, ϕ, ω and β. The members of the initial population are randomly drawn from a uniform distribution with a pre-specified range, and for each member, the residual sum of squares is calculated. Jacobsson (2009) considered an initial population of 50 members (i.e. 50 parameter vectors).

(2) *Breeding mechanism*. The 25 members with the best value of the cost function are selected from the population to be included in the breeding program. An offspring is then generated by randomly drawing two parents, without replacement, and taking the arithmetic mean of

them. Jacobsson (2009) repeated this procedure 25 times, and each pair of parents is drawn randomly with replacement, so that one parent can generate an offspring with another parent (. . .therefore, betrayals are allowed!).

(3) *Mutation mechanism*. Genetic mutations in nature play a key role in the evolution of a species, since they may increase its probability of survival, as well as introduce less favorable characteristics. In our framework, mutations perturb the previous solutions to allow new regions of the search space to be explored, so that premature convergence in local minima can be avoided.

The mutation process is implemented by computing the statistical range ($\theta_{\max} - \theta_{\min}$) for each parameter in the population. The range for each parameter is then multiplied with a factor $\pm k$, to obtain the perturbation variable ε, which is uniformly distributed over the interval $[-k \times$ (parameter range), $k \times$ (parameter range)]. Jacobsson (2009) considered $k = 2$. Twenty-five members are then drawn randomly, without replacement, from the initial population of 50 computed in the first step. Each selected member is then mutated by adding an exclusive vector of random perturbations for every parameter. Therefore, the mutation mechanism allows to compensate the problem of an inaccurate guess for the initial intervals in the solution space.

(4) *Culling mechanism*. Jacobsson (2009) merged the members generated by mutation and breeding into the population, so that a total of 100 solutions is present (50 old, 25 offsprings and 25 mutations). All of the 100 solutions are ranked according to their cost function in ascending order, and the 50 best solutions are culled and live on into the next generation. The rest is deleted.

The previous algorithm is then iterated a certain number of times till a desired termination criterion is met. Similar to the second step of the optimization process used by Johansen, Ledoit, and Sornette (2000) and described in Section 4.1, Jacobsson (2009) further refined the parameters estimated with the GA by using them as starting values for the Nelder–Mead simplex method, also known as the downhill simplex method.

4.3 *The two-step/three-step ML approach*

Fantazzini (2010a) found that estimating LPPL models for 'anti-bubbles' was much easier than estimating LPPL models for bubbles: an anti-bubble is symmetric to a bubble and represents a situation when the market peaks at a critical time t_c and then decreases following a power law with decelerating log-periodic oscillations (see Johansen and Sornette 1999; Johansen and Sornette 2000; Zhou and Sornette 2005; Fantazzini 2010b for more details). Furthermore, estimating models with log prices was much simpler than estimating models with prices in levels, and in the latter case, a much more careful choice of the starting values had to be made.

In this regard, we have already seen that the original LPPL model has a stochastic random walk component with increasing variance, so that the deterministic pattern moves away from the observable price path. Therefore, the idea given by Fantazzini (2010a) is to reverse the original times series in order to minimize the effect of the non-stationary component during the estimation process. The same idea can also be of help in the case of models with stationary error terms, like the one proposed by Lin, Ren, and Sornette (2009): a time series with strongly autocorrelated error terms is almost undistinguishable from a non-stationary process in small-to-medium-sized samples (see e.g. Stock 1994; Ng and Perron 2001; Stock and Watson 2002 for a discussion of this hotly debated issue in the econometric literature). Figure 1 shows a simulated LPPL model with AR(1) error terms and 1000 observations, as well as its reverse.

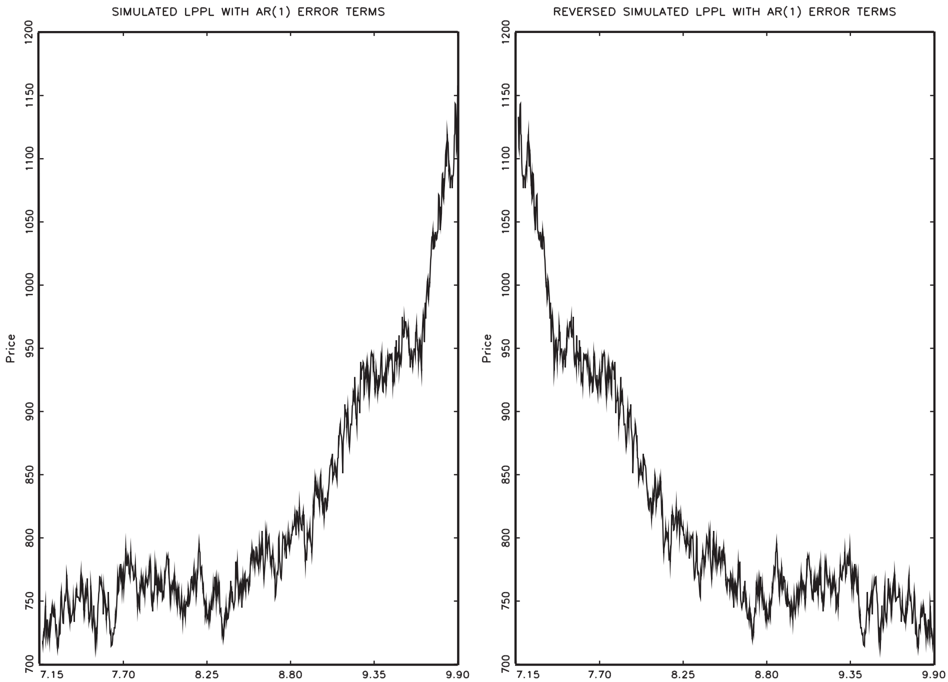

Figure 1. Simulated LPPL with AR(1) error terms and its reverse. Parameters: $A = 7.16$, $B = -0.43$, $C = 0.035$, $\beta = 0.35$, $\omega = 4.15$, $\phi = 2.07$, $t_c = 9.92$, $\rho = 0.88$, $\alpha_0 = 0.00007$.

