Journal of Economic Dynamics & Control 127 (2021) 104101

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

Speculative bubbles in present-value models: A Bayesian Markov-switching state space approach☆

Joshua C.C. Chan[a,b], Caterina Santi[c,*]

[a] Department of Economics, Purdue University, USA
[b] Department of Economics, University of Technology Sydney, Australia
[c] UCC O'Rahilly Building, College Rd, University College Cork, Cork, Ireland

ARTICLE INFO

Article history:
Received 16 September 2020
Revised 19 January 2021
Accepted 7 March 2021
Available online 10 March 2021

JEL classification:
C11
C32
G12
G14

Keywords:
Rational bubbles
Present-value model
Markov-switching model
State space model
Bayesian analysis

ABSTRACT

We estimate the dynamics of a speculative bubble subject to a surviving and a collapsing regime together with the dynamics of dividends and returns in a tractable state space specification of the present-value model. To estimate this new high-dimensional model, we develop an efficient Markov chain Monte Carlo sampler to simulate from the joint posterior distribution. We find that real-world stock price bubbles show significant Markov-switching structure. Further, the results indicate that dividend growth rates are highly predictable. Finally, we find that bubble variation explains a large share of the variation in the price-dividend ratio and unexpected return.

1. Introduction

The 2008 meltdown of the global financial market has attracted renewed attention on speculative bubbles among academics and policy-makers. Speculative bubbles might be the answer to the question posed by Shiller (1980): If not dividend growth or expected returns, what does move prices? When investors share the belief that a variable or a group of variables, not related to fundamentals, influences prices, it is rational to include this piece of information into prices (Diba and Grossman, 1988). In this context, an explosive behaviour of stock prices is still consistent with a rational behaviour of economic agents. Experimental evidence has also confirmed that bubbles are fueled by symmetrically informed traders (Asako et al., 2020). Recently, Zheng (2020) has found that investors' coordination on fundamental strategy impacts the occurrence and burst of the bubble. This paper aims at detecting speculative bubbles in stock-price data by jointly studying the return and dividend dynamics. Specifically, we want to capture the information in the present-value relations among price-dividend ratios, expected returns, expected dividend growth rates and an eventual rational bubble. Thus, we contribute to both the

☆ This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

* Corresponding author at: Part of this research was conducted while the corresponding author was visiting the department of Economics at University of Technology Sydney.
E-mail addresses: joshuacc.chan@gmail.com (J.C.C. Chan), caterina.santi@ucc.ie (C. Santi).

https://doi.org/10.1016/j.jedc.2021.104101

present-value literature in the spirit of Campbell and Shiller (1988), and the literature on the identification of speculative bubbles.

Variation through time in the price-dividend ratio on corporate stocks conveys essential information about expected returns or expected dividend growth rates (Campbell and Shiller, 1988). Recently, Binsbergen and Koijen (2010) have pioneered a latent variables approach to estimate the expected returns and expected dividend growth rates of the aggregate stock market. They find that returns and dividend growth rates are predictable with R^2 values ranging from 8.2% to 8.9% for returns and 13.9% to 31.6% for dividend growth rates. More recently, Choi et al. (2017) have shown that incorporating regime shifts in the mean of price-dividend ratios into the present value model of Binsbergen and Koijen (2010) increases in-sample predictability. Another extension of Binsbergen and Koijen (2010) is considered by Piatti and Trojani (2017), who use a latent variables approach to estimate a present-value model with time-varying risk.

Despite the relevance of the phenomenon of periodically collapsing bubbles in stock prices, present-value approaches above do not account for it. This paper proposes to incorporate a speculative bubble subject to a surviving and a collapsing regime into the state space framework by Binsbergen and Koijen (2010). Specifically, we contribute to the literature on the present-value model in the spirit of Campbell and Shiller (1988) by allowing prices to deviate from fundamentals because of a latent rational bubble component subject to a surviving and a collapsing regime. Our framework allows us to estimate expected returns and expected dividend growth rates, as well as to identify bubble's collapse dates.

This paper also contributes to the literature on the identification of speculative bubbles. Within the field, empirical papers have mainly proposed two different approaches for the detection of bubbles: indirect and direct bubble tests.[1]

The first group of studies is based on the so-called indirect bubble tests. Here, the authors apply sophisticated cointegration and unit-root tests to a dividend-price relationship (see, e.g., Bohl, 2003; Bohl and Siklos, 2004; Cerqueti and Costantini, 2011; Chen et al., 2016; Diba and Grossman, 1988; Evans, 1991; Froot and Obstfeld, 1991; Hall et al., 1999; Jiang and Lee, 2007; Kanas, 2005; McMillan, 2007; Phillips et al., 2011; Sarno and Taylor, 2003). Among the indirect tests to detect bubbles, some researchers have proposed a Bayesian approach (see, e.g., Check, 2014; Fulop and Yu, 2017; Li and Xue, 2009; Miao et al., 2015; Shi and Song, 2014).

