A statistical model of speculative bubbles, with applications to the stock markets of the United States, Japan, and China

Kazumi Asako\textsuperscript{a}, Zhentao Liu\textsuperscript{b,*}

\textsuperscript{a}Institute of Economic Research, Hitotsubashi University, 2-1 Naka, Kunitachi, Tokyo 186-8601, Japan
\textsuperscript{b}Institute for Financial and Accounting Studies, Xiamen University, Xiamen 361005, China

\textbf{Article info}

\textbf{Article history:}
Received 15 December 2011
Accepted 12 February 2013
Available online 21 February 2013

\textbf{JEL classification:}
G15
G01
C11
C51

\textbf{Keywords:}
Financial market
Speculative bubble
Bayesian recursive estimation
Crash

\textbf{ABSTRACT}

It is common knowledge that the more prices deviate from fundamentals, the more likely it is for prices to reverse. Taking this into account, we propose a simple statistical model to identify speculative bubbles in financial markets. Through the estimates of the time varying parameters, including transition probabilities, we can identify when and how newly born bubbles grow and burst over time. The model can be estimated by recursive computations, which require a huge storage capacity for standard computers. For this reason, we introduce an approximation in the computation, maintaining the recursive nature of our estimation technique. We then apply this model to the stock markets of the United States, Japan, and China, estimate its parameters and the probabilities of a bubble crash, and obtain several interesting results: the time series data of the stock price bubble show an inherently non-stationary development and the probability of a bubble crash indeed increases as the stock price becomes too high or too low.

\textcopyright 2013 Elsevier B.V. All rights reserved.

\section{1. Introduction}

Asset price bubbles have fascinated researchers in finance. Identifying speculative bubbles in, for instance, stock and land markets has been the subject of an increasing number of both theoretical and empirical papers. On the theoretical ground, it has been conjectured that such price paths reflect irrational behavior of the economic agents and, therefore, should be ruled out from a market with truly rational economic agents (Burmeister, 1980; Cass and Shell, 1983; Tirole, 1985; Diba and Grossman, 1988). Both rational theory and behavioral theory have been trying to explain why bubbles occur (Allen and Gale, 2000; Abreu and Brunermeier, 2003; Scheinkman and Xiong, 2003; Hommes, 2006). Gurkaynak (2008) surveyed the econometric tests for rational price bubbles and found that the empirical results are mixed: for each paper that finds evidence of bubbles, there are others that fit the data equally well without allowing for a bubble. Recently, many researchers have turned to the paradigm of behavioral finance (Hommes et al., 2004; Shiller, 2003). However, Brav and Heaton (2002) consider it difficult to distinguish between “… behavioral theories built on investor irrationality and rational structural uncertainty theories built on incomplete information about the structure of the economic environment.”

From the empirical point of view, on the other hand, the main interests have been to simply check whether the price series in question is better explained by assuming speculative bubbles than otherwise. Accordingly, various statistical approaches have been mainly concerned with tests for the existence and duration of speculative bubbles. They are based in one way or another on prima facie anomalies caused by the divergence or non-stationarity of prices (see, for instance, Flood and Garber, 1980; Blanchard and Watson, 1982; West, 1987). Phillips et al. (2011) and Phillips and Yu (2011) use forward recursive regression period by period to assess the evidence for unit root behavior against mildly explosive alternatives. Gutierrez (2011) proposes a method based on a bootstrap methodology to compute the finite sample probability distribution of the asymptotic tests proposed by Phillips et al. (2011). Homm and Breitung (2012) took the procedure of Phillips et al. (2011) as the benchmark and compared other test procedures of Bhargava (1986), Busetti and Taylor (2004), Kim (2000), and Kim et al. (2002) for the detection of rational bubbles. Their simulations show that the estimator derived from the Chow test is the most
reliable and powerful with finite samples. Cerqueti and Constantini (2011) analyzed international data covering 18 OECD countries using a panel unit root and cointegration methodology and they found there are bubbles in the period from 1992 to 2010. Al-Anaswah and Willing (2011) applied state-space models with Markov switching to simulated data and real international dataset.

Needless to say, a speculative bubble is characterized by a divergent sequence of a price in its early stage of development, followed by a sudden crash in due course. When a bubble grows, the time path of prices has a divergence and, when it crashes, there occurs a discrete jump in the time series data. This feature, although important in the definition of bubbles, has hardly been the subject of formal statistical analyses. McQueen and Thorley (1994) use the statistical theory of duration dependence to model speculative bubbles and examine the monthly returns of the NYSE portfolio for evidence of speculative bubbles. The results of Chan et al. (1998) and Harman and Zuehlke (2004) show that the model of McQueen and Thorley (1994) is sensitive to the data specifications. Anderson et al. (2010) present evidence of speculative bubbles in the sector indices of the S&P 500 and bubble spillovers across sectors using a regime-switching approach.

