Working Paper No. 2016008

北京大學 汇丰商学院

Peking University HSBC Business School

商界军校

P H B S

Predicting the Critical Time of Financial Bubbles

Kwangwon Ahn

Jacqueline B. Dai

Domenico Tarzia

Fan Zhang

Noname manuscript No.
(will be inserted by the editor)

Predicting the critical time of financial bubbles

K. Ahn · J. Dai · D. Tarzia · F. Zhang

Received: date / Accepted: date

Abstract Detecting and predicting financial bubbles have become crucially important because of the economic significance of endogenous market crashes. In this paper, we provide an algorithm to predict the distribution of the critical times of financial bubbles by means of a log-periodic power law. Our approach consists of a price gyration algorithm, which uses different window sizes for peak detection and a distance-based weighting approach for peak selection, and a constrained genetic algorithm. Our results show a significant improvement in the prediction of bubbles' critical times. The diagnostic analysis demonstrates the accuracy, efficiency and stability of our predictions.

Keywords Critical time · Bubbles · Log-periodic power law · Econophysics

1 Introduction

Financial crises tend to follow asset price bubbles observed in various markets throughout history. Although there is no consensus on the definition of bubbles, a notion that a bubble is a large, sustained deviation of asset price from

Kwangwon Ahn
Graduate School of Future Strategy, KAIST, Daejeon 34141, Republic of Korea;
HSBC Business School, Peking University, University Town, Shenzhen 518055, China

Jacqueline B. Dai
Computational Science Center, KIST, Seoul 02792, Republic of Korea;
HSBC Business School, Peking University, University Town, Shenzhen 518055, China

Domenico Tarzia (Corresponding Author)
HSBC Business School, Peking University, University Town, Shenzhen 518055, China
E-mail: dtarzia@phbs.pku.edu.cn
Tel: +86(0)755-2603-3375

Fan Zhang
HSBC Business School, Peking University, University Town, Shenzhen 518055, China;
Shenzhen Goofar Sanxin Fund Management, Qiushi Building, Shenzhen 518000, China

its fundamental value, prevails in academic research. Skewed asset prices fail to reflect the fundamentals well, thus in turn it may have an important effect on the resource allocations (Stiglitz 1990). Moreover, the bursting of a bubble, e.g. a dramatic collapse of the stock market, may bring the economy into even worse situation, such as the great recession and dysfunction in the financial system, which proves the importance to understand the asset price bubbles.

Given the mispricing of assets underscores the weakness of present-value models, more attention is given to introducing bubbles to asset pricing models. Predominantly, two streams of theoretical frameworks shed lights on this issue: rational models as well as behavioral models, which assume irrational behavior for at least one group of agents (Sherbina and Schlusche 2012). The rational models bound all the agents to be rational, however, the bubbles may still exist due to market imperfection like information asymmetry and sale constraints (Allen et al. 1992). As a different view, the behavioral models generate bubbles by various irrationality of market players (Miller 1977; Barberis et al. 1998; Daniel et al. 1998; Shiller 2002). Given some empirical evidence which support the validation for detecting bubbles of these models, however, none of them provide the space for predicting the critical time of financial bubbles with substantial significance.

As an alternative to explain bubbles, a framework called Log-periodic power law (LPPL) model originating in statistical physics gained a lot of attention because of many successful predictions (Johansen et al. 2000; Clark 2004; Filimonov and Sornette 2011). Besides its good performance in empirical studies, the theoretical contribution cannot be neglected. Johansen et al. (2000) reconcile the rational models and the behavioral models by (i) a macro-level setting with rational expectation hypothesis and (ii) a microscopic modelling where traders imitate their nearest neighbors irrationally. Implicit in this description is that individuals may make sub-optimal choice but the aggregate effect of expectation leads to rational expected determination. The tendency of the micro-level imitation, in other words herding behavior, increases up to a certain point called critical time, through which LPPL predicts the crash dates.

The empirical literature has employed a variety of approaches to estimating LPPL models. Johansen et al. (2000) first used taboo search and the Levenberg–Marquardt algorithm (LMA) to deduce an evolution law for stock prices before the crash in the United States and Hong Kong. Johansen and Sornette (2001) identified and analyzed 21 significant bubbles followed by large crashes or severe corrections, and found that the LPPL adequately described speculative bubbles in emerging markets. Liberatore (2011) introduced a price gyration method combined with the LMA to predict the critical time of financial bubbles for the DJIA and S&P 500. Pele (2012) proposed an extension of the approach of Liberatore (2011), with added time series peak detection, and predicted the crashes of BET-FI in 2007. Kurz-Kim (2012) applied the LPPL to detect the stock market crash in Germany, which demonstrated that the LPPL is an early warning indicator for financial crashes. Geraskin and Fantazzini (2013) introduced alternative methodologies, together with diagnostic tests and graphical tools, to investigate the gold bubble in 2009. Korzeniowski and Kuropka (2013) used a GA to fit the LPPL based on time series of the DJIA and WIG20, and found it to be useful as a forecasting tool for financial crashes.

However, previous research still encounters problems regarding forecasting the critical time of financial bubbles using the LPPL model. First, prediction results are sensitive to the initial values because the nonlinear optimization algorithm based on derivatives, e.g. gradient, curvature, etc., can easily to be trapped in local minima. Second, even though many of the LPPL forecasting results are good, the estimated parameters may be outside their reasonable ranges. Third, even if we can set some constraints for the GA to solve the problems mentioned, we do not have a proper way to provide a reasonable initial population to GA. Finally, there has been neither sufficient analysis of the LPPL model nor thorough assessment of its goodness-of-fit.