The two-step ML approach used in Fantazzini (2010a) to estimate LPPL models for financial bubbles, also allowing for an AR(1)–GARCH(1,1) model in the error terms as in Equation (19), is given below:

(1) Reverse the original time series and estimate the LPPL for the case of an anti-bubble by using the BFGS (Broyden, Fletcher, Goldfarb, Shanno) algorithm, together with a cubic or quadratic step length method (STEPBT) (see e.g. Dennis and Schnabel 1983).
(2) Keeping fixed the LPPL parameters $\hat{\Phi} = [\hat{A}, \hat{B}, \hat{C}, \hat{t}_c, \hat{\beta}, \hat{\omega}, \hat{\phi}]$ computed in the first stage, estimate the parameters of the short-term stochastic component $[\rho, \alpha_0, \alpha_1, \alpha_2]$.

In case of poor starting values, or when the bubble has just started forming (Jiang et al. 2010 and Sornette 2003a remarked that a bubble cannot be diagnosed more than one year in advance of the crash), the numerical computation can be further eased by considering one additional step. The three-step ML approach used in Fantazzini (2010a) is described below:

(1) Reverse the original time series and then consider the first temporal observation as if it was the date of the crash, that is, set $t_c = t_1$. Estimate the remaining LPPL parameters $[A, B, C, \beta, \omega, \phi]$ for the case of an anti-bubble by using the BFGS (Broyden, Fletcher, Goldfarb, Shanno) algorithm, together with a cubic or quadratic step length method (STEPBT).
(2) Use the estimated parameters in the previous step as *starting values* for estimating all the LPPL parameters by using again the reversed times series.

(3) Keeping fixed the LPPL parameters $\hat{\Phi} = [\hat{A}, \hat{B}, \hat{C}, \hat{t}_c, \hat{\beta}, \hat{\omega}, \hat{\phi}]$ computed in the second stage, estimate the parameters of the short-term stochastic component $[\rho, \alpha_0, \alpha_1, \alpha_2]$.

Being a multi-stage estimation process, the asymptotic efficiency is lower than that of the one-step full ML estimation. However, the dramatic improvement in numerical convergence and the improved efficiency in small-to-medium-sized samples more than justify the multi-step procedure. An (unreported) simulation study confirms the benefits of this procedure in small-to-medium data sets.

5. Diagnosing bubbles in the making

The main method that we have considered so far to detect financial bubbles is by fitting an LPPL model to a price series. However, in order to reduce the possibility of false alarms, it is good practice to implement a battery of tests, so that a prediction must pass all tests to be considered worthy (see e.g. Sornette and Johansen 2001; Jiang et al. 2010).

5.1 *Diagnostic tests based on the LPPL fitting residuals*

We have seen in Section 3.2 that Lin, Ren, and Sornette (2009) proposed a model for financial bubbles where the LPPL fitting residuals follow a mean-reverting Ornstein–Uhlenbeck process. This implies that the corresponding residuals follow an AR(1) process and we can test this hypothesis by using unit-root test.

Lin, Ren, and Sornette (2009) used Phillips–Perron and augmented Dickey–Fuller (ADF) unit-root tests, where the null hypothesis H_0 is the presence of a unit root. Unfortunately, a well-known shortcoming of the previous two unit-root tests is their low power when the underlying data-generating process is an AR(1) process with a coefficient close to one. Therefore, we suggest to also consider the test proposed by Kwiatkowski et al. (1992), where the null hypothesis is a stationary process. Considering the null hypothesis of a stationary process and the alternative of a unit root allows us to follow a conservative testing strategy: if we reject the null hypothesis, we can be confident that the series has indeed a unit root; but if the results of the previous two tests indicate a unit root while the result of the KPSS test indicates a stationary process, one should be very cautious and opt for the latter result.

The KPSS test is implemented in the most common statistical and econometric software (see e.g. Griffiths, Hill, and Lim 2008) for applications with Eviews as well as the Eviews User's Guide (version 5 or higher) and Pfaff (2008) for a description of unit-root tests in R, together with the many routines written for Gauss and Matlab, which can be found on the web.

5.2 *Diagnostic tests based on rational expectation models with stochastic mean-reverting termination times*

Lin and Sornette (2009) proposed two models of transient bubbles in which their termination dates occur at some potential critical time $\tilde{t}_c$, which follows a stationary process with a unconditional mean T_c. The main advantage of these models is the possibility of computing the potential critical time without the need to estimate the complex stochastic differential equation describing the underlying price dynamics. Interestingly, the rational arbitrageurs discussed in Lin and Sornette (2009) can detect bubbles, but they cannot make a deterministic forecast because they have little knowledge about the other arbitrageurs' beliefs about the process governing the stochastic critical time $\tilde{t}_c$. The heterogeneity of the rational agents' expectations determines a synchronization problem among these arbitrageurs, thus allowing the financial bubble to survive till its theoretical end time (see also Abreu and Brunnermeier 2003 for a similar model). Moreover, the two models proposed by Lin and Sornette (2009) can be tested and they allow us to diagnose financial bubbles in the making in real time. Besides, both models highlight the importance of positive feedback, that is, when a high price pushes even further the demand so that the return and its volatility tend to be a nonlinear accelerating function of the price. This positive feedback mechanism is quantified by a unique exponent m, which is larger than 1 (respectively, 2 for the second model) when we are in a bubble regime.

5.2.1 A test based on a finite-time singularity in the price dynamics with stochastic critical time

The first model proposed by Lin and Sornette (2009) views a bubble as a faster-than-exponential accelerating stochastic price, which leads to a finite-time singularity in the price dynamics at a stochastic critical time. They showed in their Proposition 1 that the price dynamics in a bubble regime follows this process:

$$\begin{aligned} &p(t) = K(\tilde{T}_c - t)^{-\beta}, \\ &\beta = \frac{1}{m-1}, \quad K = \left(\frac{\beta}{\mu}\right)^{\beta}, \quad T_c = \frac{\beta}{\mu} p_0^{-1/\beta}, \quad \tilde{T}_c = T_c + \tilde{t}_c, \end{aligned} \tag{26}$$

where μ is the instantaneous return rate and p_0 denotes the price at the start time of the bubble at $t = 0$, while the critical time $\tilde{t}_c$ follows an Ornstein–Uhlenbeck process with zero unconditional mean (see Lin and Sornette 2009 for the full derivation of the model). The last property provides that the end of the bubble cannot be forecasted with certainty but it is a stochastic variable, while the time T_c can be interpreted as the consensus forecast formed by rational arbitrageurs of the stochastic critical time $\tilde{T}_c$. In fact, we have that

$$E[\tilde{T}_c] = E[T_c + \tilde{t}_c] = T_c. \tag{27}$$

In order to build a diagnostic test for financial bubbles, Lin and Sornette (2009) inverted Equation (26) to obtain an expression for the critical time series $\tilde{T}_{c,i}$:

$$\tilde{T}_{c,i}(t) = \frac{1}{K}\frac{1}{[p(t)]^{1/\beta}} + t, \quad t = t_i - 749, \ldots, t_i, \tag{28}$$

where $\tilde{T}_{c,i}$ is defined over the time window i ending at time t_i, and they considered time windows of 750 trading days that slide with a time step of 25 days from the beginning to the end of the available financial time series. We remark that $p(t)$ is known, while the parameters K and β have to be estimated. The previous inversion aims to transform a non-stationary possibly explosive price process $p(t)$ into what should be a stationary time series $\tilde{T}_{c,i}$ in the absence of misspecification. Therefore, we can then estimate T_c according to Equation (27) by using the arithmetic average of $\tilde{T}_{c,i}(t)$:

$$T_{c,i} = \frac{1}{750}\sum_{i=1}^{750} \tilde{T}_{c,i}(t) \tag{29}$$

so that the fluctuations $\tilde{t}_{c,i}(t)$ can be computed as

$$\tilde{t}_{c,i}(t) = \tilde{T}_{c,i}(t) - T_{c,i}. \tag{30}$$

The first test proposed by Lin and Sornette (2009) consists of the following two steps:

(1) Perform a bivariate grid search over the parameter space of K and β to find the 10 best pairs (K, β) such that the resulting time series $\tilde{t}_{c,i}(t)$ given by Equation (30) rejects a standard unit-root test of non-stationarity at the 99.5% significance level. Lin and Sornette (2009) employed the ADF test, but the addition of the KPSS test would be advisable. Needless to say, only a subset of the windows will reject the null hypothesis of a unit root (for the ADF test) or will not reject the null of stationarity (for the KPSS test).

(2) If there are time windows for which there are selected pairs (K, β) according to the previous step, select the pair with the smallest variance for its corresponding time series $\tilde{t}_{c,i}(t)$. This gives the optimal pair K_i^* and β_i^* which provides the closest approximation to a stationary time series for $\tilde{t}_{c,i}(t)$ given by Equation (30). For a given window i, an alarm is declared when
   - $\beta^* > 0$, which yields $m > 1$;
   - $T_{c,i} - t_i < 750$, which implies that the termination time of the bubble is not too far. Lin and Sornette (2009) also considered two additional alarm levels: $T_{c,i} - t_i < 500$ and $T_{c,i} - t_i < 250$.

The idea of the last step is that the closer we are to the end of the financial bubble, the stronger should be the evidence for the bubble as a faster-than-exponential growth, and the alarms should be diagnosed repeatedly by several successive windows.

5.2.2 *A test based on a finite-time singularity in the momentum price dynamics with stochastic critical time*

The main disadvantage of the previous model is that the price diverges when approaching the critical time $\tilde{T}_c$ at the end of the bubble. Therefore, Lin and Sornette (2009) considered a second model where the price remains always finite and a bubble is a regime characterized by an accelerating momentum ending at a finite-time singularity with a stochastic critical time. They showed in their Proposition 3 that the log price $y(t) = \ln p(t)$ in a bubble regime follows this process:

$$
\begin{gathered}
y(t) = A - B(T_c + \tilde{t}_c(t) - t)^{1-\beta}, \\
\beta = \frac{1}{m-1}, \quad T_c = \frac{\beta}{\mu} x_0^{1/\beta}, \quad x_0 := x(t=0), \quad B = \frac{1}{1-\beta}\left(\frac{\beta}{\mu}\right)^{\beta}, \quad \tilde{T}_c = T_c + \tilde{t}_c,
\end{gathered} \tag{31}
$$

where μ is the instantaneous return rate, A is a constant and $x(t) = \mathrm{d}y/\mathrm{d}t$ denotes the effective price momentum, that is, the instantaneous time derivative of the logarithm of the price, while the critical time $\tilde{t}_c$ follows an Ornstein–Uhlenbeck process with zero unconditional mean (see Lin and Sornette 2009 for the full derivation of this model). In the second model, it is the high price momentum x which pushes the demand higher, so that the return and its volatility become nonlinear accelerating functions. Instead, in the first model, it is the price that provides a positive feedback on future prices, rather than the price momentum.

Using a procedure similar to Equation (28) to transform a non-stationary possibly explosive log-price process $y(t)$ into a stationary time series $\tilde{T}_{c,i}$, Lin and Sornette (2009) inverted Equation (31) to obtain an expression for the critical time series $\tilde{T}_{c,i}$:

$$
\tilde{T}_{c,i}(t) = \left(\frac{A - \ln p(t)}{B}\right)^{1/(1-\beta)} + t, \quad t = t_i - 899, \ldots, t_i, \tag{32}
$$

where $\tilde{T}_{c,i}$ is defined over the time window i ending at time t_i, and they considered time windows of 900 trading days that slide with a time step of 25 days from the beginning to the end of the available financial time series. We can then estimate $T_{c,i}$ according to Equation (29) by computing the arithmetic average of $\tilde{T}_{c,i}(t)$ (with 750 replaced by 900), whereas the fluctuations $\tilde{t}_{c,i}(t)$ around $T_{c,i}$ can be computed by using Equation (30). Similar to the first model, we remark that $p(t)$ is known, while the parameters A, B and β have to be estimated.

The second test proposed by Lin and Sornette (2009) consists of the following two steps:

(1) Perform a trivariate grid search over the parameter space of A, B and β to find the 10 best triplets (A, B, β) such that the resulting time series $\tilde{t}_{c,i}(t)$ given by Equation (30) rejects a standard unit-root test of non-stationarity at the 99.5% significance level. Lin and Sornette (2009) employed the ADF test, but again the addition of the KPSS test would be advisable.

(2) If there are time windows for which there are selected triplets (A, B, β) according to the previous step, select the pair with the smallest variance for its corresponding time series $\tilde{t}_{c,i}(t)$. This gives the optimal triplet A_i^*, B_i^* and β_i^* which provides the closest approximation to a stationary time series for $\tilde{t}_{c,i}(t)$ given by Equation (30). For a given window i, an alarm is declared when
   - $0 < \beta_i^* < 0$, which yields $m > 2$ and is called *level 1* filter. Lin and Sornette (2009) also considered two additional alarm levels: $m > 2.5$ (*level 2*) and $m > 3$ (*level 3*);
   - $-25 \leq T_{c,i} - t_i \leq 50$.