Harrison and Stevens (1971, 1976) proposed a statistical model called the dynamic linear model (DLM) to describe sudden changes in trend, slope, and transients. They also presented the multi-state model, which combines several DLMs with probability weights. The model proposed in the present paper can be regarded as an extension of their model towards a framework in which the probability of the choice of each model depends explicitly on the historical state of the price series of concern. While Diggle and Zeger (1989) proposed a similar model that incorporated the dependence of the probability on the state history and applied the model to endocrinological data in medicine, our model has a further complication in that it includes time dependent parameters whose transition is our prime concern. Accordingly, a simple likelihood approach to estimate the parameters is not possible and, instead, we must resort to a Bayesian recursive calculation.

In our analyses, we shall consider the recursive formula for the probability of a bubble crash as well. In fact, as will be seen in the sequel, most of the statistical complications encountered in the estimation process stem from the introduction of this state dependent probability of a crash. Even if we assume a relatively simple probability structure for the possibility of a bubble crash, the resulting recursive formula requires a vast amount of computational capacity. From a practical viewpoint, therefore, we should employ a certain approximation for the exact recursive estimation technique to reduce the computational burden.

The rest of this paper consists of four sections. We describe the model of speculative bubbles and discuss its implications in Section 2. Section 3 develops the recursive estimation technique of the model. It turns out that the exact recursive method requires considerable computational power, which easily exceeds the capacity of standard computers even for a moderate length of sample data. To overcome this difficulty, an approximation method is proposed. Section 4, then, applies our model to examine the speculative bubbles in the stock markets of the United States, Japan, and China. Section 5 concludes our discussion.

2. Description of the model

The model of speculative bubbles we consider in this paper is a stochastic one and is given by

\[ x_t = \begin{cases} 
\beta_1 x_{t-1} + u_t & \text{(A) with probability } \pi_t, \\
\mu & \text{(B) with probability } 1 - \pi_t,
\end{cases} \quad (1) \]

where \( x_t \) denotes the sequence of prices measured as deviations from their fundamental values and \( \pi_t \) denotes the probability that \( x_t \) follows model (A) depending on \( x_{t-1} \). A newly arisen bubble \( u_t \) is a serially independent and normally distributed random variable with mean 0 and a constant variance \( \sigma_u^2 \) which is unknown to us. The coefficient \( \beta \) is a time dependent parameter whose variation is given by the following random walk process:

\[ \beta_t = \beta_{t-1} + \nu_t, \quad \nu_t \sim N(0, \sigma_\nu^2). \quad (2) \]

Like \( \sigma_u^2 \), the constant variance of innovations \( \sigma_\nu^2 \) is unknown to us.

Let us consider briefly the implications of this model. Since \( x_t \) is a sequence from which the influence of the fundamental factors is removed, \( x_t \) is expressed by a divergent time series model when a speculative bubble continues. We describe this phenomenon by the autoregressive model (A) with parameter \( \beta_t \) exceeding unity. As implied by a speculative bubble, the divergent sequence will suddenly crash at a certain unknown time. We formulate this event by a systematic and probabilistic switch from model (A) to model (B).

Namely, we assume that the probability of bubble continuation can be expressed by

\[ \pi_t = e^{-\alpha x_{t-1}} \times \gamma > 0, \quad \alpha, \gamma > 0 \],

where \( \alpha \) and \( \gamma \) are unknown parameters. This formulation implies that \( \pi_t \) decreases as \( x_t \) becomes greater in absolute value. To put it another way, the probability of a bubble crash, \( 1 - \pi_t \), is an increasing function of how distant the observed bubble deviates from market fundamentals. When \( \alpha = 0 \), \( \pi_t \) is independent of \( x_{t-1} \), and therefore the probability of a crash is constant, which corresponds to the formulation given by Blanchard and Watson (1982). When \( \alpha = \gamma = 0 \), the whole process is described by the autoregressive process (A) and when \( \gamma \) is large or \( \beta_t = 0 \), the process reduces to a simple white noise process and there is no speculative bubble. Thus, by investigating the parameter estimates, we may statistically test the properties of the process.

In principle, we can generalize our formulation by considering a broader class of stochastic models for \( u_t \), such as ARMA processes or by introducing the fundamental values into the functional form of the transition probability (3). However, we have tried to keep our model as simple as possible because this paper is only meant to be a first step in this research direction. The specification, (3), of the probability turns out to be one of the few analytically tractable formulations in the following analyses.

When the probability structure of the crashes is taken into consideration, we see that the bubble cannot continue forever. As it grows, the probability of a crash approaches unity and \( x_t \) will sooner or later be pulled back to its mean level, which is taken to be zero here. In this way, the time series of the price never diverges, but exhibits more or less a stable behavior in the longer run.

3. Recursive estimation

In this section, we describe a Bayesian recursive technique to estimate the parameters of our model. In the Appendix, we give rather detailed explanations of the estimation process at period 1, when only the initial value \( x_0 \) and the first data point \( x_1 \) are given, and find the estimators of \( \beta_1 \) and its variance \( \sigma_{\beta_1}^2 \). In the same way as the estimation process at period 1, here, we proceed to derive the general recursive formulae for estimation from period \( t = 1 \) to \( t \), and thereby obtain the parameter estimates given the time series data up to period \( t \).