In this paper, we investigate whether it is possible to predict market crashes by analyzing fluctuating financial bubbles, and whether it is possible to capture a shift over time in the log-periodic oscillations of stock prices that are associated with market crashes. Our work contributes to the existing literature by establishing an algorithm that can provide a series of reasonable and stable initial values for LPPL estimation. Our extensive approach avoids being trapped in local minima, and provides a good and robust forecast, with the imposition of constraints on LPPL parameters. The results show that our algorithm provides superior performance in regard to capturing financial bubbles. Our predictions of critical times are highly concentrated around the actual times when crashes took place. Using diagnostic analysis, we also show a relatively small and stationary residual.

The remainder of this paper is organized as follows. In Section 2, we propose a model of rational imitation in which stock prices are characterized by an LPPL evolution. We introduce the main feature of an LPPL, describe the underlying mechanism of the LPPL model, and define the main rationale behind the model. In Section 3, we explain the methodology and data that we use to fit the LPPL parameter. We implement an algorithm, which is extensively described, and highlight the distinctive key properties. In Section 4, we present the prediction results, diagnostic tests, and model comparisons of our algorithm to fit the log-price data of financial bubbles of the S&P 500 in 2000, Nikkei 225 in 1989, HSI in 2007 and SSEC in 2015. In the final section, we conclude the paper and discuss future research.

2 The LPPL Model

The LPPL model simultaneously estimates the continuation and termination of a bubble. The notion that financial crashes are manifestations of power law accelerations essentially suggests that endogenously induced stock market crashes might obey a particular power law, with log-periodic fluctuations. The basic form of the LPPL model can be written as

$$y_t = A + B\,(t_c - t)^{\beta}\left\{1 + C\cos\left[\omega \ln\left(t_c - t\right) + \phi\right]\right\}, \tag{1}$$

where $y_t > 0$ is the price, or the log of the price, at time t, $A > 0$ is the price at the critical time t_c, $B < 0$ is the increase in y_t over the time before the crash when C is close to 0, $C \in (-1, 1)$ controls the magnitude of oscillations around the exponential trend, $t_c > 0$ is the critical time, $\beta \in [0, 1]$ is the exponent of the power law growth, $\omega > 0$ is the frequency of the fluctuations during the bubble, and $\phi \in [0, 2\pi]$ is a phase parameter.

The term $B\,(t_c - t)^{\beta}$ characterizes super-exponential growth that leads to a critical time. The term $C\cos\left[\omega \ln\left(t_c - t\right) + \phi\right]$ acts as an accelerating oscillation as the critical time approaches. The residuals follow a mean-reverting Ornstein-Uhlenbeck (OU) process captured by

$$v_{t+1} - v_t = -\alpha v_t + u_t\ , \tag{2}$$

where v_t is the residual of the LPPL model, α is a positive coefficient, and u_t is Gaussian white noise.

The underlying mechanism of the LPPL model is based on rational expectations. The dynamics of the asset price before the crash are given by the following stochastic differential equation:

$$dp_t = \mu_t p_t dt - \kappa p_t dj\ , \tag{3}$$

where μ_t is the time-dependent drift, κ is the proportion by which the price is expected to decrease if a crash occurs, and j is a jump process whose value is zero before the crash and one after the crash.

We assume no arbitrage in the market so that the price process satisfies the martingale condition $E_t\left[dp_t\right] = \mu_t p_t dt - \kappa p_t h_t dt = 0$, then

$$\mu_t = \kappa h_t\ , \tag{4}$$

where h_t denotes the hazard rate at time t, which is the probability per unit of time that the crash will occur during the next unit of time, if it has not occurred yet.

Substituting Eq. (4) into Eq. (3), we obtain the differential equation before the crash given by $d\left(\ln p_t\right) = \kappa h_t dt$, whose solution is

$$\ln\left(\frac{p_t}{p_{t_0}}\right) = \kappa \int_{t_0}^{t} h_s ds\ . \tag{5}$$

Eq. (5) shows the manner in which the hazard rate is a critical component for price behavior: The higher the hazard rate, the faster the price increases before the crash.

Following Liggett (1997) and Johansen et al. (2000), we introduce a dynamic stochastic model in which each trader i (for $i = 1, \cdots, n$) can either buy (+1) or sell (−1). The current state is determined by

$$s_i = sign\left(K \sum_{j \in N(i)} s_j + \sigma \varepsilon_i\right), \tag{6}$$

where K is the tendency toward imitation (coupling strength), $N(i)$ is the set of traders who influence trader i, σ is the tendency toward idiosyncratic behavior, and ε_i is a random variable.

In Eq. (6), K and σ are critical parameters because an increase in K forces the order in the network to increase, whereas σ works in the opposite direction. There exists a critical point, $Kc = K(t_c)$, that determines the separation between the different regimes: (i) if $K < K_c$, the system is in a disordered state and the sensitivity to small global perturbations is low; (ii) as K approaches K_c, imitation forces traders to act collectively, most of the traders have the same state, and the system becomes more sensitive to small global perturbations; and (iii) when $K > K_c$, the tendency toward imitation is so intense that there exists a strong predominance of one state.

Following Johansen et al. (2000), we assume that traders are placed on a two-dimensional grid and each trader has four neighbors. The susceptibility of the system near the critical value K_c can be shown as follows:

$$X \propto (K_c - K)^{-\gamma},$$

where $\gamma > 0$ is the critical exponent of susceptibility according to a power law.