The stronger upper bound on $T_{c,i} - t_i$ stems from the fact that the finite-time singularity in the price momentum is a weaker singularity which can only be observed only close to the critical time. The lower bound of -25 days is due to the fact that the analysis is performed in sliding windows with a time step of 25 days. Using a data set covering the last 30 years of the SP500 index, the NASDAQ composite index and the Hong Kong Hang Seng index, Lin and Sornette (2009) found that the second diagnostic method was more reliable and with fewer false alarms than the first method analyzed in Section 5.2.1.

5.3 *Graphical tools: the crash lock-in plot*

Fantazzini (2010a) proposed a graphical tool that proved to be useful to track the development of a bubble and to understand whether a possible crash is in sight or at least a bubble deflation. The idea is to plot on the horizontal axis the *date of the last observation in the estimation sample* and on the vertical axis the *estimated crash date* $\hat{t}_c$ computed by fitting the LPPL to the data: if a change in the stock market regime is approaching, then the recursively estimated $\hat{t}_c$ should

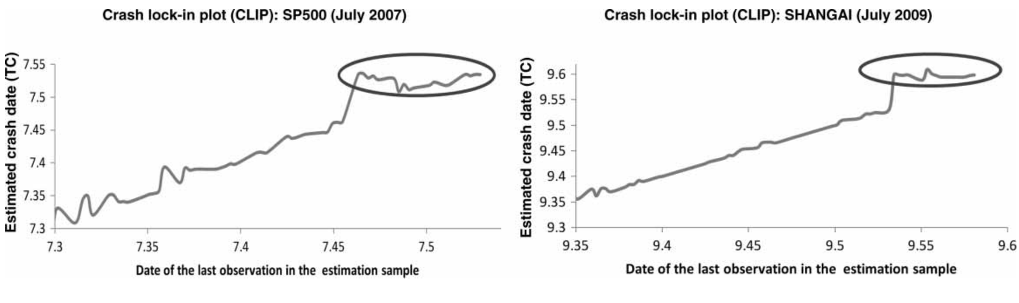

Figure 2. CLIPs for the SP500 and Shanghai composite index.

stabilize around a constant value close to the critical time. Fantazzini (2010a) called such a plot the crash lock-in plot (CLIP).

This idea can be easily justified theoretically by resorting to the models proposed by Lin and Sornette (2009), in which the critical time $\tilde{T}_c$ follows an Ornstein–Uhlenbeck process.

We report in Figure 2 the CLIPs for the Chinese Shanghai Composite Index in July 2009, a case which was analyzed in detail in Bastiaensen et al. (2009) and Jiang et al. (2010), and for the SP500 in July 2007, with these being the peak of the market in the decade. We used data spanning from the global minima till one day before the market peak.

6. An application: the burst of the gold bubble in December 2009

The gold market peaked on 2 December 2009, hitting the record high at $1216.75 an ounce in Europe, and then started falling on 4 December 2009, losing more than 10% in two weeks. The main concerns cited to be behind this bubble were the future prospects for a weak dollar as well as inflationary fears (see e.g. Mogi 2009; White 2009). However, there were also some worried calls about the possibility of a gold bubble: the prestigious magazine *Fortune* wrote on 12 October 2009 that '. . .Signs of gold fever are everywhere. . .' but '. . .amid the buying frenzy and after a decade-long run-up that has seen the price quadruple, is gold still a smart investment? The simple answer: Wherever the price of gold is headed in the long term, several market watchers say the fundamentals indicate that gold is poised to fall' (Cendrowski 2009). Interestingly, on the day the gold price peaked, that is, 2 December 2009, Hu Xiaolian, a vice-governor at the People's Bank of China, told reporters in Taipei that '. . .gold prices are currently high and markets should be careful of a potential asset bubble forming. . .', see the original report by Tung (2009). The gold price, starting from 12 November 2008 (which represents the global minima over a three-year span) till the end of January 2010 is reported in Figure 3. This figure also reports the 'Search Volume Index' by Google Trends, which computes how many searches have been done for the term 'Gold Price' on Google over time.[3]

The 'Search Volume Index' is an interesting tool because it allows us to get some insights as to when the bubble started: looking at Figure 3, we can see that a massive interest around gold started to build during the year 2008, just before the price minima in November 2008. Therefore, we expect that a possible LPPL model can be fitted using data starting from the year 2008.

6.1 *LPPL fitting with varying window sizes*

Jiang et al. (2010) tested the stability of LPPL estimation parameters by varying the size of the estimation samples and adopting the strategy of fixing one endpoint and varying the other one (see also Sornette and Johansen 2001). By sampling many intervals as well as by using bootstrap techniques, they obtained probabilistic predictions on the time intervals in which a given bubble may end and lead to a new market regime (which may not necessarily be a crash, but also a transition to a plateau or a slower decay). Following their example, we fit the logarithm of the gold price by using the LPPL Equation (22) in *shrinking windows* and in *expanding windows*. The shrinking windows have a fixed end date $t_2 = 1$ December 2009, while the starting date t_1 increases from 12 November 2008 to 17 August 2009 in steps of 5 (trading) days. The expanding windows have a fixed starting date $t_1 = 12$ November 2008, while the end date t_2 increases from 17 August 2009 to 1 December 2009 in steps of 5 (trading) days.

Given the stochastic nature of the initial parameter selection and the noisy nature of the underlying generating processes, we employed four estimation algorithms: the original Taboo Search