The second group of studies, which are more relevant to this work, implements direct tests for speculative bubbles by explicitly formulating the existence of a bubble in the alternative hypothesis (see Al-Anaswah and Wilfling, 2011; Balke and Wohar, 2009; Lammerding et al., 2013; West, 1987; Wu, 1997). The basic idea in the seminal paper of West (1987) is to compare two alternative estimators for the set of parameters needed to compute the expected present discounted values of a stock's dividend stream, where expectations is conditional on current and past dividends. Specifically, West (1987) constructs one set of estimates by regressing the stock price on a suitable set of lagged dividends. Instead, the other set of estimates is obtained using a pair of equations with one being an arbitrage equation yielding the discount rate, and the other being the ARIMA equation of the dividend process. Then the Hansen and Sargent (1980) formulas may be applied to this pair of equations to obtain a second set of estimates of the expected present discounted value parameters. Under the null hypothesis of no-bubble, the two sets of estimates should be the same, apart from sampling error. West (1987) finds that the test usually rejects the null hypothesis of no bubbles for the US market. More recently, Wu (1997) suggests a state space representation of the deviations of stock prices from the present-value model in which the bubble is included as an unobservable component. In Wu (1997), dividends are assumed to follow an autoregressive process. The analysis attributes large portions of stock price movements to speculative bubbles in the S&P 500. Al-Anaswah and Wilfling (2011); Lammerding et al. (2013) extend the state space model in Wu (1997) by allowing the bubble to switch between two alternative regimes, namely an explosive and a stationary regime. Al-Anaswah and Wilfling (2011) adopt the methodology of Kim and Nelson (1999) to identify regime-switching of speculative bubbles in stock price monthly data. Differently, Lammerding et al. (2013) propose a Bayesian approach to estimate the Markov-switching state space model of speculative bubbles in oil price data. This paper extends the state space model in Al-Anaswah and Wilfling (2011) and Lammerding et al. (2013) by adding a bubble to the state space framework of Binsbergen and Koijen (2010) which model expected returns and expected dividend growth rates as latent variables. Thus, our state space model includes three latent variables namely, expected returns, expected dividend growth rates, and a rational bubble as opposed to one (Al-Anaswah and Wilfling, 2011; Lammerding et al., 2013; Wu, 1997). Moreover, we adopt a Markov-switching approach to identify the bubble's collapsing and surviving regimes. In contrast to Al-Anaswah and Wilfling (2011) and Lammerding et al. (2013), we allow regime-switching also in fundamentals.

Markov-switching models (Hamilton, 1989) have been used extensively in the bubble literature. Driffill and Sola (1998) allow fundamentals to switch between alternative regimes in a stock price model which includes an intrinsic bubble. Hall et al. (1999) propose a univariate Markov-switching Augmented-Dickey-Fuller test to detect bubble episodes, later extended by Shi (2013) to allow for heteroskedasticity. Further, Brooks and Katsaris (2005) show that a three-regime model that allows for dormant, explosive and collapsing speculative behaviour can explain the dynamics of the S&P 500. Shi and Song (2014) propose an infinite hidden Markov model, which allows for an infinite number of regimes to detect, date stamp and estimate speculative bubbles. In a recent contribution, Fulop and Yu (2017) have developed a new model for real time bubbles detection where the dynamic structure of the asset price, after the fundamental value is removed, is subject to two different regimes.

[1] Readers are referred to Gürkaynak (2008) for a survey of econometric tests on asset price bubbles.

A common drawback of the bubble literature is that rejection of the present-value model that are interpreted as evidence of the presence of bubbles can still be explained by alternative structures of the fundamentals. In this paper we mitigate this issue in two ways. First, we restrict our analysis to rational bubbles which impose fairly strong restrictions on the dynamics of the bubble component. Hence these restrictions can help us identify the non-fundamental component in the data. Second, we use a less restrictive fundamentals model. Indeed, our econometric procedure allows us to analyse a more complex model with time-varying discount rates and regime-switching in fundamentals and the bubble. In doing so, we allow the fundamentals part to fit the data better, leaving less room for a bubble. This is an improvement with respect to the model in Al-Anaswah and Wilfling (2011), which allows regime-switching only in the bubble process and assumes constant return rates. Moreover, in their specification the dividend process follows a pure random walk.

In line with previous research on the identification of speculative bubbles, we employ artificial as well as real-world datasets. The artificial bubble processes are defined in the sense of Evans (1991), whereas the real-world datasets are drawn from Datastream. We consider 20 years of monthly data (November 1997–October 2017) for the price index, the dividend yield and the market value. We use data for five countries: United States, United Kingdom, Malaysia, Japan and Brazil. We choose this set of countries because of their economic relevance and the severe bubble episodes experienced in the past (Kindleberger and Aliber, 2003). The advantage of the artificial datasets with respect to real-world data is that the bubbles' collapse dates are known, hence they allow us to assess the accuracy of our bubble-detection method. For the real-world datasets, we rely on what economic historians have classified as bubble periods (Kindleberger and Aliber, 2003).