Before proceeding to this task, we put \( X = (x_0, x_1, \ldots, x_t) = (x_0, x_1^{t-1}) \) the set of data observations up to period \( t \), and by \( I = (i_1, i_2, \ldots, i_t) \), we denote the set of ordered integer indices where each \( i_s (s = 1, 2, \ldots, t) \) is either 1, 2 or 3.
With these new notations, we write down the joint density for $\alpha$, $\beta_t$, and $\gamma$ conditioned on $X_t^t$:

$$P(x, \beta, \gamma|X_t^t) = \frac{\prod_{i=1}^{t} g(t^i) f(t^i)(x, \beta, \gamma|X_{t_i-1}^t) N_{\beta_t}(\hat{\mu}(t^i), \sigma^2(t^i-1) + \sigma_b^2)}{\prod_{i=1}^{t} g(t^i)}$$

where $f(t^i)$ and $g(t^i)$ are the joint prior and density functions for $\alpha$ and $\gamma$ conditioned on $X_t^t$, and $P_{\beta_t} (\beta, \gamma|X_t^t)$ is the joint prior density function for $\beta$ and $\gamma$ conditioned on $X_t^t$.

Now our main task is to calculate the updated posterior density $P(x, \beta_t, \gamma|X_t^t)$ by Bayesian theorem:

$$P(x, \beta_t, \gamma|X_t^t) = \frac{P(x, \beta, \gamma|X_t^t) P(x_t|\beta, \gamma, X_t^t)}{P(X_t^t)}$$

where, from (1) and the normality of $u_t$, we have

$$P(x_t|\beta, \gamma, X_t^t) = e^{-\gamma x_t^2/2} \frac{1}{\sqrt{2\pi \sigma_a^2}} e^{-\beta x_t^2/2}$$

Therefore, in view of (6), multiplying (5) by (7) yields the updated formula of (5) for period $t$ if and only if we have

$$P_{t}(x, \gamma|X_t^t) = \left[ \tilde{c}_0 + \sum_{i=0}^{t-1} \hat{\theta}(i) \right] e^{-\sum_{i=0}^{t-1} \hat{\theta}(i)}$$

and

$$f(t^i-1, 1) = f(t^i-1) N_{\beta_t}(\hat{\mu}(t^i-1), (\sigma^2(t^i-1) + \sigma_b^2) \sigma^2_{\beta_t})$$

$$f(t^i-1, 2) = -f(t^i-1, 3) = \frac{-f(t^i-1)}{P(X_{t_i}^i)} N_{0, \sigma^2_{\beta_t}}$$

$$g(t^i) = g(t^i)$$

$$\hat{\mu}(t^i, 1) = \frac{\hat{\mu}(t^i-1) \sqrt{\sigma^2(t^i-1) + \sigma_b^2} + \beta_t x_t}{\sqrt{\sigma^2(t^i-1) + \sigma_b^2}}$$

$$\hat{\mu}(t^i, 2) = \hat{\mu}(t^i, 3) = \hat{\mu}(t^i-1)$$

Finally, the estimate of the variance of $\beta_t$ is given by

$$\hat{\sigma}^2_{\beta_t} = \text{Var}(\hat{\beta}_t|X_t) = \sum_{i=1}^{t} f(t^i) g(t^i) \hat{\sigma}^2(t^i)$$

As can be easily checked, each expression from (5) to (20) for period $t$ is the generalization of the corresponding expression at period 1, namely, each expression from (27) to (43), (45) to (46) in the Appendix. This should establish, by way of mathematical induction, that the obtained general recursive formulae are valid for all periods $t \geq 1$. However, although the recursive procedure is theoretically applicable to the data set of any length, the required computer capacity increases in an explosive way. In order to reduce the computational burden for practical use, therefore, we make use of a simple approximation method which corresponds to what Harrison and Stevens (1976) refer to as the collapsing process.\(^1\)

Finally, to carry out the recursive procedure explained above, two variance parameters are to be specified. These are the dispersion parameters for the random terms in (1) and (2), i.e., $\sigma_x^2$ and $\sigma_b^2$. The likelihood function for these parameters can be obtained in the following way.

Let us put $\sigma^2 = (\sigma_x^2, \sigma_b^2)$ for simplicity. The likelihood function for $\sigma^2$ with $T$ periods of data is defined as

$$P_t(x_1, x_2, \ldots, x_T|x_0, \sigma^2) = \prod_{i=1}^{T} P_t(x_i|X_{i-1}^i, \sigma^2)$$

$$\cdots \cdots P_1(x_1|x_0, \sigma^2) = \prod_{i=1}^{T} P_t(x_i|X_{i-1}^i, \sigma^2)$$

On the other hand, since

$$P(x, \beta_t, \gamma|X_t^t, \sigma^2) = P(x, \beta_t, \gamma|X_t^t, \sigma^2) P(x_t|\beta, \gamma, X_t^t, \sigma^2)$$

\(^1\) See Condensation in the Appendix.
and

\[ P_t(x_t | \mathbf{X}^{-1}, \sigma^2) = \int \int \int P(x_t, \alpha, \beta_t, \gamma | \mathbf{X}^{-1}, \sigma^2) \, dx \, d\beta_t \, d\gamma, \]

we have

\[ P_t(x_t | \mathbf{X}^{-1}, \sigma^2) = \sum_{t} f(t | g(t)). \]

Therefore, the log likelihood function of \(\sigma^2\) can be expressed by

\[ \log P_t(\mathbf{X}^{t} | x_0, \sigma^2) = \sum_{t=1}^{T} \log \left( \sum_{t} f(t | g(t)) \right) \]

and the resulting set of variances \(\hat{\sigma}^2 = (\hat{\sigma}_{x}^2, \hat{\sigma}_{y}^2)\) which maximize (25) are the desired estimates.