If K evolves smoothly, we can apply a first-order Taylor expansion around the critical point t_c. Then, prior to t_c, we have the following approximation:

$$K_c - K \propto (t_c - t).$$

Given that the hazard rate of the crash behaves in the same way as the susceptibility in the neighborhood of the critical point, we get

$$h_t \propto (t_c - t)^{-(1-\beta)}, \tag{7}$$

where $1-\beta > 0$ is the critical exponent of hazard rate like that of susceptibility.

Substituting Eq. (7) into Eq. (5) and integrating, we finally obtain a power law growth model

$$\ln p_t = A + B (t_c - t)^{\beta}.$$

A hierarchical diamond lattice is an appropriate structure of financial markets, and describes our model of rational imitation (Derrida et al. 1983; Johansen et al. 2000). The hierarchical structure can be interpreted as follows: (i) start with two original traders linked to each other; (ii) substitute each link with a diamond with four links and the two new vertices diagonally occupied by two new traders; and (iii) after n iterations of the second process, there are $N = \frac{2}{3}(2 + 4^n)$ traders and $L = 4^n$ links.

The lattice structure has a general solution of susceptibility given by a first-order expansion as follows:

$$X \approx B_0 (K_c - K)^{-\gamma} + B_0 C_0 (K_c - K)^{-\gamma} \cos[\omega \ln (K_c - K) + \psi].$$

Thus, we obtain the hazard rate with the approximation of a first-order expansion as follows:

$$h_t \approx B_1 (t_c - t)^{-(1-\beta)} + B_1 C_1 (t_c - t)^{-(1-\beta)} \cos[\omega \ln (t_c - t) + \phi]. \tag{8}$$

Substituting Eq. (8) into Eq. (5) and integrating provides Eq. (1), which is known as the LPPL model.

3 Methodology and Data

3.1 Fitting the LPPL Parameters

The basic form of the LPPL, given by Eq. (1), requires the estimation of seven parameters. The parameter set must be such that the root mean square error (RMSE) between the observation and predicted value of the LPPL model is minimized as follows:

$$RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(Y_t - y_t)^2}$$

where Y_t denotes the observation at time t and T the number of trading days in the dataset.

Let $f_t = (t_c - t)^\beta$ and $g_t = (t_c - t)^\beta \cos[\omega \ln(t_c - t) + \phi]$, then the LPPL model, Eq. (1), can be rewritten as

$$y_t = A + Bf_t + BCg_t.$$

It is straightforward that the linear parameters A, B and C can be estimated using ordinary least squares (OLS) given the four parameters β, ω, t_c and ϕ. Thus, determining suitable values for these four parameters is critical for fitting the LPPL model.

We fit the LPPL parameters using two steps. In the first step, we produce the initial values for the parameters with a price gyration method. In the second step, we optimize these parameters using a nonlinear optimization algorithm, GA.

Liberatore (2011) defined a price gyration method to produce the initial values of the LPPL parameters by visually inspecting stock prices as follows:

1. Identify three consecutive stock price peaks: i, j and k.
2. Estimate the initial values of t_c, ω and ϕ from price gyrations as follows:
$$t_c = \tfrac{\rho k - j}{\rho - 1},\ \omega = \tfrac{2\pi}{\ln \rho} \text{ and } \phi = \pi - \omega \ln(t_c - k) \text{ with } \rho = \tfrac{j-i}{k-j}.$$
3. Set the other initial values, i.e. $\beta = 1$ and $C = 0$.
4. Estimate the initial values of A and B using an OLS fit:
$$p_t = A + B(t_c - t) + \varepsilon_t.$$

Pele (2012) extended the approach of Liberatore (2011) using an automatic time series peak detection algorithm (Palshikar 2009) described as follows:

1. Define a peak function S_i which associates a score with element $p(i)$ and distance κ

$$\begin{aligned} S_i\left[\kappa, i, p\left(i\right)\right] = & \frac{1}{2}\left\{\max\left[p\left(i\right) - p\left(i-1\right), \cdots, p\left(i\right) - p\left(i-\kappa\right)\right]\right. \\ & \left. + \max\left[p\left(i\right) - p\left(i+1\right), \cdots, p\left(i\right) - p\left(i+\kappa\right)\right]\right\}. \end{aligned} \tag{9}$$

2. Screen the series of S_i using $S_i > 0$ and $S_i - m > hs$, where m and s are the mean and standard deviation of S_i, and h is a positive coefficient.[1]
3. Then, retain only one peak with the largest value from any set of peaks within distance κ and finally obtain the peak series.

Once the peaks are detected, price gyration might encounter following problems. The prediction results are not stable for different window sizes and the estimation of critical times is not sufficiently accurate if peaks are far from the end point of the sample. To eliminate these issues, we relax and improve the idea of a fixed window size and equally weighted peaks. Our window size κ for peak detection is no longer fixed, which allows us to test for different possibilities of a fluctuating cycle of LPPL growth. Because more recent data have made a greater contribution to forecasting, we implement a distance-based weighting (DBW) approach for peak selection. After detecting the peaks, we obtain a series of peaks. We assume that the sample size of the index time series is T, then the initial weight of each peak i is

$$w_{0,i} = \frac{1}{T-i}. \tag{10}$$

We standardize the value so that the sum of the weights of all the peaks equals 1. Then, the weight of each peak i is

$$w_i = \frac{w_{0,i}}{\sum_{j=1}^{n} w_{0,j}}, \tag{11}$$

where n is the total number of peaks.

