Journal of Social and Economic Statistics

No.2, Vol.1, Winter 2012

AN LPPL ALGORITHM FOR ESTIMATING THE CRITICAL TIME OF A STOCK MARKET BUBBLE[*]

Daniel Traian PELE[a]

Abstract

LPPL models have been widely used to describe the behaviour of stock prices during an endogenous bubble and to predict the most probable time of the regime switching. Although their utility has been proved in many papers, there is still a lack of consensus on the statistical robustness, as the estimators are obtained through a nonlinear optimization algorithm and they are sensitive to the initial values. In this paper we propose an extension of the approach from Liberatore (2011), using a time series peak detection algorithm.

Keywords: LPPL, stock market crash, speculative bubble.
JEL Codes: G - Financial Economics, G01 - Financial Crises.

Author's Affiliation

[a] PhD Lecturer, Department of Statistics and Econometrics, Bucharest University of Economic Studies, danpele@ase.ro.

[*] An earlier version of this paper was presented at *The 6th International Conference on Applied Statistics*, November 2012, Bucharest.

1. Introduction

The behaviour of stock market price during an endogenous bubble is a subject widely debated in the literature, especially in the past years of the financial crisis.

Johansen et al. (2000) compares seismic activity to the evolution of speculative bubbles, and deduces the evolution law for stock prices before and during the crash, which is seen as a critical time. Thus, the trading price before the crash follows a log-periodic power law(LPPL):

$$\ln p(t) \approx A + B(t_c - t)^{\beta}\{1 + C\cos[\omega \ln(t_c - t)^{\beta} + \phi]\}, \quad (1)$$

where *p(t)* is the price at moment *t*, t_c is the critical time (the most probable moment of the crash), and $\beta, B_0, B_1, \omega, \phi$ are the parameters of the model which give its log-periodic feature.

In order to have a proper specification of the model, there are several constrains applied to the parameters:

- $A>0$ - usually this the price at the critical time t_c;
- $B<0$;
- $C \neq 0, |C|<1$ – this parameter controls the magnitude of oscillations around the exponential trend;
- $0 < \beta < 1$ – controls the growth rate of the magnitude and is the most important feature capturing the imminence of a regime switching, as his value is close to zero;
- $\omega \in (0, \infty)$ - controls for the amplitude of oscillations;
- $\phi \in [0, 2\pi]$ – a phase parameter.

An alternative form of the model (1) uses raw data for stock prices instead of log-prices:

$$p(t) \approx A + B(t_c - t)^{\beta}\{1 + C\cos[\omega \ln(t_c - t)^{\beta} + \phi]\}. \quad (2)$$

Johansen, Ledoit and Sornette (2000) have applied these models to successfully predict famous crashes like the one in October 1987 and for the Brazilian market, Cajueiro, Tabak and Werneck (2009) have applied these models to predict the catastrophic behaviour of the price series of 21 stocks. The Financial Crisis Observatory (ETH – Zurich) has released during the past few years predictions about the bubble behaviour of different assets and they have succeeded to predict two famous events of this type: Oil Bubble – 2008 and Chinese Index Bubble – 2009.

Fantazzini and Geraskin (2011) provide an extensive review of theoretical background behind the LPPL models, estimation methods and various applications, pointing out that although the literature on this subject is heterogeneous, LPPL fit for asset bubbles could be a useful tool in predicting the catastrophic behaviour of capital markets as a whole.

Moreover, even using such a model, the prediction of critical time is not very accurate, Kurz-Kim (2012) shows that LPPL models could be used as a early warning mechanism of regime switching in case of a stock market.

2. Fitting LPPL models

Fitting log-period power laws is not straightforward (Fantazzini and Geraskin (2011)); there are several methods used to estimate the parameters of the model (1) or model (2):

- 2-step Nonlinear Optimization(Johansen et al. (2000)), using in the first step a Taboo search to find the initial values of the parameters and in the second step a Levenberg-Marquardt nonlinear least squares algorithm;
- Genetic Algorithms(Jacobsson (2009));
- The 2-step/3-step ML (Fantazzini (2010)), using a maximum likelihood approach to estimate the parameters of an anti-bubble.

The 2-step Nonlinear Optimization method consists in the following steps:

- Reparametrization of the model (1) or (2) as $y_t = A + Bf_t + Cg_t$, where $y_t = \ln p(t)$ or $y_t = p(t)$, $f_t = (t_c - t)^{\beta}$ and $g_t = (t_c - t)^{\beta} \cos[\omega \ln(t_c - t)^{\beta} + \phi]$.