4. Applications to the stock markets of the United States, Japan, and China

In this section, we apply our model to the stock markets of the United States, Japan, and China. The monthly time series data we have chosen are the Dow Jones Industrial Average Stock Price Index (hereinafter DJ),\(^2\) the Nikkei225 index (hereinafter Nikkei225) for the Japanese stock market, and the Shang Zheng Zhong Zhi index (hereinafter SZZZ) for the Chinese stock market.\(^3\)

The theoretical fundamentals of the stock price equal the discounted sum of expected future dividends if the stock market is efficient. However, such measures in practice are hardly computable since the expected values of future variables are not observable. Moreover, it is to be noted that the efficiency of the stock markets of the United States and other countries has long been questioned, based on a number of empirical findings (e.g., Summers, 1986; West, 1988; Poterba and Summers, 1988). As a consequence, there have been a number of attempts that have sought systematic relationships between the stock price index and macroeconomic indicators, either successful or unsuccessful. For example, Engle and Rangel (2009) show the relationship between low frequency volatility in nearly fifty countries’ equity markets and the macroeconomic indicators of GDP, inflation, and short-term interests. Chen et al. (1986) test the effects of interest rates, inflation, industrial production, and the spreads between high- and low-grade bonds on stock market returns. Flannery and Protopapadakis (2002) identify three nominal variables (CPI, PPI, and Money Aggregate) and three real variables (Balance of Trade, Employment Report, and Housing Starts) affecting stock returns.

Here, for simplicity and to maintain consistency between our applications to the United States, Japan, and China, we adopt as benchmark studies only two factors reflecting the fundamentals: i.e., the nominal GDP and a representative interest rate. Thus, in order to remove the movement of the part of the variations attributable to market fundamentals, we regress the original monthly data (figures measured as the average of each month) of stock price indexes in logarithms on the level of the interest rate and the logarithm of the nominal GDP\(^4\) as follows:

\[ \ln(\text{Index}) = \text{Constant} + a \text{ Interest Rate} + b \ln(\text{GDP}). \]

4.1. The United States stock market

The sample period for DJ is from January 1980 to December 2009. The results estimated for the fundamentals equation are

\[ \ln(P_t) = -16.55 + 0.01 r_t + 1.83 \ln Y_t, R^2 = 0.953, \]

where \(P_t, r_t, \) and \(Y_t\) represent, respectively, the DJ index, the 3-month Treasury Bill (TB) rate, and the nominal GDP, and the numbers in parentheses are the t-statistics. We can say that the GDP is the main

\(^2\) We have also used the closing prices of the DJ, Nikkei225, and SZZZ at the end of each month and the results are basically the same as listed below.

\(^3\) We do not assume any form of moving average process when estimating return indices, though both French (1980) and Jaffe and Westerfield (1985) report evidence of calendar effects. Compared to the smoothed ones, the unsmoothed indices best preserve the information on market conditions, especially on the deviation of a stock price from its fundamental.

\(^4\) The quarterly GDP are first seasonally adjusted using the X11 method, then transformed to be of monthly frequency by interpolation with the cubic spline method.
The determinant of the stock market. The interest rate also affects the stock market. However, its marginal effect is very small compared to that of the GDP. The realized DJ index (thick line) and its fitted values (thin line) are plotted in Fig. 1. A rather high coefficient of determination $R^2$ indicates that the DJ index reflects the US economy’s fundamentals quite well, at least as late as the mid 1990s. The resultant residuals, shown in Fig. 2 in dollars, are the concern of our analysis. From these two figures we can find that the DJ index deviated from its fundamental value positively from 1995 through 2002 and attained its peak around 1998 (the dot com or IT bubble), and negatively from 2005 through 2010 and reached the bottom around 2008 (at the subprime loan crisis).

Before beginning the estimations, some parameters have to be specified. First, as for the initial values, $\hat{\beta} = 1$, $\hat{\sigma}_0 = \gamma_0 = 0.01$ and $\hat{\sigma}_0 = 0.01$ are adopted. Although we have tried several other values, the effects of the initial values disappeared rapidly and the estimation results stayed almost the same as those given below. Second, two unknown variances $\sigma_u^2$ and $\sigma_v^2$ are to be specified as well. These are estimated by the maximum likelihood method.

5 The exclusion of this variable will not affect our results significantly. It is taken into account here in order to keep our regression similar to the previous literature.