The second cornerstone of our algorithm is the use of a GA to fit the LPPL. Compared with other nonlinear optimization algorithms, such as quasi-Newton and the LMA, a GA has many advantages. It avoids some local minima because the search for solutions runs in parallel and does not require additional information about the shape of the calculated plane. Moreover, the objective function does not need to be continuous or smooth. The GA is implemented using the following steps:

1. Each member of the initial population is a vector of the seven LPPL parameters (A, B, C, t_c, β, ω, ϕ) generated by our improved price gyration algorithm. The RMSE is calculated for each member.
2. An offspring is produced by randomly drawing two parents, without replacement, and calculating their arithmetic mean. If any parameter value is outside the constraints, it is set as the closest boundary value.

[1] h is typically set within $1 \leq h \leq 3$ (Liberatore, 2011). In this paper we choose $h = 1.5$.

3. A mutation perturbs the solutions so that new regions of the search space can be explored. The mutation process is performed by adding a perturbation variable for each coefficient in the current population. As the perturbation may drive the parameters out of the constraints, the closest boundary value will be given to these parameters as in the step 2.
4. After breeding and mutation, we merge the newly generated individuals into the population. All the solutions are ranked according to their RMSE in ascending order and only half of the best solutions can survive to the next generation.
5. We iterate this procedure and choose the best fit as the final solution.

Johansen and Sornette (2001) found that whether an LPPL model can capture crashes well depends, to some extent, on the specific bounds of the critical parameters β and ω. Based on their finding, we impose constraints on the LPPL parameters that are consistent with previous literature. Table 1 defines the constraints on the LPPL parameters.

Table 1: Restrictions on LPPL parameters

Parameter	Constraint	Literature
A	$(\max(P), +\infty)$	Korzeniowski et al. 2013
B	$(-\infty, 0)$	Lin et al. 2014
C	$(-1, 1)$	Lin et al. 2014
t_c	(t, ∞)	Korzeniowski et al. 2013
β	$[0.1, 0.9]$	Lin et al. 2014
ω	$[5, 15]$	Johansen 2002; Lin et al. 2014
ϕ	$[0, 2\pi]$	Lin et al. 2014

To summarize, the algorithm we propose implements the following steps:

1. Detect the peaks of the sample with window size κ following Palshikar (2009) and Pele (2012).
2. Assign the weight of each peak using the DBW approach.
3. Randomly select three consecutive peaks based on the weights.
4. Use these three consecutive peaks for price gyration and obtain the initial values of t_c, ω and ϕ.
5. Set the initial values $\beta = 1$ and $C = 0$, and estimate the initial values of A and B using OLS.
6. Repeat steps 3 to 5, and obtain a series of initial values for the seven LPPL parameters.
7. Find the LPPL parameters using the GA with the initial population of the parameters from step 6 and constraints.
8. Repeat steps 1 to 7, change the window size κ, and obtain the prediction interval of the critical time t_c.

Our algorithm combines the price gyration method with the GA, which extends existing research using a floating window size for peak detection and the DBW approach for peak selection.

3.2 Data

First, we need to identify financial bubbles and crashes. A financial bubble occurs when the asset price continues to increase for a long period of time beyond its fundamental value, whereas a financial crash is defined as a substantial decrease of the asset price when the bubble bursts. Stock bubbles and market crashes are identified according to Brée and Joseph (2013) as follows: (i) stock prices increase by more than 25% for a period of 252 weekdays prior to the peak; and (ii) stock prices decrease by more 25% for a period of 126 weekdays.

Based on these criteria, we choose four stock bubbles and market crashes that occurred during different time periods and in different financial markets: the dot-com bubble in the late 1990s,[2] the Japanese asset price bubble in the late 1980s,[3] the Hong Kong stock bubble of 2007,[4] and the Chinese stock bubble of 2015.[5] Table 2 summarizes the characteristics of price series for these four periods of financial bubbles. The daily closing prices of four indices are from the WIND database.

Table 2: Historic stock market bubbles and crashes

Index	Period	Observations	Critical Time
S&P 500	11/05/1994 - 17/03/2000	1479	24/03/2000
Nikkei 225	23/07/1984 - 22/12/1989	1330	29/12/1989
HSI	25/04/2003 - 23/10/2007	1123	30/10/2007
SSEC	27/06/2013 - 05/06/2015	475	12/06/2015

[2] During the dot-com bubble in the late 1990s, stock markets saw their equity value increase rapidly from growth in the Internet sector and related fields. The collapse of the bubble occurred during the period 1999–2001.

[3] In the Japanese asset price bubble in the late 1980s, real estate and stock market prices were greatly inflated. The bubble was characterized by a rapid acceleration of asset prices and overheated economic activity, in addition to an uncontrolled money supply and credit expansion. By August 1990 (the fifth monetary tightening by the Bank of Japan), the Nikkei stock index had plummeted to half of its peak price.

[4] The announcement of the "through train" scheme by SAFE caused a frenzied boom in the Hong Kong stock market in 2007. In less than 2 months, the Hang Seng Index increased from 20,000 to a peak of 31,958 at the end of October. As the Central Government of China postponed the "through train" scheme indefinitely, the Hong Kong stock market encountered a substantial daily decrease of 1,526 points on December 5, 2007.

[5] During the Chinese stock bubble of 2015, the Shanghai Stock Exchange Composite Index increased by approximately 150% within a year. Because of a series of government policies to deleverage the stock market, a third of the value of A-shares on the Shanghai Stock Exchange was lost within a month after June 2015. Subsequently, the market underwent two further crashes in August 2015 and January 2016.