The linear parameters A, B, C are estimated using OLS from the system below:

$$\begin{pmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} y_t f_t \\ \sum_{t=1}^{T} y_t g_t \end{pmatrix} = \begin{pmatrix} T & \sum_{t=1}^{T} f_t & \sum_{t=1}^{T} g_t \\ \sum_{t=1}^{T} f_t & \sum_{t=1}^{T} f_t^2 & \sum_{t=1}^{T} f_t g_t \\ \sum_{t=1}^{T} g_t & \sum_{t=1}^{T} f_t g_t & \sum_{t=1}^{T} g_t^2 \end{pmatrix} \begin{pmatrix} A \\ B \\ C \end{pmatrix} . \quad (3)$$

If $b = \begin{pmatrix} A \\ B \\ C \end{pmatrix}$, $X = \begin{pmatrix} 1 & f_1 & g_1 \\ \ldots & \ldots & \ldots \\ 1 & f_T & g_T \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix}$, then the estimators are

$$\hat{b} = (X'X)^{-1} X'y \quad . \quad (4)$$

Thus we obtain 3 free parameters(*A*, *B*, *C*), assuming that all the others have been estimated.

- Taboo search algorithm requires an initial estimation of 10 potential solutions such as $B < 0$, $0 < \beta < 1$ and $t_c > T$; other methods, like grid search, could be employed in this stage(Fantazzini and Geraskin (2011)).
- The 10th initial solutions found at the previous stage are used as initial valued for a nonlinear Levenberg-Marquardt algorithm, the critical time t_c being selected as the one that minimizes RMSE.

Liberatore (2011) propose an estimation method based on the same Levenberg-Marquardt nonlinear least squares algorithm, but the initial values of the algorithm are based on the peaks of the stock price time series.

Basically, the algorithm involves the following steps for estimating the parameters of the model (2):

- Identify three consecutive price peaks i, j, k (either visually or using an automatic peak detection algorithm).
- Estimate the initial values of t_c, ω and φ from price gyrations: $t_c = (\rho k - j)/(\rho - 1)$, $\omega = 2\pi / \log(\rho)$ and $\phi = \pi - \omega \log(t_c - k)$, where $\rho = (j - i)/(k - j)$.
- Estimate the initial values of A and B using an OLS fit: $p(t) = A + B(t_c - t) + \varepsilon_t$.
- Estimate the other parameters of the model using the Levenberg-Marquardt nonlinear least squares algorithm.

Palshikar(2009) proposes an algorithm for peak detection, using the following function to compute the average of the maximum among the signed distances of p_t from its *k* left neighbours and the maximum among the signed distances of p_t from its *k* right neighbours:

$$S_1[k, i, p(i)] = \frac{\max\{p(i) - p(i-1), ..., p(i) - p(i-k)\} + \max\{p(i) - p(i+1), ..., p(i) - p(i+k)\}}{2}.$$

procedure peaks(T,k,h)
input *T = p(1),...,p(n), N* // input time-series of *N* points
input *k* // window size around the peak
input *h* // typically 1 ≤ *h* ≤ 3
output *O* // set of peaks detected in *T*
begin
O = ∅ // initially empty
for (i = 1; i < n; i++) **do**
a[i] = $S_1[k, i, p(i)]$ // compute peak function value for each of the *N* points in *T*
end for
Compute the mean *m′* and standard deviation *s′* of all positive values in array a;

```
for (i = 1; i < n; i++) do // remove local peaks which are "small" in global context
if (a[i] > 0 && (a[i] – m′) >( h * s′)) then O = O ∪ {p(i)}; end if
end for
Order peaks in O in terms of increasing index in T
// retain only one peak out of any set of peaks within distance k of each other
for every adjacent pair of peaks p(j) and p(j) in O do
if |j – i| ≤ k then remove the smaller value of {p(i),p(j)} from O end if
end for
```