6 Compared with the larger deviation since 1995, the bubble crash of Black Monday in 1987 seems unimportant through the regression for the sample period from 1980 to 2010. We have divided the full sample into four sub-samples and found our empirical models fit very well. The results are not reported to save space, but are available on request.
using (25). The estimated values are obtained as $\hat{\sigma}_0^2 = 0.0872$ and $\hat{\sigma}_1^2 = 0.223^2$, respectively.

To begin with, the estimated values of the coefficient $\hat{b}_t$ are plotted in Fig. 3, and they fluctuate rather widely, reflecting the fact that the maximum likelihood estimate of $\sigma_0$ can be as large as 0.223. The intertemporally changing estimate $\hat{b}_t$ records its maximum of 1.615 at November 2001 and its minimum, −0.476, at January 1994.

The cells in Table 1 marked by “*” are those periods when $\hat{b}_t > 1$. Out of the total 360 months, there are 119 months when $\hat{b}_t > 1$. And there are 21 months when the hypothesis $\hat{b}_t > 1$ is statistically significant. The two longest lasting bubbles, as long as six months, are recorded from February to July in both 1982 and 1984. Three five-month periods, from November 1995 to March 1996, from September 1998 to January 1999, and from November 2007 to March 2008, are the second longest runs of

**Table 1**

The period of bubbles ($\hat{b}_t > 1$) for the US stock market.

<table>
<thead>
<tr>
<th>y/m</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crs</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1987</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Crs</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1988</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1989</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1992</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1993</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1995</td>
<td>*</td>
<td>Crs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1996</td>
<td>*</td>
<td></td>
<td>Crs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1997</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>*</td>
<td>Crs</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>Crs</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2005</td>
<td>*</td>
<td>Crs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2006</td>
<td>*</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2007</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2008</td>
<td>*</td>
<td>Crs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* means a period when $\hat{b}_t > 1$.

**Fig. 4.** Plot of $t = \frac{\hat{b}_t - 1}{\bar{b}_t}$ for DJ.
bubble. On the other hand, the longest spell of fads, or a period of $\beta_t < 1$, is observed from November 1990 to October 1994, as long as 48 months. We can also confirm these findings through the plot of the $t$-values or $(\beta_t - 1)/\sigma_\alpha$ in Fig. 4, where the two horizontal dashed lines represent $t = \pm 1.96$ as the benchmark critical values of the two-tailed 5% significance level.

Fig. 5 plots the estimates of the probability of a bubble’s continuation, $\pi_t$, which fluctuates over a relatively narrow range between 0.979 (May 1999) and 0.995 (May 2004). These are in part due to the small and stable estimates for $\alpha$ and $\gamma$ throughout the sample period, implying that most of the variations in $\pi_t$ are caused by $|x|$. In fact, we can see a reasonable resemblance in the time series profiles of Figs. 3 and 4 once the vertical axis of either one is turned upside down.

4.2. The Japanese stock market

The sample period for the Nikkei 225 is from May 1979 to January 2010. The results estimated for the fundamental equation are

$$\ln(P_t) = -16.81 + 0.11 r_t + 2.01 \ln Y_t, R^2 = 0.461,$$

where $P_t$, $r_t$, and $Y_t$ represent Nikkei225, the 1-year deposit interest rate, and the nominal GDP, respectively, and the numbers in parentheses are the $t$-statistics. We can see that the nominal GDP is also the main determinant for the Japanese stock market, although the adjusted coefficient of determination $R^2$ is smaller than that for the United States. The smaller $R^2$ might indicate that it is common for the Japanese stock market to deviate widely from fundamentals.
and the bubble part dominates the movements of the stock market. What appears strange is that the interest rate affects the stock market positively, which might be caused by the long-term downside momentum in both Japanese stock prices and interest rates since the end of the 1990s. The realized Nikkei225 and its fitted value are plotted in Fig. 6 and the resulting residuals are shown in Fig. 7. We use $r_u^{2} = 0.4692$ and $r_v^{2} = 0.0112$ as the maximum likelihood estimates.

Fig. 8 shows that the estimated $\hat{b}_t$ (thin line) fluctuates widely in Japan since the latter part of the 1980s. The estimated minimum $\hat{b}_t$ equals 0.615, recorded in April 1990, which rejects the hypothesis of $b_t = 0$. Figs. 9 and 10 support the assertion that the Japanese stock market around 1990 was truly in a bubble state, and the probability of $\hat{p}_t$ decreases sharply before the bubble burst.

The cells in Table 2 marked by “•” are the periods when $\hat{b}_t$ exceeds unity. There are 84 months out of the 188 samples when the estimated $\hat{b}_t > 1$, among which $\hat{b}_t > 1$ is statistically significant in 14 months.

### 4.3. The Chinese stock market

The sample period for SZZZ is from January 1997 to April 2010. The results estimated for the fundamental equation are presented here.