After identifying the financial bubbles, we need to select carefully the time window to estimate the LPPL model. Following Johansen and Sornette (2001) and Brée and Joseph (2013), we select the forecasting window as follows: (i) the time window starts at the end of the previous crash, that is, the lowest point since the last crash; (ii) the day with the peak value of the index is the actual critical point of the financial bubble; and (iii) the endpoint is divided into four groups, which are from 1 to 4 weeks before the actual critical point.

4 Quantitative Analysis

4.1 Empirical Results

We applied our algorithm to predict the critical times of financial bubbles in the aforementioned four markets. Table 3 summarizes the forecasting results of our algorithm. The last column P_{100} denotes the percentage of predicted critical times within 100 weekdays around the actual critical time. IQR is the interquartile range of prediction interval, which was computed by subtracting the first quartile from the third quartile. According to the table, our algorithm provides a good forecast of the four stock market bubbles.

Table 3: Summary of predicted critical times

Index	End Date	95% Prediction Interval	IQR	P_{100}
S&P 500	17/03/2000	[06/07/2000-14/09/2000]	18.91	77.78%
	10/03/2000	[11/07/2000-20/09/2000]	19.05	70.21%
	03/03/2000	[19/06/2000-05/09/2000]	21.19	76.47%
	25/02/2000	[21/06/2000-25/08/2000]	15.67	94.44%
Nikkei 225	22/12/1989	[03/04/1990-15/06/1990]	16.69	89.58%
	15/12/1989	[30/03/1990-21/06/1990]	18.41	88.64%
	08/12/1989	[22/03/1990-04/06/1990]	21.43	93.88%
	01/12/1989	[23/03/1990-18/05/1990]	10.92	98.39%
HSI	23/10/2007	[17/10/2007-14/02/2008]	36.28	100%
	16/10/2007	[15/10/2007-03/03/2008]	42.17	100%
	09/10/2007	[26/10/2007-05/02/2008]	15.50	100%
	02/10/2007	[25/10/2007-05/02/2008]	17.76	100%
SSEC	05/06/2015	[05/05/2015-24/07/2015]	7.61	100%
	29/05/2015	[28/04/2015-15/07/2015]	8.10	100%
	22/05/2015	[23/04/2015-09/07/2015]	14.12	100%
	15/05/2015	[27/04/2015-07/07/2015]	18.87	100%

Note: The end date is divided into four different groups, which are from 1 to 4 weeks before the actual critical point.

We forecast a 95% prediction interval of critical times with a period of no more than 100 weekdays for all the stock market bubbles. The interquartile ranges of the predicted critical times for all the bubbles are within 50 weekdays, which means that our predictions are highly concentrated. Even though the last observation of the sample changes from 1 week to 4 weeks before the actual critical time, most of our prediction intervals demonstrate stable behavior. More than two thirds of the predicted critical times are within 100 weekdays of the actual critical time for all the bubbles. For the short-term bubbles, for example HSI and SSEC, all the predicted critical times are within 100 weekdays of the actual critical time.

In addition to forecasting the critical time, it is important to qualify the LPPL calibration, i.e. to check the stylized features of LPPL which are reflected by the restrictions on the parameters mentioned in Table 1. Table 4 shows the best LPPL fits for the four financial bubbles estimated by our algorithm. Since we use the GA with constraints, all the estimated parameters are within the boundaries. The two conditions $B < 0$ and $0.1 \leq \beta \leq 0.9$ ensure a faster-than-exponential acceleration of the log-price with a vertical slope at the critical time t_c (Lin et al. 2014). The positive hazard rate always holds because the absolute value of C is restricted in one unit. The values of ω are close to the lower bound 5, which corroborates the existing studies such as Johansen (2002), who found that $\omega \approx 6.36 \pm 1.56$ for 30 crashes on major financial markets.

Table 4: LPPL parameters of best fit

Index	A	B	C	β	ω	t_c	ϕ	$RMSE$
S&P 500	7.59	−0.008	−0.061	0.714	5.11	1566	4.318	0.055
Nikkei 225	10.79	−0.009	−0.115	0.700	5.05	1389	5.191	0.059
HSI	10.20	−0.013	0.074	0.609	5.12	1126	0.601	0.050
SSEC	10.73	−1.660	0.034	0.109	5.00	479	4.574	0.050

4.2 Diagnostic Tests

To reduce the possibility of false alarms, it was necessary to conduct diagnostic analysis to demonstrate our predictions. We conducted the diagnostic analysis by considering the relative errors, unit root test of LPPL residuals, sensitivity analysis of LPPL parameters and crash lock-in plot (CLIP) analysis.

The relative error analysis of the best fits shows that the bubbles are well-captured by our model. In the analysis of Johansen et al. (2000), most of the relative errors of their best fits were below 5%. Our algorithm significantly improves their result, demonstrating more accurate performance in capturing the four financial bubbles. The significance of our findings is quite evident and immediate in Figure 1, where the relative errors of all the fitting points of these four indices are well below 3%.