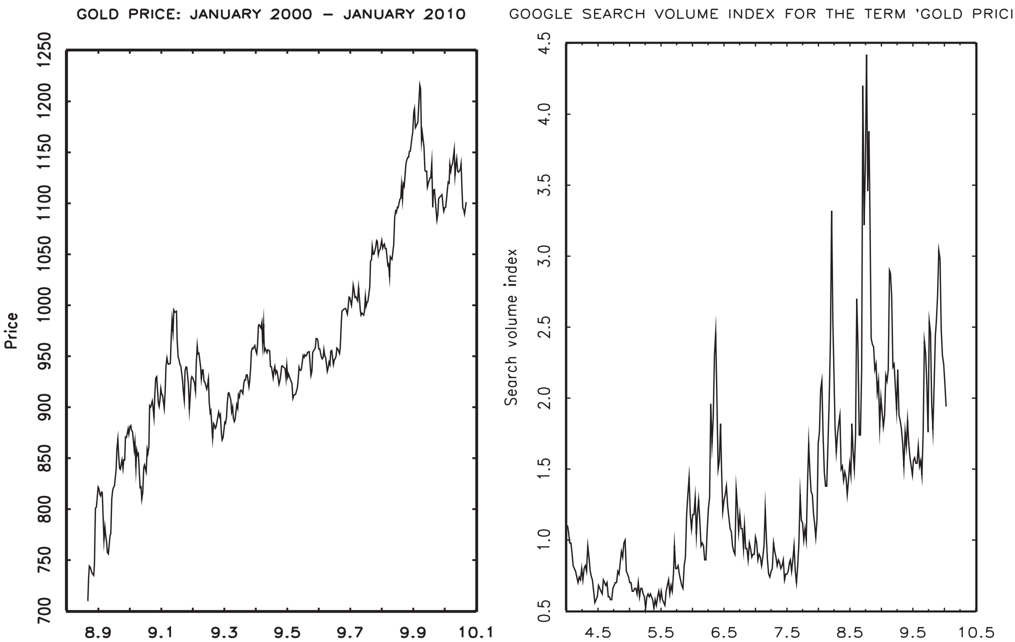

Figure 3. Gold price and Google search volume index. Time t converted in units of 1 year.

algorithm proposed by Cvijović and Klinowski (1995), the two-step nonlinear optimization proposed by Johansen, Ledoit, and Sornette (2000), and the pure random search (PRS) and the three-step ML approach proposed by Fantazzini (2010a).[4]. The estimation results are then filtered by the following LPPL conditions, which were also used in Jiang et al. (2010) for the case of Chinese bubbles: $\hat{t}_c > t_2$, $B < 0$ and $0 < \beta < 1$. The selected $\hat{t}_c$ are then used to compute the 20%/80% and 5%/95% quantile range of values of the crash dates, which are reported in Figure 4: the left plot shows the ranges which are obtained by considering the filtered results from all four estimation methods, whereas the right plot shows the ranges obtained by considering only the two-step nonlinear optimization and the three-step ML approach.

As expected, the original Taboo Search and the PRS are very inefficient methods compared with the competing two-step and three-step approaches and deliver very large quantile ranges. Nevertheless, the two medians $\hat{t}_c$, equal to 11 December 2009 for the left plot and 5 December 2009 for the right plot, are very close to the actual market peak date, which is 2 December 2009 (i.e. 9.9206 when converted in units of one year). Moreover, if we consider only the most efficient methods, the 20%/80% quantile interval is rather close and precise and diagnoses that the critical time t_c for the end of the bubble and the change of market regime lies in the time sample 3 December 2009–11 December 2009 (the market started to fall on 4 December 2009).

6.2 *Diagnostic tests based on the LPPL fitting residuals*

We discussed in Section 3.2 that Lin, Ren, and Sornette (2009) proposed a model for financial bubbles where the LPPL fitting residuals follow a mean-reverting Ornstein–Uhlenbeck process. Therefore, the corresponding residuals should follow a stationary AR(1) process and this hypothesis can be tested by using unit-root tests. We employed ADF and KPSS tests: a rejection of the

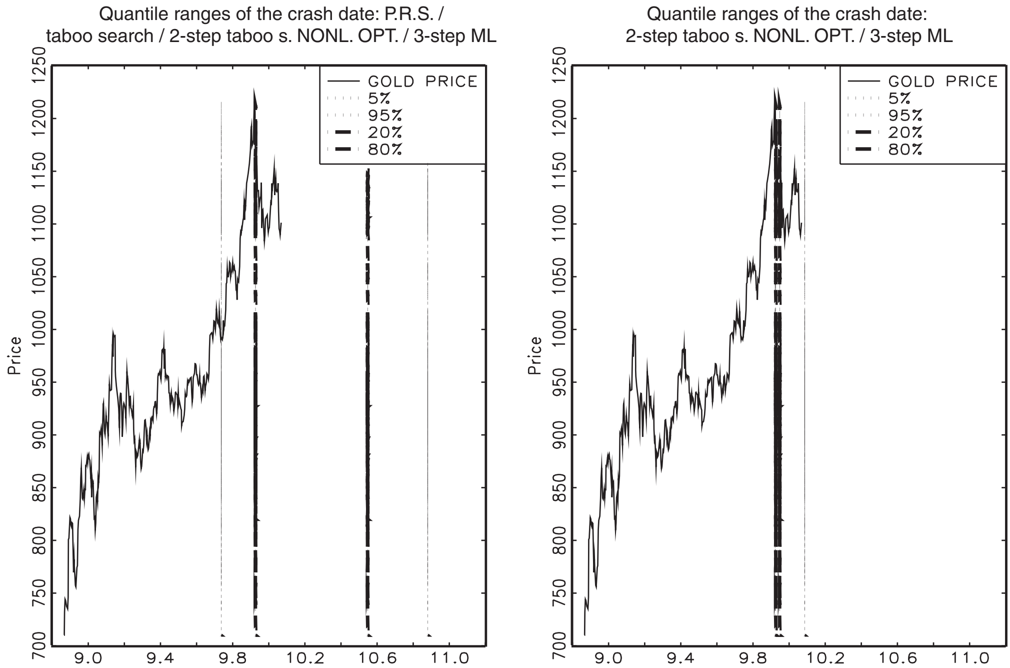

Figure 4. Quantile ranges of the crash date.

null hypothesis in the first test, together with a failure to reject the null in the second test, indicates that the residuals are stationary and thus compatible with an O–U process.

We used the residuals resulting from the previous estimation windows and numerical algorithms, that is, 48 shrinking windows, 19 expanding windows and four estimation methods, which give a total of 268 calibrations. The fraction P_{LPPL} of these different windows that met the LPPL conditions was equal to $P_{\text{LPPL}} = 60.1\%$. The conditional probability that, out of the fraction P_{LPPL} of windows that satisfied the LPPL conditions, the null hypothesis of non-stationarity was rejected for the residuals was equal to $P_{\text{Stat.Res.}|\text{LPPL}} = 100\%$ when using the ADF test at the probability level $\alpha = 0.001$. As for the KPSS test, the null of stationarity was not rejected at the 10% level or higher in all cases which satisfied the LPPL conditions. Therefore, this empirical evidence is comparable with the results reported by Jiang et al. (2010) for the case of the 2005–2007 and 2008–2009 Chinese stock market bubbles.