7 From here on, to save space, we put the time series data for Japan and China in the same figures, except for the fundamentals series.
The realized SZZZ and its fitted values are plotted in Fig. 11. The resulting residuals are plotted in Fig. 7 by the thick line. We obtain $\sigma_t^2 = 0.061^2$ and $\sigma_t^2 = 0.123^2$. Comparing the estimated $\hat{\sigma}_t$ of these three indices, we know the one for DJ is the largest, and the medium is for Nikkei225 and the smallest is for SZZZ. This ranking might be related to daily price variation constraints in the Japanese and Chinese stock markets.

The estimated $\hat{\beta}_t$, plotted in Fig. 8 by the thick line, fluctuates up and down more widely than that for Japan, but not so frequently as that in Fig. 3. From Fig. 9, we can see that most of the $t$-values of $(\hat{\beta}_t - 1)/\sigma_t$ are limited between the two dash lines for $t = \pm 1.96$: only five points break through the top dashed lines, which means there is a bubble period in 2007.

The cells in Table 3 marked by “*” are the periods when $\beta_t$ exceeds unity. The period from June 2004 to July 2005 is the longest
Table 2
The period of bubbles ($\beta_t > 1$) for the Japanese stock market.

<table>
<thead>
<tr>
<th>y/m</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>1979</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1980</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1982</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1983</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1986</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1987</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1988</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1989</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1990</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1991</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1992</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>1993</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1994</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* means a period when $\beta_t > 1$.

** means a period when the hypothesis $\beta_t > 1$ is statistically significant.

Crs means a bubble crash.

Fig. 11. SZZZ and its fundamentals.

Table 3
The period with bubbles ($\beta_t > 1$) for the Chinese stock market.

<table>
<thead>
<tr>
<th>y/m</th>
<th>January</th>
<th>February</th>
<th>March</th>
<th>April</th>
<th>May</th>
<th>June</th>
<th>July</th>
<th>August</th>
<th>September</th>
<th>October</th>
<th>November</th>
<th>December</th>
</tr>
</thead>
<tbody>
<tr>
<td>1997</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1999</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2000</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2001</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2002</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2003</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2004</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2005</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2006</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2007</td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2008</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2009</td>
<td>*</td>
<td>*</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td>*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* means a period when $\beta_t > 1$.

** means a period when the hypothesis $\beta_t > 1$ is statistically significant.

Crs means a bubble crash.
continuation of a bubble. However, only for five months in 2007 are the results statistically significant.

5. Concluding remarks

This paper presents a simple statistical model for speculative bubbles, which can be estimated by recursive computations. The explicit formulation of the probability of a crash, depending on the value of the deviation from fundamentals, reflects the effects of past experiences on the future expectations of investors. Our model demonstrates the fact that no bubble can grow or live for ever, partly because the majority of investors are in one way or another ready to escape from the bubble crash at the very moment it is likely to crash, and partly because monetary policy or whatever interventions are undertaken by the relevant economic agency will make the deviant prices return back to fundamentals.

The exact recursive formula requires a huge storage capacity, which is hard for standard computers even for a reasonable length of the sample period. For this reason, we introduce an approximation in the computation, yet maintaining the recursive nature of our estimation technique.

We then apply the model to the stock markets of the United States, Japan, and China, and estimate the parameters and the probabilities of a bubble crash. Although the empirical investigation is of preliminary nature, we have obtained some interesting results: the probability of a bubble crash indeed increases as the stock price becomes too high or too low; usually, even if \( \beta_i > 1 \), beta fades out without actually creating a bubble; the time series data of the stock price has been inherently non-stationary; fundamentals dominates the United States stock market but not the Japanese and Chinese stock markets, for both of which the bubble part causes their stock prices to deviate widely and frequently.

Since \( \beta_i \) could be a function of \( \pi_i \), and since the crash of a bubble could extend over a period of several consecutive time-steps, more complicated models are required for future research.

Acknowledgments

We are grateful to Ike Mathur and an anonymous referee for their helpful and extensive comments which greatly improved the paper. The usual disclaimer applies: errors are only ours. Kazumi Asako acknowledges the financial support from Grants-in-Aid for Scientific Research of the Japanese Government and Zhentao Liu acknowledges the financial support from Ministry of Education of China (Grant No. 11YJAZ90095). We benefit a great deal from discussions with Professor Satoru Kanoh at Hitotsubashi University who passed away in 2007.

Appendix A. Recursive estimation of the bubble model

We now develop the recursive estimation procedure of the bubble model (1). For a more detailed discussion of the model, see Asako et al. (1990) and Liu et al. (2011).