In order to estimate this new high-dimensional model, we adopt the Bayesian approach and use Markov chain Monte Carlo (MCMC) methods to simulate from the joint posterior distribution. Indeed, when the number of model parameters is large, standard maximum likelihood estimation tends to be numerically unstable and may result in estimates that are locally, but not globally, maximal. In contrast, MCMC methods are numerically more robust and can handle a large number of parameters and latent variables. In addition, one novel feature of our implementation is that it builds upon the band and sparse matrix algorithms for state space models developed in Chan and Jeliazkov (2009), McCausland et al. (2011) and Chan (2013), which are shown to be more efficient than the conventional Kalman filter-based algorithms.

We find that our new bubble-detection method is able to correctly identify 92.27% of all the bubble collapsing dates in the artificial datasets. Moreover, it never signals a bubble when there is none in the price process. These results represent an improvement with respect to the methodology discussed in Al-Anaswah and Wilfling (2011) which correctly identifies around 50% of all the bubble collapsing dates. Also, we find that our framework is able to identify most of the bubble periods classified as such by Kindleberger and Aliber (2003). Consistent with Al-Anaswah and Wilfling (2011) and Lammerding et al. (2013), we document the existence of statistically significant Markov-switching in the data-generating process of real-world stock price bubbles.

Our framework is also able to predict dividend growth rates as well as returns with R^2 values ranging from 74.07% to 78.89% for dividend growth rates and 4.04% and 20.71% for returns in the artificial datasets. In the real-world datasets, the R^2 values for dividend growth rates are quite high, the highest value is recorded for the US where it is equal to 70.49% while the lowest value is registered for Brazil where it is equal to 49.10%. However, the R^2 values for returns are less than 1% with the exception of Brazil where it is above 3%.

We show that present-value models should not ignore the bubble component of stock prices. Indeed, we find that in the surviving bubble regime most of the variation in the price-dividend ratio is related to the bubble variation. Specifically, bubble variation accounts for more than 50% of the price-dividend variation in all the countries under study with the exception of Brazil where it accounts for about 36%. Further, bubble variation explains also a large share of unexpected return variation in the surviving bubble regime.[2]

The paper is structured as follows: next section reviews the present-value model by Campbell and Shiller (1988). In Section 3, we present the econometric model, and Section 4 discusses the posterior sampler. In Section 5, we present the data sources and some descriptive statistics. Section 6 discusses the results, and Section 7 concludes.

2. Economic model

In this section we briefly review the log-linearized present-value model in the spirit of Campbell and Shiller (1988) in which both expected returns and expected dividend growth rates are treated as latent variables as suggested by Binsbergen and Koijen (2010).

Let denote with pd_t and Δd_{t+1} respectively the log price-dividend ratio and the log dividend growth rate

$$\begin{aligned} pd_t &\equiv \log\left(\frac{P_t}{D_t}\right), \\ \Delta d_{t+1} &\equiv \log\left(\frac{D_{t+1}}{D_t}\right). \end{aligned}$$

[2] This result is consistent with Balke and Wohar (2009) which find that the bubble component is substantially important in explaining fluctuations in the log price-dividend ratio when there are no permanent components in market fundamentals.

The log gross return, denoted as r_{t+1}, is defined as follows

$$r_{t+1} \equiv \log\Big(\frac{P_{t+1}+D_{t+1}}{P_t}\Big) = \log\big(P_{t+1}+D_{t+1}\big) - \log\big(P_t\big). \tag{1}$$

Eq. (1) is nonlinear since it involves the log of the sum of price and dividend. However, using the first order Taylor expansion it can be well approximated by

$$r_{t+1} \simeq \kappa + \rho\, pd_{t+1} + \Delta d_{t+1} - pd_t, \tag{2}$$

where κ and ρ are parameters of linearizations, $\kappa = \log(1+exp(\bar{pd})) - \rho\bar{pd}$ and $\rho = \frac{exp(\bar{pd})}{1+exp(\bar{pd})}$, $\bar{pd} = \mathbf{E}[pd]$ (Campbell and Shiller, 1988).

Iterating forward Eq. (2) and imposing the transversality condition, we obtain the unique no-bubble solution

$$pd_t^f = \frac{\kappa}{1-\rho} + \sum_{j=1}^{\infty} \rho^{j-1}\mathbf{E}_t[\Delta d_{t+j} - r_{t+j} \mid \Psi_t], \tag{3}$$

where Ψ_t denotes the economic agents' information set at time t. Similar to Binsbergen and Koijen (2010), we assume that both expected returns ($\mu_t \equiv \mathbf{E}_t[r_{t+1} \mid \Psi_t]$) and expected dividend growth rates ($g_t \equiv \mathbf{E}_t[\Delta d_{t+1} \mid \Psi_t]$) follow an AR(1) process

$$\mu_{t+1} = \delta_0 + \delta_1(\mu_t - \delta_0) + \epsilon_{t+1}^{\mu}, \tag{4}$$

$$g_{t+1} = \gamma_0 + \gamma_1(g_t - \gamma_0) + \epsilon_{t+1}^{g}. \tag{5}$$

The dividend growth rate and the return rate are respectively equal to their expected value plus an orthogonal shock:

$$\Delta d_{t+1} = g_t + \epsilon_{t+1}^{d}, \tag{6}$$

$$r_{t+1} = \mu_t + \epsilon_{t+1}^{r}. \tag{7}$$

Assuming that $\lim_{j\to\infty} \rho^j pd_{t+j} = 0$ and taking expectations conditional upon time t we obtain the fundamental price-dividend ratio:

$$\begin{aligned}
pd_t^f &= \tfrac{\kappa}{1-\rho} + \textstyle\sum_{j=1}^{\infty} \rho^{j-1}\mathbf{E}_t\big[\,\Delta d_{t+j} - r_{t+j} \,\big|\, \Psi_t\big] \\
&= \tfrac{\kappa}{1-\rho} + \textstyle\sum_{j=1}^{\infty} \rho^{j-1}\mathbf{E}_t\big[\,g_{t+j-1} - \mu_{t+j-1} \,\big|\, \Psi_t\big] \\
&= \tfrac{\kappa}{1-\rho} + \textstyle\sum_{j=0}^{\infty} \rho^{j}\mathbf{E}_t\big[\,g_{t+j} - \mu_{t+j} \,\big|\, \Psi_t\big] \\
&= \tfrac{\kappa}{1-\rho} + \textstyle\sum_{j=0}^{\infty} \rho^{j}\mathbf{E}_t\big[\,\gamma_0 + \gamma_1^j(g_t - \gamma_0) - \delta_0 - \delta_1^j(\mu_t - \delta_0) \,\big|\, \Psi_t\big] \\
&= \tfrac{\kappa}{1-\rho} + \tfrac{\gamma_0-\delta_0}{1-\rho} + \textstyle\sum_{j=0}^{\infty} \rho^{j}\mathbf{E}_t\big[\,\gamma_1^j(g_t - \gamma_0) - \delta_1^j(\mu_t - \delta_0) \,\big|\, \Psi_t\big] \\
&= \tfrac{\kappa}{1-\rho} + \tfrac{\gamma_0-\delta_0}{1-\rho} + \tfrac{g_t-\gamma_0}{1-\rho\gamma_1} - \tfrac{\mu_t-\delta_0}{1-\rho\delta_1},
\end{aligned} \tag{8}$$

which uses

$$\mathbf{E}_t[x_{t+j}] = \alpha_0 + \alpha_1^j(x_t - \alpha_0), \tag{9}$$

provided that

$$x_{t+1} = \alpha_0 + \alpha_1^j(x_t - \alpha_0) + \epsilon_{t+1}. \tag{10}$$

Finally, the fundamental price-dividend ratio can be written

$$pd_t^f = A - B_1\hat{\mu}_t + B_2\hat{g}_t, \tag{11}$$

where $A = \frac{\kappa-\delta_0+\gamma_0}{1-\rho}$, $B_1 = \frac{1}{1-\rho\delta_1}$, $B_2 = \frac{1}{1-\rho\gamma_1}$, $\hat{\mu}_t = \mu_t - \delta_0$, and $\hat{g}_t = g_t - \gamma_0$.

It is important to stress that if the transversality solution does not hold, the no-bubble solution pd^f in (11) represents only a particular solution to the difference Eq. (2), and the general solution has the form

$$pd_t = pd_t^f + b_t. \tag{12}$$

where b_t is a rational speculative bubble, that is a deviation of the stock price from fundamentals generated by extraneous factors or rumors and driven by self-fulfilling expectations. The bubble component of the price-dividend ratio satisfies the homogeneous difference equation

$$\mathbf{E}_t[b_{t+i} \mid \Psi_t] = \frac{b_t}{\rho^i}. \tag{13}$$

In line with the literature (i.e., Al-Anaswah and Wilfling, 2011; Lammerding et al., 2013; Wu, 1997), we assume that the bubble component follows a linear AR(1) process

$$b_{t+1} = \frac{b_t}{\rho} + \epsilon_{t+1}^{b}, \quad \epsilon^b \backsim N(0, \sigma_b^2). \tag{14}$$

When estimating the price-dividend Eq. (12), we are confronted with the problem that expected returns, expected dividend growth rates, and the bubble component are unobservables. Hence, we have to express our model in state space form.

3. Econometric model

Bubbles are empirically plausible only if they are likely to collapse after reaching high levels. For instance, Al-Anaswah and Wilfling (2011) and Lammerding et al. (2013) allow the bubble in the present-value model in Wu (1997) to switch between two alternative regimes: an explosive and a stationary regime. Using both stock and oil price data, they document statistically significant Markov-switching between these two regimes. Their findings motivate us to extend the present-value model of Binsbergen and Koijen (2010) to incorporate a speculative bubble that switches between two regimes. The two regimes aim to represent the two distinct phases in the bubble process, namely, one in which the bubble survives and one in which it collapses.

Differently from Al-Anaswah and Wilfling (2011), we allow all the model parameters in the bubble and fundamentals equations to switch between two distinct regimes $S_t \in \{1, 2\}$. The regime indicator S_t, which is independent of all the other shocks in our model, is governed by a first-order Markov process with constant transition probabilities,

$$\Pi = \begin{pmatrix} p_{11} & 1-p_{11} \\ 1-p_{22} & p_{22} \end{pmatrix},$$

where $p_{11} = P(S_t = 1|S_{t-1} = 1)$ and $p_{22} = P(S_t = 2|S_{t-1} = 2)$ are between 0 and 1.