6.3 *Diagnostic tests based on rational expectation models with stochastic mean-reverting termination times*

We employed the two diagnostics proposed by Lin and Sornette (2009) to detect the presence of a bubble (and reviewed in Section 5.2) in the gold price time series, from 12 November 2008 to 1 December 2009. As discussed previously, we considered both the ADF and KPSS unit-root tests. Moreover, we also used time windows of 500 and 250 trading days to compute the critical time series $\tilde{T}_{c,i}$ in Equations (28)–(29) and Equation (32), together with the original 750 trading days for the first diagnostic and 900 for the second one. The rationale for this choice is that a long time span may include data which are not observed during a bubble regime but during a

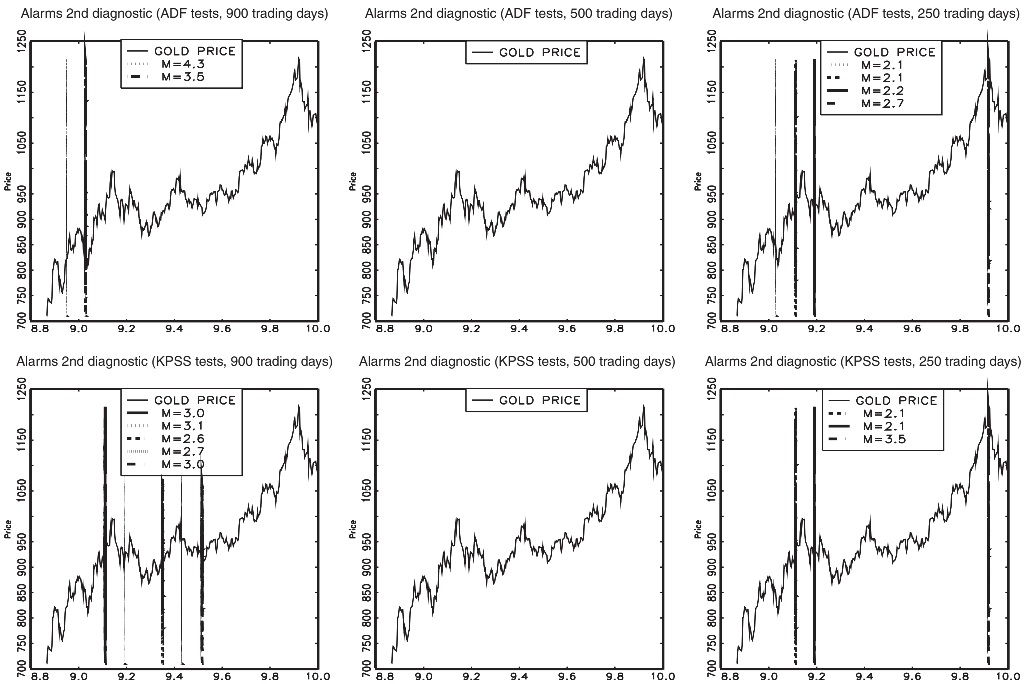

Figure 5. Logarithm of the gold price and corresponding alarms as vertical lines indicating the ends of the windows of T trading days, in which the second diagnostic flags an alarm for the presence of a bubble. The value of the exponent m for each alarm is reported in the legend.

standard geometric Brownian motion regime (or other regimes). Of course, reducing the time window implies a loss of efficiency.

Interestingly, the first procedure did not flag any alarm for the presence of a bubble, whereas the second one flagged three series of alarms close to three important price falls (Figure 5): the first group of alarms was centered around the local market peak on 20 February 2009 when gold reached the value of $995.3 an ounce, very close to the important psychological barrier of $1000, and after two days, it started falling, losing more than 10% in two weeks. The second group of alarms was centered around the local market peak on 2 June 2009 when gold reached the value of $982.9, and after two days, it started falling, losing more than 5% in a week. Finally, the third group of alarms was centered around the global market peak on 2 December 2009 when gold reached the value of $1216 an ounce.

This empirical evidence seems to suggest that the KPSS test provides more precise alarms than the ADF test, which is not a surprise, given the well-known limitations of the latter test. Moreover, a time window of 250 observations delivers more reliable flags for the presence of a bubble (or imminent price falls) than longer time spans, with it being more robust to market regime changes. However, a time window of 900 observations still provides useful information.

6.4 *CLIPs for the gold bubble*

The CLIP plots on the horizontal axis the date of the last observation in the estimation sample and on the vertical axis the estimated crash date $\hat{t}_c$ computed by fitting the LPPL to the data.

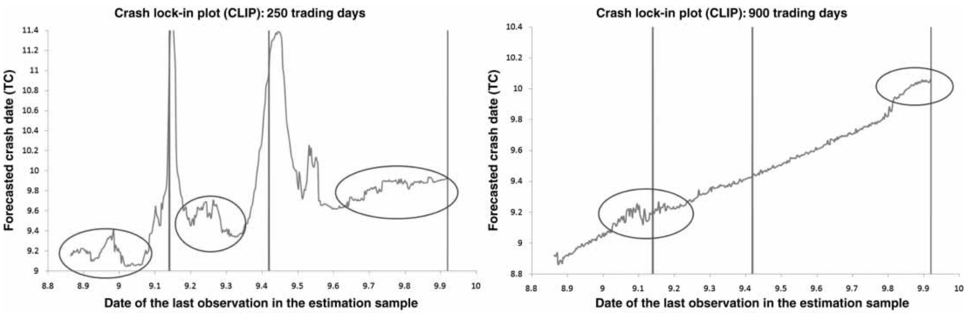

Figure 6. CLIPs for the gold price series. The three vertical lines correspond to the two local market peaks on 20 December 2009 and 2 June 2009 and to the global market peak on 2 December 2009, respectively.

Following the previous empirical evidence as well as the one reported in Lin and Sornette (2009) and Fantazzini (2010a), we computed the CLIP by fitting the data with two rolling estimation windows of 900 and 250 days and by using the simple average of the estimated $\hat{t}_c$ resulting from the four estimation algorithms discussed in Section 6.1. We used data spanning from 12 November 2008 till 1 day before the global market peak on 2 December 2009. The two CLIPs are shown in Figure 6.