A.1. Initial stage

Whenever \( Y \) and \( X \) are random variables defined on the same sample space, we write \( P(Y|X) \) for the conditional probability of \( Y \) conditioned on \( X \). We will allow both \( X \) and \( Y \) to represent vectors. With this notation, we first assume that the prior densities for the parameters \( \alpha, \gamma \) in (3) and \( \beta_0 \) in (1) given by

\[
\begin{align*}
P(x_0|X_{\alpha}) &= \hat{a}_0 \exp^{-\hat{a}_0 x_0}, \quad \hat{a}_0 > 0, \\
P(b_0|x_0) &= N_{b_0}(\hat{b}_0, \hat{\sigma}_b^2), \\
P(\gamma|x_0) &= \hat{c}_0 \exp^{-\hat{c}_0 \gamma}, \quad \hat{c}_0 > 0,
\end{align*}
\]

(26)

where \( N_{b_0}(\cdot, \cdot) \) denotes the normal density with mean \( \zeta \) and variance \( \zeta \). The initial prior distributions of \( \alpha \) and \( \gamma \) are represented by exponential distributions with means the reciprocals of \( \hat{a}_0 \) and \( \hat{c}_0 \), and that of \( b_0 \) by a normal distribution with mean \( \hat{b}_0 \) and variance \( \hat{\sigma}_b^2 \). Moreover, the random variables \( (\alpha|x_0), (\beta|x_0) \), and \( (\gamma|x_0) \) are assumed independent of each other, so that

\[
P(\alpha, \beta, \gamma|x_0) = P(\alpha|x_0)P(\beta|x_0)P(\gamma|x_0) = \hat{a}_0 \exp^{-\hat{a}_0 x_0} \hat{c}_0 \exp^{-\hat{c}_0 \gamma} N_{b_0}(\hat{b}_0, \hat{\sigma}_b^2).
\]

(27)

Then it is easy to see from (2) that

\[
P(\alpha, \beta, \gamma|x_0) = P(\alpha|x_0)P(\beta|x_0)P(\gamma|x_0) = \hat{a}_0 \exp^{-\hat{a}_0 x_0} \hat{c}_0 \exp^{-\hat{c}_0 \gamma} N_{b_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2).
\]

(28)

From the specification of our model (1), we can state that the random variable \( (X_1|\alpha, \beta, \gamma, x_0) \) has the density \( N_{x_0}(\alpha_1 x_0, \sigma_a^2) \) with probability \( \pi_1 \) and \( N_{x_0}(0, \sigma_b^2) \) with probability \( 1 - \pi_1 \). More explicitly, we have

\[
P(x_1|\alpha, \beta, \gamma, x_0) = e^{-\frac{(x_1 - \alpha_1 x_0)^2}{2\sigma_a^2}} + (1 - e^{-\frac{(x_1 - \alpha_1 x_0)^2}{2\sigma_a^2}}) N_{x_0}(0, \sigma_b^2).
\]

(29)

But, since we know the truth

\[
P(x_1|\alpha, \beta, \gamma, x_0) = P(x_1|\alpha, \beta, \gamma, x_0)P(\alpha, \beta, \gamma|x_0),
\]

(30)

we obtain, from (28) and (29), that the left hand side of (30) can be rewritten as

\[
\hat{c}_0 e^{-(\hat{c}_0 + 1)/\hat{a}_0} \exp^{-\frac{(x_1 - \alpha_1 x_0)^2}{2\sigma_a^2}} N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2)
\]

\[
- \hat{c}_0 e^{-\frac{(x_1 - \alpha_1 x_0)^2}{2\sigma_a^2}} N_{x_0}(0, \sigma_b^2) N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2) + \hat{c}_0 e^{-\frac{x_1^2}{2\sigma_b^2}} N_{x_0}(0, \sigma_b^2) N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2).
\]

(31)

Then, by using the relation

\[
N_{x_0}(\beta|x_0) = N_{x_0}(\beta|x_0, \sigma_\beta^2) N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2)
\]

\[
\times N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2)
\]

(32)

and by straightforwardly computing the following marginal density

\[
P(x_1|x_0) = \int P(x_1|x_0) dx_0 P(\alpha, \beta, \gamma|x_0) = P(x_1|x_0) dx_0 P(\alpha, \beta, \gamma|x_0)
\]

\[
= \frac{\hat{c}_0}{(\hat{c}_0 + 1)} N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2) + \frac{\hat{c}_0}{(\hat{c}_0 + 1)} N_{x_0}(0, \sigma_b^2).
\]

(33)

we obtain the posterior density

\[
P(\alpha, \beta, \gamma|x_1, x_0) = \frac{P(x_1|x_0, \alpha, \beta, \gamma, x_0)}{P(x_1|x_0)} = \frac{\sum_{i=1}^3 f(i) g(i) P(x_1, \alpha, \beta, \gamma|x_0) N_{x_0}(\hat{b}_0, \hat{\sigma}_b^2 + \hat{\sigma}_b^2)}{P(x_1|x_0)}.
\]

(34)

This posterior density, together with the relationship (2), can be used to calculate the joint prior for \( \alpha, \beta_2, \) and \( \gamma \) given \( x_0 \) and \( x_1 \). Then the obtained density takes the place of (27) in the next stage to continue the recursive estimation. In (34), each joint density of \( \alpha \) and \( \gamma \) for given \( x_0 \) and \( x_1 \) takes the form

\[
P_i(\alpha, \gamma|x_0, x_1) = \left( (\hat{c}_0 + \delta(i)) e^{-\frac{(x_1 - \alpha_1 x_0)^2}{2\sigma_a^2}} \right) \left( (\hat{b}_0 + x_0) \delta(i) e^{-\frac{x_1^2}{2\sigma_b^2}} \right).
\]