By the end of time t or at the beginning of time $t+1$, economic agents observe S_t but not future states. Thus, economic agents' information set at the end of time t is specified as

$$\Psi_t = \{I_t; S_t\}, \tag{15}$$

where I_t consists of the observed data up to time t.

The model transition equations can be written as:

$$\begin{aligned} \hat{g}_t &= \gamma_{1,S_{t+1}} \hat{g}_{t-1} + \epsilon_t^g, \\ \hat{\mu}_t &= \delta_{1,S_{t+1}} \hat{\mu}_{t-1} + \epsilon_t^\mu, \\ b_t &= 1/\rho_{S_{t+1}} b_{t-1} + \epsilon_t^b. \end{aligned} \tag{16}$$

The dividend growth rate is then equal to

$$\Delta d_{t+1} = \gamma_{0,S_{t+1}} + \hat{g}_t + \epsilon_{t+1}^d, \tag{17}$$

and the price-dividend equation is

$$\begin{aligned} pd_{t+1} &= A_{S_{t+1}} + B_{2,S_{t+1}} \hat{g}_{t+1} - B_{1,S_{t+1}} \hat{\mu}_{t+1} + b_{t+1} \\ &= A_{S_{t+1}} + B_{2,S_{t+1}} \tilde{\gamma}_{1,S_{t+1}} \hat{g}_t - B_{1,S_{t+1}} \tilde{\delta}_{1,S_{t+1}} \hat{\mu}_t + 1/\tilde{\rho}_{S_{t+1}} b_t + B_{2,S_{t+1}} \epsilon_{t+1}^g \\ &\quad - B_{1,S_{t+1}} \epsilon_{t+1}^\mu + \epsilon_{t+1}^b + \epsilon_{t+1}^e, \end{aligned} \tag{18}$$

where, for $i, j \in \{1, 2\}$ and $j \neq i$, we have defined:

$$\begin{aligned} \tilde{\gamma}_{1,S_{t+1}} &= \mathbf{E}_{t+1}[\gamma_{1,S_{t+2}} \,|\, S_{t+1} = i] = p_{ii}\gamma_{1,i} + (1-p_{ii})\gamma_{1,j}, \\ \tilde{\delta}_{1,S_{t+1}} &= \mathbf{E}_{t+1}[\delta_{1,S_{t+2}} \,|\, S_{t+1} = i] = p_{ii}\delta_{1,i} + (1-p_{ii})\delta_{1,j}, \\ \tilde{\rho}_{S_{t+1}} &= \mathbf{E}_{t+1}[\rho_{S_{t+2}} \,|\, S_{t+1} = i] = p_{ii}\rho_{1,i} + (1-p_{ii})\rho_{1,j}. \end{aligned} \tag{19}$$

Notice that we have added to the equation for the price-dividend an orthogonal error ϵ_{t+1}^e. Indeed when we substitute the transition variables at $t+1$, we are confronted with the fact that next period regime is unknown. Hence, we use their expectation conditioned on the information available at time $t+1$ which generate an error, measured by ϵ_{t+1}^e.

In line with Al-Anaswah and Wilfling (2011) and Lammerding et al. (2013), we express the price-dividend equation in first difference to circumvent nonstationarity problems:

$$\begin{aligned} \Delta pd_{t+1} &= A_{S_{t+1}} - A_{S_t} + (B_{2,S_{t+1}} \tilde{\gamma}_{1,S_{t+1}} - B_{2,S_t}) \hat{g}_t - (B_{1,S_{t+1}} \tilde{\delta}_{1,S_{t+1}} - B_{1,S_t}) \hat{\mu}_t + \\ &\quad (1/\tilde{\rho}_{S_{t+1}} - 1) b_t + B_{2,S_{t+1}} \epsilon_{t+1}^g - B_{1,S_{t+1}} \epsilon_{t+1}^\mu + \epsilon_{t+1}^b + \epsilon_{t+1}^e. \end{aligned} \tag{20}$$

Concerning the return process, approximation (2) together with Eq. (12) imply that the return shock has the following form:

$$\epsilon_{t+1}^r = \epsilon_{t+1}^d + \rho \epsilon_{t+1}^{pd}, \tag{21}$$

where $\epsilon_{t+1}^{pd} = B_2 \epsilon_{t+1}^g - B_1 \epsilon_{t+1}^\mu + \epsilon_{t+1}^b$. Since the series of market returns is fully described by the dividend growth rates and the price-dividend ratio, we omit it from our state space specification. Let α_t denote the 7×1 vector of unobservable variables, and y_t be the 2×1 vector of observable variables:

$$\alpha_{t+1} = \begin{pmatrix} \hat{g}_t & \hat{\mu}_t & b_t & \epsilon_{t+1}^g & \epsilon_{t+1}^\mu & \epsilon_{t+1}^d & \epsilon_{t+1}^b \end{pmatrix}', y_{t+1} = \begin{pmatrix} \Delta d_{t+1} & \Delta pd_{t+1} \end{pmatrix}'.$$