In order to insure a proper specification of the model and to control for sensitivity of estimates to the initial values, a combined approach is proposed in this paper, using both price gyrations and peak detection algorithm.

```
procedure rolling_peaks_price_gyrations(k,h,w)
input  w // time window
input p//local time step
    procedure peaks(T,k,h)
        input T = p(1),…,p(n), N // input time-series of N points
        input k // window size around the peak
        input h // typically 1 ≤ h ≤ 3
        output O // set of peaks detected in T
        begin
        O = ∅ // initially empty
        for (i = 1; i < n; i++) do
        a[i] = S_1[k,i,p(i)] // compute peak function value for each of the N points in T
        end for
        Compute the mean m′ and standard deviation s′ of all positive values in array a;
        for (i = 1; i < n; i++) do // remove local peaks which are "small" in global context
        if (a[i] > 0 && (a[i] – m′) >( h * s′)) then O = O ∪ {p(i)}; end if
        end for
        Order peaks in O in terms of increasing index in T
        // retain only one peak out of any set of peaks within distance k of each other
        for every adjacent pair of peaks p(j) and p(j) in O do
        if |j – i| ≤ k then remove the smaller value of {p(i),p(j)} from O end if
        end for
    end peaks

    for m=1 to n-p do
    input T(w)= p(1),…,p(w+m+p), N // input time-series of w+m+p stock prices
    call peaks(T(w),3,1.5)
    u=rand('Uniform')//uniform random number in [0,1]
    nn=int(u*#(O)) // the first peak selected randomly
    i=nn // peak i
    j=nn+1 // peak j
    k=nn+2 // peak k
    Estimate the initial values of t_c, ω and φ from price gyrations: t_c = (ρk − j)/(ρ − 1),
    ω = 2π/log(ρ) and φ = π − ωlog(t_c − k), where ρ = (j − i)/(k − j).
```

Estimate the initial values of A and B using an OLS fit: $p(t) = A + B(t_c - t) + \varepsilon_t$.

Estimate $p(t) \approx A + B(t_c - t)^{\beta} \{1 + C\cos[\omega \ln(t_c - t)^{\beta} + \phi]\}$ using LMA, with initial values for t_c , ω, φ, A and B, estimated in the previous steps.

end for

end rolling_peaks_price_gyrations

The proposed method uses random consecutive peaks in order to estimate the initial values of the parameters t_c , ω and φ, diminishing the likelihood of a subjective choice.

Moreover, applying this procedure in an iterative way, with different starting values for Levenberg-Marquardt algorithm one can obtain the distribution of the critical time for an on-going bubble.

3. Numerical results for BET-FI Index of Bucharest Stock Exchange

The above algorithm was applied to the time series of BET-FI Index of Bucharest Stock Exchange, for the period 3.01.2001 – 23.12.2008(1978 daily observations).

Fig.1. Closing price of the BET-FI Index

Actually, from 2001 to 2007, the BET-FI index exhibits a near exponential behaviour, reaching its historical maximum on 25th July 2007.

Atfer this point, the evolution of the index followed a descending trend, until the turbulent period from October 2008, when the daily return was lower than -10% for several days and the local mimimum value of the index was reached on 27.10.2008.

In order to estimate the most probable time of the regime switching, the algorithm using price gyrations and peak detection was applied for the raw BET-FI Index time series.

The initial sample for fitting LPPL model in the case of BET-FI index for predicting the phase transition from January 2008 was 31.10.2000 – 28.06.2007 (1640 daily observations); starting from the last observation in the initial sample, we extended the sample using a rolling window with fixed lower limit, so we estimated at every step the LPPL model for $t \in [1,T+k]$, $k=1\ldots30$:

$$p(t) = A_k + B_k(t_c - t)^{\beta_k}\{1 + C_k \cos[\omega_k \ln(t_c - t)^{\beta_k} + \phi_k]\}\,. \tag{3}$$

The algorithm was applied for 100 iterations, obtaining an empirical distribution of the estimated critical time.

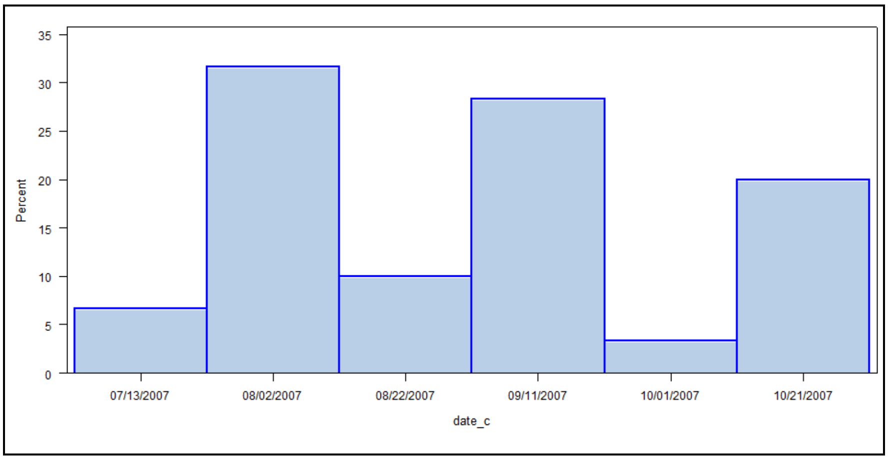

Fig.2. The empirical distribution of the estimated critical time

The mean of this distribution corresponds to the 30th of august 2007, while 5% and 95% quantiles are 2.08.2007 and 26.10.2007.