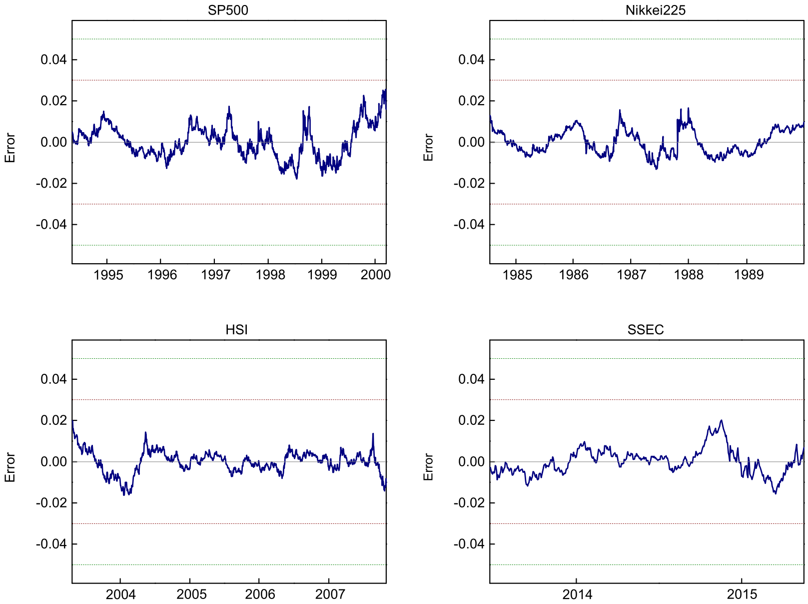

Fig. 1: Relative error analysis of the best LPPL fits

One key property of the LPPL model is that the residuals follow a mean-reverting OU process, Eq. (2). Table 5 shows the results of unit root tests. Both ADF and PP tests with two lags for the best fit of the four indices reject the null hypothesis H_0 at a 1% significance level, which means that the residuals do not have a unit root but are stationary and thus compatible with a mean-reverting OU process.

Table 5: Unit-root test for LPPL residuals

Index	ADF	PP
S&P 500	-3.117^{**}	-3.176^{**}
Nikkei 225	-2.921^{**}	-3.108^{**}
HSI	-3.532^{***}	-3.962^{***}
SSEC	-2.610^{**}	-2.658^{**}

We also investigated how sensitive the RMSE is to variations in the LPPL parameters. Because A, B and C are always estimated given the four parameters β, ω, t_c and ϕ, we examined the sensitivity of the LPPL fit to variations in these four parameters. We let t_c and one parameter among β, ω and ϕ vary for the S&P 500, while the remaining parameters were fixed. The results in Figure 2 show that the variation of the RMSE is relatively smooth with respect to β, t_c and ϕ, and that of the RMSE is highly sensitive to small fluctuations in ω. Figure 2 shows significant evidence supporting the choice of a GA instead of the LMA to fit the LPPL model. If the initial value of ω is close to a local minimum, the searching algorithm of the LMA can be easily trapped because it achieves a local optimal solution. Compared with the LMA, the search of a GA for solutions runs in parallel, which means that even when a local minimum has been found, small variations in the parameters might avoid the search procedure becoming trapped.

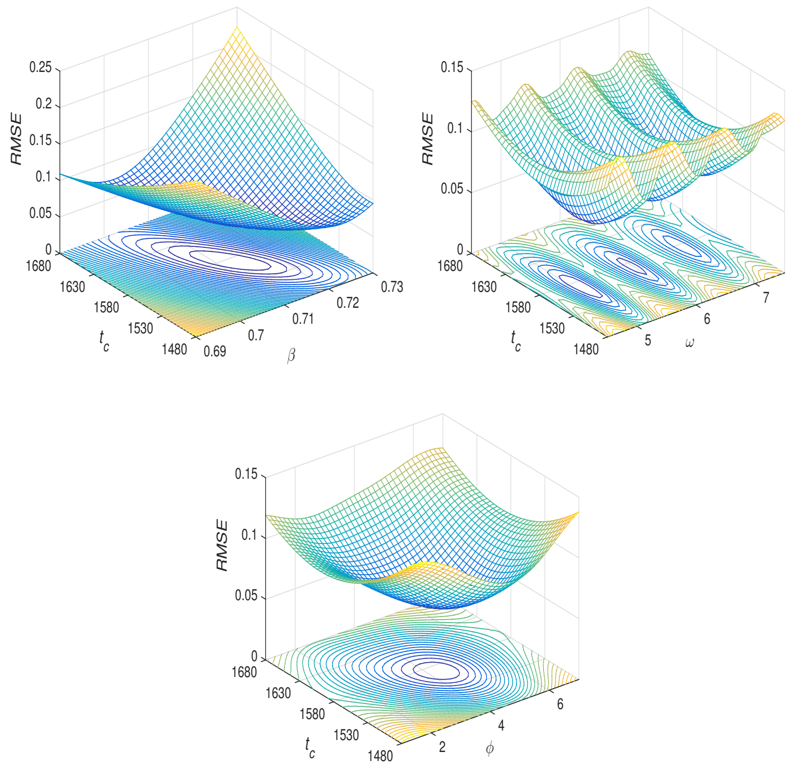

Fig. 2: Sensitivity of the RMSE to LPPL parameters for S&P 500

Finally, we analyze CLIP following Fantazzini (2010). A CLIP is a useful tool for tracking the development of a bubble and understanding whether a possible crash is imminent. The main idea of a CLIP is to plot the date of the last observation in the estimation sample on the horizontal axis and the estimated crash date t_c on the vertical axis. If a regime change in the stock market is approaching, then the estimated t_c should be stabilized around a constant value close to the critical time. To implement a CLIP, we continued changing the last observation of our estimation sample from 1 week to 8 weeks before the actual critical time and made predictions for the four financial bubbles. From Figure 3, we can see that our predicted results for t_c are stable, especially in the last 4 weeks before the crash. Moreover, the predicted t_c are also very close to the actual critical time. These results indicate that when the crash is imminent, our algorithm provides a robust and precise forecast of the critical time.