We can now express the model in matrix form:

$$\begin{aligned} \alpha_{t+1} &= G_{S_{t+1}}\alpha_t + \Gamma\xi_{t+1}, \\ y_{t+1} &= M_{S_{t+1},S_t} + Z_{S_{t+1},S_t}\alpha_{t+1} + \eta_{t+1}. \end{aligned} \tag{22}$$

where $\alpha_1 \sim N(0, \Gamma V_1 \Gamma')$, G, M, and Z are time invariant matrices of the appropriate dimensions, and ξ_t and η_t are (4×1) and (2×1) vector of disturbances, respectively

$$\xi_{t+1} = \begin{pmatrix} \epsilon^g_{t+1} & \epsilon^\mu_{t+1} & \epsilon^d_{t+1} & \epsilon^b_{t+1} \end{pmatrix}', \eta_{t+1} = \begin{pmatrix} 0 & \epsilon^e_{t+1} \end{pmatrix}',$$

with

$$\xi_t \sim N(0, V_{S_t}),$$

$$\eta_t \sim N(0, R_{S_t}).$$

The model matrices are defined as follows:

$$G_{S_t} = \begin{pmatrix} \gamma_{1,S_t} & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \delta_{1,S_t} & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1/\rho_{S_t} & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \ \Gamma = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \ \Omega_{S_t} = \begin{pmatrix} \sigma^2_{g,S_t} & \sigma_{g\mu,S_t} & \sigma_{gd,S_t} \\ \sigma_{g\mu,S_t} & \sigma^2_{\mu,S_t} & \sigma_{\mu d,S_t} \\ \sigma_{gd,S_t} & \sigma_{\mu d,S_t} & \sigma^2_{d,S_t} \end{pmatrix},$$

$$V_{S_t} = \begin{pmatrix} \Omega_{S_t} & 0 \\ 0 & \sigma^2_{b,S_t} \end{pmatrix}, R_{S_t} = \begin{pmatrix} 0 & \\ & \sigma^2_{e,S_t} \end{pmatrix}, M_{S_t,S_{t-1}} = \begin{pmatrix} \gamma_{0,S_t} \\ A_{S_t} - A_{S_{t-1}} \end{pmatrix},$$

$$Z_{S_t,S_{t-1}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ B_{2,S_t}\tilde{\gamma}_{1,S_t} - B_{2,S_{t-1}} & -B_{1,S_t}\tilde{\delta}_{1,S_t} + B_{1,S_{t-1}} & 1/\tilde{\rho}_{S_t} - 1 & B_{2,S_t} & -B_{1,S_t} & 0 & 1 \end{pmatrix}.$$

Given that the bubble process is exogenous, we have assumed $\sigma_{gb} = \sigma_{\mu b} = \sigma_{db} = 0$.

4. Bayesian estimation

In this section we describe a Bayesian approach for estimating our Markov-switching state space model. Since the number of model parameters is quite large, standard maximum likelihood estimation tends to be numerically unstable and may result in estimates that are locally, but not globally, maximal. For this reason we apply MCMC methods which are numerically more robust. A key novel feature of our approach is that it builds upon the band and sparse matrix algorithms for state space models developed in Chan and Jeliazkov (2009), McCausland et al. (2011) and Chan (2013), which are shown to be more efficient than conventional Kalman filter-based algorithms.

In what follows we use the index i to denote the regime, $i \in \{1, 2\}$. There are two sets of regime-specific parameters. When it does not cause confusion, we would drop the regime index i. For estimation, we split the latent states and model parameters into 7 blocks: states α, covariance matrices Ω_i, variances $(\sigma^2_{b,i}, \sigma^2_{e,i})$, parameters $\Theta_1 = (\gamma_{0,1}, \delta_{0,1}, \gamma_{0,2}, \delta_{0,2})$, $\Theta_2 = (\rho_1, \gamma_{1,1}, \delta_{1,1}, \rho_2, \gamma_{1,2}, \delta_{1,2})$,[3] regime indicators S, and Markov regime-switching probabilities p_{11} and p_{22}.

We assume the following prior distributions: i. $\Omega_i \sim IW(\nu_{01}, S_{01})$; ii. $\sigma^2_{k,i} \sim IG(\nu_{02}, S_{0k})$, $k = \{b, e\}$; iii. $\Theta_1 \sim N(\underline{\Theta}_1, V_{\Theta_1})$; iv. $\Theta_2 \sim N(\underline{\Theta}_2, V_{\Theta_2})$; v. $p_{11} \sim Beta(u_{11}, u_{12})$, and $p_{22} \sim Beta(u_{22}, u_{21})$. For brevity, we use Θ to denote the vector (Θ_1, Θ_2).