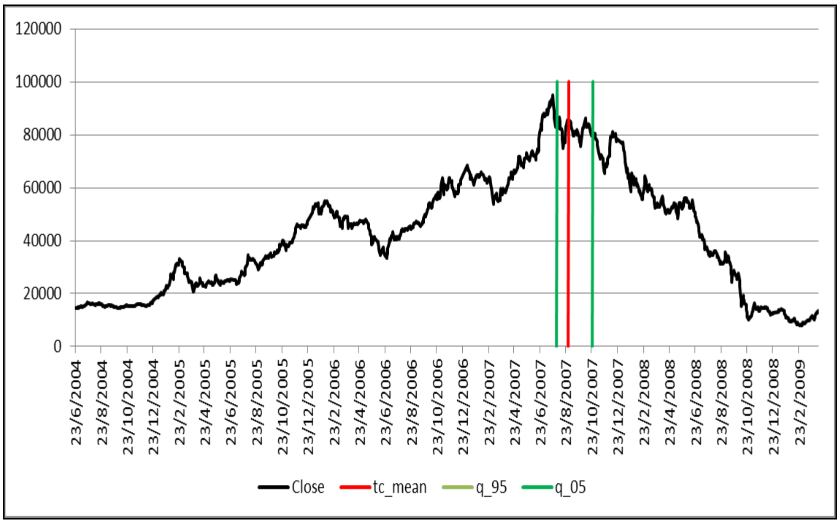

Fig. 3. The critical time for BET-FI Index

The distribution of the critical time could be useful for estimating a confidence interval for the most probable time of the expected regime switching. According to this, the most probable time for the phase transition is 30th of August 2007, which is a reasonable approximation of the reality.

Conclusions

In order to insure a proper specification of the LPPL model and to control for sensitivity of estimates to the initial values, a combined approach is proposed in this paper, using both price gyrations and peak detection algorithm.

The results obtained for the BET-FI Index shows that this method could be useful in estimating the most probable time of the regime switching for an endogenous stock market bubble.

A recommendation arising from these results is to implement an iterative estimation method for LPPL models, allowing to asses periodically the probability of a phase transition in the stock market. The research in this direction needs to be improved, in order to define a clear, standardized method to recognize an ongoing stock market bubble.

Acknowledgements

This paper was cofinanced from the European Social Fund, through the Human Resources Development Operational Sectorial Program 2007-2013, project number POSDRU/89/1.5/S/59184, „Performance and excellence in the post-doctoral economic research in Romania”.

References

Cajueiro, D. O., B. M. Tabak, & F. K. Werneck (2009). Can We Predict Crashes? The Case of the Brazilian Stock Market, *Physica A*, 388(8), 1603-1609.

Fantazzini, D. &Geraskin, P.(2011). Everything You Always Wanted to Know About Log Periodic Power Laws for Bubble Modelling But Were Afraid to Ask, *European Journal* *of Finance*, available at http://www.tandfonline.com/doi/abs/10.1080/1351847X.2011.601657.

Fantazzini, D. (2010). Modelling Bubbles And Anti-Bubbles In Bear Markets: A Medium-Term Trading Analysis. In *Handbook of Trading*, ed. G. Gregoriou, 365-388. New York: McGraw-Hill.

Jacobsson, E. (2009). How to predict crashes in financial markets with the Log-Periodic Power Law, Master diss., Department of Mathematical Statistics, Stockholm University.

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

Kurz-Kim, J.R. (2012). Early warning indicator for financial crashes using the log periodic power law, *Applied Economics Letters*, 19:15, 1465-1469.

Liberatore, V. (2011). Computational LPPL Fit to Financial Bubbles, available at http://arxiv.org/abs/1003.2920.

Palshikar, G.K.(2009). Simple Algorithms for Peak Detection in Time-Series, *Proc. 1st Int. Conf. Advanced Data Analysis, Business Analytics and Intelligence*, 2009. [Online]. Available https://sites.google.com/site/girishpalshikar/Home.

Pele, D.T.(2012). LPPL fit to stock market bubbles, *Proceedings of the VIth International Conference on Applied Statistics*, Bucharest, 2012, ISSN 2069-2498.