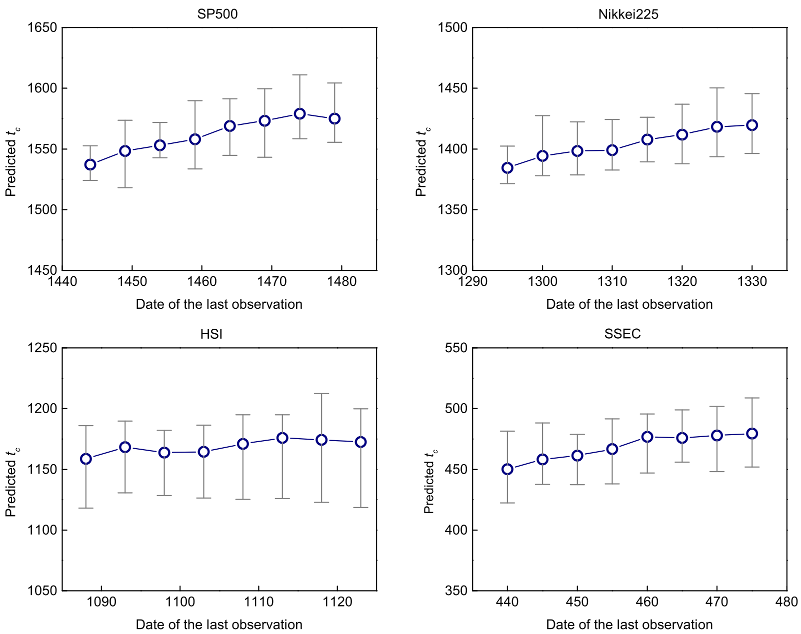

Fig. 3: Crash lock-in plots with rolling estimation windows

4.3 Model Comparison

As a final step, we compared the prediction accuracy of critical time t_c for the following two models: (i) price gyration algorithm for searching initial values and the LMA for optimization, M1; and (ii) the improved price gyration algorithm, which is applied with different window sizes and a DBW approach, and a GA with constraints for optimization, M2.

For both methods, we forecast the critical time t_c using different last observations of the estimation sample from 1 week to 4 weeks before the actual critical time and merge the results of four groups into one sample. We computed the 95% prediction interval of t_c, IQR and P_{100}. The prediction results of t_c are shown in Figure 4, in which we immediately observe that M1 cannot predict critical times particularly well because most of the predicted t_c occur somewhat earlier than the actual critical time. The 95% prediction intervals of t_c are too large for forecasting. Except SSEC, all the other indices have less than one-third of the predicted t_c within 100 weekdays of the actual critical time, which means that this approach is unable to predict the crash effectively. M2 predicts a relatively accurate t_c compared with M1. Most of the predicted t_c are close to each other and form a 95% prediction interval of t_c of less than 100 weekdays in addition to an interquartile range of less than 30 weekdays, which means that our predicted results are highly concentrated. In addition, more than 80% of the predicted t_c are within 100 weekdays of the actual critical time for all four financial bubbles, which also demonstrates that our method provides a better forecast of the critical time.

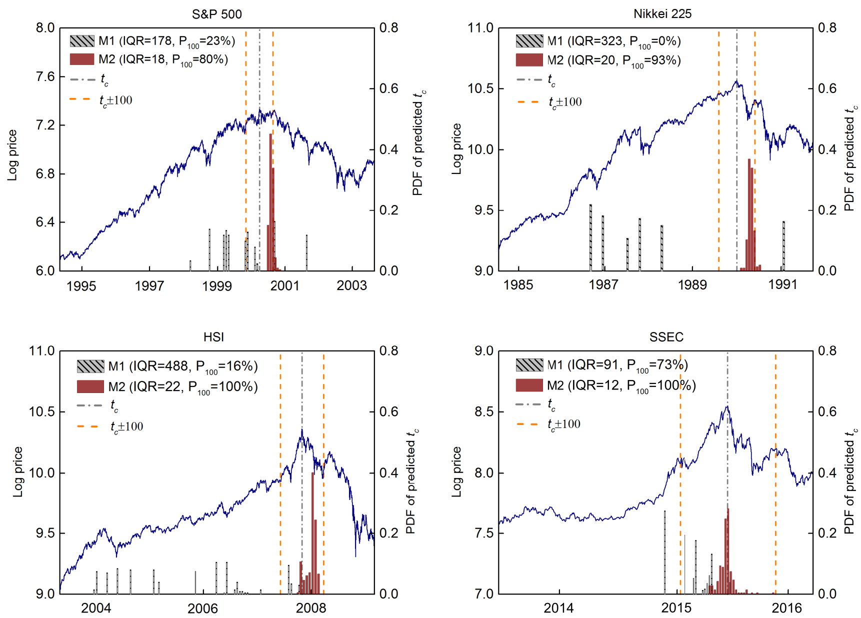

Fig. 4: Logarithm of the stock price and corresponding alarms

5 Conclusion

This paper provides an algorithm to predict the critical time of financial bubbles with an LPPL model. The parameters are estimated by minimizing the cost function by means of a nonlinear optimization method. This algorithm consists of two steps, (i) a price gyration method to generate an initial candidate of parameters and (ii) a genetic algorithm to find the optimal solution. Specifically, we go beyond the price gyration method in the previous literature. In our case, different window sizes are applied to peak detection since the fixed window size may omit the possible variation in cycle of LPPL growth. Given the peaks detected, we use a DBW method to assign the weights on each peak according to its distance to the crash day. The DBW method makes the estimation accord with reality, i.e. the recent data have more influence on forecasting.