We define Regime 1 as the bubble surviving regime, while Regime 2 represents the bubble collapsing regime. The main model parameter is ρ, which governs the growth rate of the bubble process. When ρ increases, the bubble's growth rate decreases. In particular, when $\rho \leq 1$, the bubble is explosive; when $\rho > 1$, the bubble follows a stationary AR(1) process. We assume a Normal prior for ρ with mean equal to 0.75 in Regime 1 and 1.25 in Regime 2. For other parameters we assume the same priors across the two regimes. The values of the hyperparameters are informed by the the estimation results of previous studies (Binsbergen and Koijen, 2010; Choi et al., 2017; Piatti and Trojani, 2017). Finally, we adopt a conjugate prior $Beta(15, 1)$ for the transition probabilities p_{11} and p_{22}. Table 1 summarizes the priors and starting values for the MCMC algorithm.

We implement the following 7-block Metropolis-within-Gibbs sampler to simulate from the joint posterior distribution:

1. Sample from $f(\alpha|Y, \Theta, \Omega, \sigma^2_b, \sigma^2_e, S, p_{11}, p_{22})$.
   It can be shown that the full conditional distribution of α is Gaussian. As a first step, we rewrite the transition and the measurement equations in matrix form:

$$H_G\alpha = \tilde{\Gamma}\xi, \ \tilde{\Gamma}\xi \sim N(0, W), \tag{23}$$

[3] The parameter of linearization κ is expressed as a function of ρ; $\kappa(\rho) = \log(1 + exp(\bar{pd})) - \rho\bar{pd}$, where $\bar{pd}$ is the unconditional expected price-dividend ratio. We set it equal to the sample average of the price-dividend ratio of each dataset.

Table 1
Priors and Starting values.

Parameters	Regime 1		Regime 2	
	Prior	Starting value	Prior	Starting value
Ω	$IW(3+2, 0.01*I_3)$	$0.001*I_3$	$IW(3+2, 0.01*I_3)$	$0.001*I_3$
σ_b^2	$IG(5, 0.04)$	0.001	$IG(5, 0.04)$	0.001
σ_e^2	$IG(5, 0.0004)$	0.001	$IG(5, 0.0004)$	0.001
ρ	$N(0.75, 0.05^2)$	0.900	$N(1.25, 0.05^2)$	1.100
γ_0	$N(0.00, 0.05^2)$	0.000	$N(0.00, 0.05^2)$	0.000
γ_1	$N(0.50, 0.05^2)$	0.500	$N(0.50, 0.05^2)$	0.500
δ_0	$N(0.02, 0.05^2)$	0.000	$N(0.02, 0.05^2)$	0.000
δ_1	$N(0.80, 0.05^2)$	0.800	$N(0.80, 0.05^2)$	0.800
$p_{11},\ p_{22}$	$Beta(15, 1)$	0.800		

$$Y = \tilde{M} + H_Z\alpha + \eta,\ \ \eta \sim N(0, \Phi), \tag{24}$$

where $H_G = \begin{pmatrix} I_7 & & & \\ -G_{S_2} & I_7 & & \\ & \ddots & \ddots & \\ & & -G_{S_T} & I_7 \end{pmatrix}, W = \begin{pmatrix} \Gamma V_{S_1}\Gamma' & & & \\ & \Gamma V_{S_2}\Gamma' & & \\ & & \ddots & \\ & & & \Gamma V_{S_T}\Gamma' \end{pmatrix}$

$$H_Z = \begin{pmatrix} Z_{S_1,S_0} & & & \\ & Z_{S_2,S_1} & & \\ & & \ddots & \\ & & & Z_{S_T,S_{T-1}} \end{pmatrix}, \Phi = \begin{pmatrix} R_{S_1} & & & \\ & R_{S_2} & & \\ & & \ddots & \\ & & & R_{S_T} \end{pmatrix}$$

$$\alpha = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_T \end{pmatrix}, \tilde{\Gamma} = \begin{pmatrix} \Gamma \\ \vdots \\ \Gamma \end{pmatrix}, \xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_T \end{pmatrix}, Y = \begin{pmatrix} y_1 \\ \vdots \\ y_T \end{pmatrix}, \tilde{M} = \begin{pmatrix} M_{S_1,S_0} \\ \vdots \\ M_{S_T,S_{T-1}} \end{pmatrix}, \eta = \begin{pmatrix} \eta_1 \\ \vdots \\ \eta_T \end{pmatrix}.$$

Then,[4] the conditional posterior $[\alpha|Y, \Theta, \Omega, \sigma_b^2, \sigma_e^2, S, p_{11}, p_{22}] \sim N(\hat{\alpha}, P^{-1})$, where

$$\begin{aligned} P &= H_G' W^{-1} H_G + H_Z' \Phi^{-1} H_Z, \\ \hat{\alpha} &= P^{-1}(H_Z' \Phi^{-1}\big(Y - \tilde{M}\big)). \end{aligned} \tag{25}$$

To simulate from $N(\hat{\alpha}, P^{-1})$, we first obtain the Cholesky factor C of P such that $C'C = P$. Then, given $u \sim N(0, I)$, we solve $Cx = u$ for x by back substitution and take $\alpha = \hat{\alpha} + x$. It can be shown that $\alpha \sim N(\hat{\alpha}, P^{-1})$; see, e.g., Chan and Jeliazkov (2009).