(35)
where $\delta(i)$ denotes

$$\delta(i) = \begin{cases} 1 & \text{for } i = 1, 2, \\ 0 & \text{for } i = 3. \end{cases}$$

(36)

The other coefficients are defined as follows:

$$f(1) = \frac{1}{\pi(x_0|\gamma)} N_{\gamma}(\hat{\mu}_0 x_0, (\sigma_0^2 + \sigma_1^2) x_0^2 + \sigma_2^2),$$

$$f(2) = -f(3) = -\frac{1}{\pi(x_0|\gamma)} N_{\gamma}(0, \sigma_2^2).$$

(37)

$$g(i) = \left[ \frac{\hat{\alpha}_0}{\hat{\alpha}_0 + |x_0|\hat{\alpha}(i)} \right] \left[ \frac{\hat{\gamma}_0}{\hat{\gamma}_0 + \hat{\alpha}(i)} \right].$$

(38)

$$\mu(1) = \left( \frac{\sigma_0^2 + \sigma_1^2}{\sigma_0^2 + \sigma_1^2 x_0^2 + \sigma_2^2} \right) \hat{\mu}_0,$n

$$\mu(2) = \hat{\mu}(3) = \hat{\beta}_0.$$

(39)

$$\sigma^2(1) = \left( \frac{\sigma_0^2 + \sigma_1^2}{\sigma_0^2 + \sigma_1^2 x_0^2 + \sigma_2^2} \right) \sigma_2^2,$n

$$\sigma^2(2) = \sigma^2(3) = \left( \frac{\sigma_0^2 + \sigma_2^2}{\sigma_0^2 + \sigma_2^2 x_0^2 + \sigma_2^2} \right).$$

(40)

The expected values of $\alpha, \beta_1, \gamma$ at period 1 are then employed as the estimates for these parameters and they are obtained from (34) by

$$\hat{x}_1 = E(x|\mu_1, \gamma),$$

$$\hat{\beta}_1 = E(\beta_1|x_1, \gamma),$$

$$\hat{\gamma}_1 = E(\gamma|x_1, \gamma).$$

Now, $\hat{x}_1$ and $\hat{\gamma}_1$ can be obtained to compute a point of the probability of the continuation of a bubble from period 0 to 1:

$$\tilde{\pi}_1 = e^{-\gamma_1 - \hat{x}_1|x_0}.$$ (41)

Alternatively, since we can conveniently compute explicitly the expected value of this probability as

$$\tilde{\pi}_1 = E(e^{-\gamma_1 - x_0|\gamma}|x_1, \gamma),$$

$$= \sum_{i=1}^{3} f(i) g(i) \left[ \frac{\hat{\gamma}_0 + |x_0|\hat{\gamma}(i)}{\hat{\alpha}_0 + |x_0| (1 + \hat{\alpha}(i))} \right] \left[ \frac{\hat{\gamma}_0 + |x_0|\hat{\gamma}(i)}{\hat{\alpha}_0 + |x_0| (1 + \hat{\alpha}(i))} \right].$$

(45)

$\tilde{\pi}_1$, rather than $\tilde{\pi}_1$, may be used to assess the probability of the persistence of the bubble. Jensen's inequality yields that $\tilde{\pi}_1$ is no greater than $\tilde{\pi}_1$ implying that the use of (45) underestimates the probability of the crash of a bubble.

As for the estimate of $\beta_1$, we may obtain the variance

$$\sigma^2_1 = \text{Var}(\beta_1|x_1, \gamma) = \sum_{i=1}^{3} f(i) g(i),$$

(46)

as well as the mean (42). Then, the Studentized value may be calculated so as to evaluate the significance of $\beta_1$.

### A.2. Condensation

To reduce the computational burden, here we make use of a simple approximation method which Harrison and Stevens (1971) refer to as the collapsing process. Concretely, we approximate the posterior density (34) by a joint density of the form (27), where the first and second moments of the marginal densities for each parameter are equated. That is, at the initial period, (34) is approximated by

$$\tilde{c}_1 e^{-\gamma_1 - \hat{x}_1|x_0} N_{\gamma}(\hat{\beta}_1, \sigma^2_1),$$

(47)

so that the joint prior density at the second period can be written as

$$P(x_1|x_0) = \tilde{c}_1 e^{-\gamma_1 - \hat{x}_1|x_0} N_{\gamma}(\hat{\beta}_1, \sigma^2_1 + \sigma^2_2),$$

(48)

where $\gamma_1$ and $\sigma^2_1$ are equated to, respectively, the reciprocal of the mean estimates (41) and (43):

$$\gamma_1 = 1/\hat{\gamma}_1,$n

$$\sigma^2_1 = 1/\hat{\sigma}^2_1.$$ (49)

(50)

whereas $\beta_1$ and $\sigma^2_1$ are the estimates given by (42) and (46). This procedure can be repeated at each stage.

### References


\footnote{Our simulation results prove this approximation is feasible and accurate. To save space, we do not here report the simulation results, which are available on request.}