Not surprisingly, the indications provided by the two CLIPs are rather similar to those provided by the second diagnostic test proposed by Lin and Sornette (2009) discussed in the previous section: the recursive forecasted crash dates computed with time windows of 250 trading days stabilize around three constant values which are very close to the dates corresponding to the two local market peaks on 20 February 2009 and 2 June 2009 and to the global market peak on 2 December 2009. The indications from the second CLIP computed with time windows of 900 trading days are somewhat weaker, but confirm the previous alarms. As expected, the estimates computed with smaller time spans are more noisy than those computed with longer time spans.

7. Conclusions

We presented an easy-to-use and self-contained guide for modeling and detecting financial bubbles with LPPLs, which contains the sufficient steps to derive the main models and discusses the important aspects for practitioners and researchers. We reviewed the original JLS model and we discussed early criticism to this approach and recent generalizations proposed to answer these remarks. Moreover, we described three different estimation methodologies which can be employed to estimate LPPL models. We then examined the issue of diagnosing bubbles in the making by using a set of different techniques, that is, by considering diagnostic tests based on the LPPL fitting residuals and diagnostic tests based on rational expectation models with stochastic mean-reverting termination times, as well as graphical tools useful for capturing bubble development and for understanding whether a crash is in sight or not. We finally presented a detailed empirical application devoted to the burst of the gold bubble in December 2009, which highlighted how a series of different diagnostics flagged an alarm for the presence of a bubble before prices started to fall.

Notes

1. In the context of the alignment of atomic spins to create magnetization, this model represented by Equation (4) is identical to the so-called two-dimensional *Ising model* which was solved explicitly by Onsager (1944), and where the disorder parameter is represented by the temperature of the system.
2. The lower bound on ω is rather strong, given that Sornette (2003a) found that $\omega = 6.36 \pm 1.56$ by using a large collection of empirical evidence. Moreover, the recent work about Chinese market bubbles by Jiang et al. (2010) considers only the former set of conditions.
3. See http://www.google.com/intl/en/trends/about.html for more details. In this case, the time span starts from 2004, which is the first year available for this analysis.
4. Similar to Cvijović and Klinowski (1995), we found that GA has a performance in between Taboo Search and PRS. However, even though PRS is computationally inefficient, it has the benefit to potentially visit regions of the parameter space that sometimes are not visited by the previous algorithms. This is why we consider it in our analysis in the place of GA.

References

Abreu, D., and M.K. Brunnermeier. 2003. Bubbles and crashes. *Econometrica* 71, no. 1: 173–204.

Bastiaensen, K., P. Cauwels, D. Sornette, R. Woodard, and W.X. Zhou. 2009. The Chinese equity bubble: ready to burst. http://arxiv.org/abs/0907.1827.

Bauwens, L., M. Lubrano, and J.F. Richard. 2000. *Bayesian inference in dynamic econometric models*. New York: OUP.

Bernardo, J.M., and A.F.M. Smith. 1994. *Bayesian theory*. New York: Wiley.

Blanchard, O.J. 1979. Speculative bubbles, crashes and rational expectations. *Economics Letters* 3, no. 4: 387–89.

Cendrowski, S. 2009. Beware the gold bubble. http://money.cnn.com/2009/10/06/pf/gold_investing_bubble.fortune/index.htm.

Chang, G., and J. Feigenbaum. 2006. A Bayesian analysis of log-periodic precursors to financial crashes. *Quantitative Finance* 6, no. 1: 15–36.

Cvijović, D., and J. Klinowski. 1995. Taboo search – an approach to the multiple minima problem. *Science* 267, no. 5: 664–6.

Dennis, Jr., J.E., and R.B. Schnabel. 1983. *Numerical methods for unconstrained optimization and nonlinear equations*. Englewood Cliffs, NJ: Prentice-Hall.

Derrida, B., L. De Seze, and C. Itzykson. 1983. Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics* 33, no. 3: 559–69.

Fantazzini, D. 2010a. Modelling bubbles and anti-bubbles in bear markets: a medium-term trading analysis. In *Handbook of trading*, ed. G. Gregoriou, 365–88. New York: McGraw-Hill.

Fantazzini, D. 2010b. Modelling and forecasting the global financial crisis: initial findings using heteroskedastic log-periodic models. *Economics Bulletin* 30, no. 3: 1833–41.

Financial Crisis Observatory website. http://www.er.ethz.ch/fco.

Gazola, L., C. Fernandes, A. Pizzinga, and R. Riera. 2008. The log-periodic-AR(1)-GARCH(1,1) model for financial crashes. *The European Physical Journal B* 61, no. 3: 355–62.

Goldenfeld, N. 1992. *Lectures on phase transitions and the renormalization group*. Reading, MA: Addison-Wesley.

Greenspan, Alan. 2002. Remarks by Chairman Alan Greenspan at a symposium sponsored by the Federal Reserve Bank of Kansas City, Jackson Hole, Wyoming, 30 August.

Griffiths, W.E., R. Hill, and G.C. Lim. 2008. *Using EViews for principles of econometrics*. 3rd edn. Hoboken, NJ: Wiley.

Gulsten, M., E.A. Smith, and D.M. Tate. 1995. A genetic algorithm approach to curve fitting. *International Journal of Production Research* 33, no. 7: 1911–23.

Gürkaynak, R. 2008. Econometric tests of asset price bubbles: taking stock. *Journal of Economic Surveys* 22, no. 1: 166–86.

Hayashi, F. 2000. *Econometrics*. Princeton, NJ: PUP.

Jacobsson, E. 2009. How to predict crashes in financial markets with the log-periodic power law. Master diss., Department of Mathematical Statistics, Stockholm University.

Jiang, Z.Q., W.H. Zhou, D. Sornette, R. Woodard, K. Bastiaensen, and P. Cauwels. 2010. Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. *Journal of Economic Behavior and Organization* 74: 149–62. http://arxiv.org/abs/0909.1007.

Johansen, A. 2002. Comment on ‘Are financial crashes predictable?’. *Europhysics Letters* 60, no. 5: 809–10.

Johansen, A. 2003. Characterization of large price variations in financial markets. *Physica A* 324, no. 1: 157–66.

Johansen, A., O. Ledoit, and D. Sornette. 2000. Crashes as critical points. *International Journal of Theoretical and Applied Finance* 3, no. 2: 219–55.

Johansen, A., and D. Sornette. 1999. Financial anti-bubbles: log-periodicity in gold and Nikkei collapses. *International Journal of Modern Physics C* 10, no. 4: 563–75.

Johansen, A., and D. Sornette. 2000. Evaluation of the quantitative prediction of a trend reversal on the Japanese stock market in 1999. *International Journal of Modern Physics C* 11, no. 2: 359–64.