For validation, we performed an ex-ante prediction on the time of crashes on four stock market indices. The critical time of the bubbles, when the crashes may happen with significant probability, is one of the parameters in the LPPL model. Our predictions on critical times are highly concentrated around the actual time. Moreover, diagnostic analysis demonstrates our results in different aspects. First, we generate a smaller prediction error with large randomness. Second, our prediction is stable with respect to moving the termination time of the observation period. Third, as for the degree of concentration and accuracy, we present a more significant improvement than an existing algorithm.

Our work focuses on bubbles in stock price, and it can be extended to other assets without loss of generalization, such as bubbles in real estate and credit assets. Also, the LPPL model mainly focuses on the time period prior to crashes. However it leaves blank the price behavior in the post-crash periods when mispricing may still exist, which is of policy significance to stabilize and boost the economy once crashes occur. In addition, since parameters in LPPL characterize stock markets, they have a large potential to be related to economic fundamentals. Thus as future research, additional analysis on whether and how the LPPL model reflects macroeconomic conditions is of interest.

References

1. Allen, F., Morris, S., & Postlewaite, A. (1992). Finite bubbles with short sale constraints and asymmetric information. *Journal of Economic Theory*, *61*(2), 206-229.
2. Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics*, *49*(3), 307-343.
3. Bhattacharya, U., Galpin, N., Ray, R., & Yu, X. (2009). The role of the media in the internet IPO bubble. *Journal of Financial and Quantitative Analysis*, *44*(03), 657-682.
4. Brée, D. S., & Joseph, N. L. (2013). Testing for financial crashes using the Log Periodic Power Law model. *International Review of Financial Analysis*, *30*, 287-297.
5. Clark, A. (2004). Evidence of log-periodicity in corporate bond spreads. Physica A Statistical Mechanics & Its Applications, *338*(3), 585-595.
6. Daniel, K., Hirshleifer, D., & Subrahmanyam, A. (1998). Investor psychology and security market under-and overreactions. *Journal of Finance*, *53*(6), 1839-1885.

7. Derrida, B., De Seze, L., & Itzykson, C. (1983). Fractal structure of zeros in hierarchical models. *Journal of Statistical Physics*, *33*(3), 559-569.
8. Fantazzini, D. (2010). Modelling bubbles and anti-bubbles in bear markets: a medium-term trading analysis. *HANDBOOK OF TRADING*, McGraw-Hill Finance and Investing, 365-388.
9. Filimonov, V., & Sornette, D. (2011). A stable and robust calibration scheme of the log-periodic power law model. Physica A Statistical Mechanics & Its Applications, *392*(17), 36983707.
10. Geraskin, P., & Fantazzini, D. (2013). Everything you always wanted to know about log-periodic power laws for bubble modeling but were afraid to ask. *European Journal of Finance*, *19*(5), 366-391.
11. Glaeser, E. L., Gyourko, J., & Saiz, A. (2008). Housing supply and housing bubbles. *Journal of Urban Economics*, *64*(2), 198-217.
12. Johansen, A., Ledoit, O., & Sornette, D. (2000). Crashes as critical points. *International Journal of Theoretical and Applied Finance*, *3*(02), 219-255.
13. Johansen, A., & Sornette, D. (2001). Bubbles and anti-bubbles in Latin-American, Asian and Western stock markets: An empirical study. *International Journal of Theoretical and Applied Finance*, *4*(06), 853-920.
14. Johansen, A. (2002). Characterization of large price variations in financial markets. *Physica A: Statistical Mechanics & Its Applications*, *324*(12), 157-166.
15. Korzeniowski, P., & Kuropka, I. (2013). Forecasting the critical points of stock markets indices using log-periodic power law. *Ekonometria*, *1*, 100-110.
16. Kurz-Kim, J. R. (2012). Early warning indicator for financial crashes using the log periodic power law. *Applied Economics Letters*, *19*(15), 1465-1469.
17. Liberatore, V. (2011). Computational LPPL fit to financial bubbles. arXiv preprint arXiv:1003.2920.
18. Liggett, T. M. (1997). Stochastic models of interacting systems. *Annals of Probability*, *25*(1), 1-29.
19. Lin, L., Ren, R. E., & Sornette, D. (2014). The volatility-confined LPPL model: A consistent model of explosive financial bubbles with mean-reverting residuals. *International Review of Financial Analysis*, *33*, 210-225.
20. Miller, E. (1977). Risk, uncertainty, and divergence of opinion. *Journal of Finance*, *32*, 11511168.
21. Palshikar, G. (2009). Simple algorithms for peak detection in time-series. In *Proc. 1st Int. Conf. Advanced Data Analysis, Business Analytics and Intelligence*.
22. Pele, D. T. (2012). An LPPL algorithm for estimating the critical time of a stock market bubble. *Journal of Social and Economic Statistics*, *1*(2), 14-22.
23. Scherbina, A., & Schlusche, B. (2014). Asset price bubbles: a survey. *Quantitative Finance*, *14*(4), 589-604.
24. Shiller, R. (2002). Bubbles, Human Judgment, and Expert Opinion. *Financial Analysts Journal*, *58*(3), 18-26.
25. Stiglitz, J. E. (1990). Symposium on bubbles. *Journal of Economic Perspectives*, *4*(2), 13-18.