Johansen, A., and D. Sornette. 2004. Endogenous versus exogenous crashes in financial markets. In *Contemporary Issues in International Finance*. Nova Science Publishers. Reprinted as Shocks, crashes and bubbles in financial markets. *Brussels Economic Review (Cahiers economiques de Bruxelles)* 49, no. 3/4, Special Issue on Nonlinear Analysis, 2006.

Kaizoji, T., and D. Sornette. In press. Market bubbles and crashes. In *Encyclopedia of quantitative finance*. Wiley. http://www.wiley.com//legacy/wileychi/eqf/ (long version available at http://arXiv.org/abs/0812.2449).

Koop, G. 2003. *Bayesian econometrics*. Chichester: Wiley.

Krugman, P. 1998. The confidence game, how Washington worsened Asia's crash. *The New Republic*, 5 October.

Kwiatkowski, D., P. Phillips, P. Schmidt, and Y. Shin. 1992. Testing the null hypothesis of stationary against the alternative of a unit root. *Journal of Econometrics* 54, nos. 1–3: 159–78.

Laloux, L., M. Potters, R. Cont, J.P Aguilar, and J.P. Bouchaud. 1999. Are financial crashes predictable? *Europhysics Letters* 45, no. 1: 1–5.

Lin, L., R.E. Ren, and D. Sornette. In press. A consistent model of explosive financial bubbles with mean-reversing residuals. *Quantitative Finance*. http://arxiv.org/abs/0905.0128 and http://papers.ssrn.com/abstract=1407574.

Lin, L., and D. Sornette. 2009. Diagnostics of rational expectation financial bubbles with stochastic mean-reverting termination times. ETH working paper. http://arxiv.org/abs/0911.1921.

Mogi, C. 2009. Gold hits record high as dollar sets new lows. *Reuters*, 26 November. http://www.reuters.com/article/goldMktRpt/idUSSP7486520091126.

Ng, S., and P. Perron. 2001. Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69, no. 6: 1519–54.

Onsager, L. 1944. Crystal statistics. A two-dimensional model with an order-disorder transition. *Physics Review* 65, nos. 3–4: 117–49.

Pfaff, B. 2008. *Analysis of integrated and cointegrated time series with R*. New York: Wiley.

Reinhart, C., and K. Rogoff. 2009. *This time is different: eight centuries of financial folly*. Princeton, NJ: PUP.

Shefrin. H. 2005. *A behavioral approach to asset pricing*. Academic Press Advanced Finance Series. Burlington, VT: Academic Press.

Sornette, D. 2003a. *Why stock markets crash (critical events in complex financial systems)*. Princeton, NJ: PUP.

Sornette, D. 2003b. Critical market crashes. *Physics Reports* 378, no. 1: 1–98.

Sornette, D. 2009. Dragon-kings, black swans and the prediction of crises. *International Journal of Terraspace Science and Engineering* 2, no. 1: 1–18.

Sornette, D., and A. Johansen. 1997. Large financial crashes. *Physica A* 245, nos. 3–4: 411–22.

Sornette, D., and A. Johansen. 2001. Significance of log-periodic precursors to financial crashes. *Quantitative Finance* 1, no. 4: 452–71.

Sornette, D., A. Johansen, and J.P. Bouchaud. 1996. Stock market crashes, precursors and replicas. *Journal de Physique I* 6, no. 1: 167–75.

Sornette, D., and R. Woodard. 2010. Financial bubbles, real estate bubbles, derivative bubbles, and the financial and economic crisis. In *'Proceedings of APFA7 (Applications of Physics in Financial Analysis)', New approaches to the analysis of large-scale business and economic data*, ed. Misako Takayasu, Tsutomu Watanabe, and Hideki Takayasu. Springer. http://www.thic-apfa7.com/en/htm/index.html, http://arxiv.org/abs/0905.0220, http://web.sg.ethz.ch/wps/CCSS-09–00003/, and http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1407608.

Sornette, D., R. Woodard, and W.X. Zhou. 2008. The 2006–2008 oil bubble and beyond. ETH Zurich preprint. http://arXiv.org/abs/0806.1170.

Sornette, D., R. Woodard, and W.X. Zhou. 2009. The 2006–2008 oil bubble: evidence of speculation, and prediction. *Physica A* 388, no. 8: 1571–76.

Sornette, D., and W.X. Zhou. 2006. Predictability of large future changes in major financial indices. *International Journal of Forecasting* 22, no. 1: 153–68.

Stanley, H.E. 1971. *Introduction to phase transitions and critical phenomena*. New York: OUP.

Stock, J.H. 1994. Unit roots, structural breaks and trends. In *Handbook of econometrics*, ed. R.F. Engle and D.L. McFadden, vol. 4, chap. 46, 2740–831. Amsterdam: Elsevier.

Stock, J.H., and M.W. Watson. 2002. *Introduction to econometrics*. Addison Wesley.

Tung, R. 2009. China c.Bank: Gold prices high, warns of bubble. *Reuters*, 2 December. http://www.reuters.com/article/idUSTPU00193020091202.

White, G. 2009. Golden times for precious metals but pick carefully. *The Daily Telegraph*, 8 September. http://www.telegraph.co.uk/finance/markets/questor/6156983/Golden-times-for-precious-metals-but-pick-carefully.html.

Zhou, W.X., and D. Sornette. 2003. 2000–2003 Real estate bubble in the UK but not in the USA. *Physica A* 329, nos. 1–2: 249–63.

Zhou, W.X., and D. Sornette. 2005. Testing the stability of the 2000–2003 US stock market antibubble. *Physica A* 348: 428–52.

Zhou, W.X., and D. Sornette. 2006. Is there a real-estate bubble in the US? *Physica A* 361, no. 1: 297–308.

Zhou, W.X., and D. Sornette. 2008. Analysis of the real estate market in Las Vegas: bubble, seasonal patterns, and prediction of the CSW indexes. *Physica A* 387, no. 1: 243–60.

Zhou, W.X., and D. Sornette. 2009. A case study of speculative financial bubbles in the South African stock market 2003–2006. *Physica A* 388, no. 1: 869–80.

Zhou, Xin and Alan Wheatley (2009) China's CIC wealth fund muscles up as markets recover. Reuters, Beijing, 28 August. http://www.reuters.com/article/cusir/idUSTRE57S0D420090829.