2. Sample from $f(\Omega|y, \alpha, \Theta, \sigma_b^2, \sigma_e^2, S, p_{11}, p_{22})$. This step is standard as Ω follows an Inverse-Wishart distribution:

$$\left[\Omega_i\middle|y, \alpha, \Theta, \sigma_b^2, \sigma_e^2, S, p_{11}, p_{22}\right] \sim IW\left(\nu_{01} + \sum_{t=1}^{T}\mathbb{I}(S_t = i), S_{01} + \sum_{t=1}^{T}\left(e_t e_t^{'}\right)\mathbb{I}(S_t = i)\right),$$

where IW stands for the Inverse-Wishart distribution, $e_t = (\epsilon_t^g, \epsilon_t^\mu, \epsilon_t^d)$.

3. Sample from $f(\sigma_k^2|y, \alpha, \Theta, \Omega, S, p_{11}, p_{22})$, $k = \{b, e\}$. This step is standard as each of the variances follows an Inverse-Gamma distribution:

$$[\sigma_{k,i}^2|y, \alpha, \Theta, \Omega, S, p_{11}, p_{22}] \sim IG\left(\nu_{02} + \frac{\sum_{t=1}^{T}\mathbb{I}(S_t = i)}{2}, S_{0k} + \frac{\sum_{t=1}^{T}(\epsilon_t^k)^2\mathbb{I}(S_t = i)}{2}\right),$$

where IG stands for the Inverse-Gamma distribution.

4. Sample from $f(\Theta_1|y, \alpha, \Theta_2, \Omega, \sigma_b^2, \sigma_e^2, S, p_{11}, p_{22})$. This step is also standard as Θ_1 follows a Normal distribution. To see that, we first write the measurement equation as

$$y_t = M_{S_t,S_{t-1}} + (Z_{S_t,S_{t-1}}G_{S_t})\alpha_{t-1} + Z_{S_t,S_{t-1}}\Gamma\xi_t + \eta_t. \tag{26}$$

[4] Note that the first three diagonal elements of $\Gamma V_{S_t}\Gamma'$ are zero, hence matrix W is singular. We substitute the zero elements with 10^{-8} in order to preserve the invertibility of W.

We can then express the constant term $M_{S_t,S_{t-1}}$ as $M_{S_t,S_{t-1}} = C_{S_t,S_{t-1}} + X_{S_t,S_{t-1}}\Theta_1$, where $C_{S_t,S_{t-1}} = \big(0,\ \kappa(\rho_{S_t})/(1-\rho_{S_t}) - \kappa(\rho_{S_{t-1}})/(1-\rho_{S_{t-1}})\big)'$,

$$X_{1,1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, X_{1,2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \frac{1}{1-\rho_1} & -\frac{1}{1-\rho_1} & -\frac{1}{1-\rho_2} & \frac{1}{1-\rho_2} \end{pmatrix},$$

$$X_{2,2} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, X_{2,1} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ -\frac{1}{1-\rho_1} & \frac{1}{1-\rho_1} & \frac{1}{1-\rho_2} & -\frac{1}{1-\rho_2} \end{pmatrix}.$$

Using standard linear regression results, one can show that the conditional posterior is $[\Theta_1|y,\alpha,\Theta_2,\Omega,\sigma_b^2,\sigma_e^2,S,p_{11},p_{22}] \sim N(\hat{\Theta}_1, K_{\Theta_1})$, where

$$\begin{aligned} K_{\Theta_1} &= \big(V_{\Theta_1}^{-1} + \tilde{X}'\Sigma^{-1}\tilde{X}\big)^{-1}, \\ \hat{\Theta}_1 &= K_{\Theta_1}\big(V_{\Theta_1}^{-1}\underline{\Theta}_1 + \tilde{X}'\Sigma^{-1}\big(Y - \tilde{C} - H_{ZG}\alpha\big)\big). \end{aligned} \tag{27}$$

The matrix $\tilde{X}$ is a $2T \times 4$ matrix $\tilde{X} = (X_{S_1,S_0}, ..., X_{S_T,S_{T-1}})'$, $\tilde{C} = (C_{S_1,S_0}, ..., C_{S_T,S_{T-1}})'$, and Y, α are the stacked vectors of y_t and α_t respectively, H_{ZG} and Σ are defined as:

$$H_{ZG} = \begin{pmatrix} 0 & & & \\ Z_{S_2,S_1}G_{S_2} & 0 & & \\ & \ddots & \ddots & \\ & & Z_{S_T,S_{T-1}}G_{S_T} & 0 \end{pmatrix},$$

$$\Sigma = \begin{pmatrix} \big(Z_{S_1,S_0}\big(\Gamma V_{S_1}\Gamma'\big)Z'_{S_1,S_0} + R_{S_1}\big) & & \\ & \ddots & \\ & & \big(Z_{S_T,S_{T-1}}\big(\Gamma V_{S_T}\Gamma'\big)Z'_{S_T,S_{T-1}} + R_{S_T}\big) \end{pmatrix}.$$

[5]

5. Sample from $f(\Theta_2|Y,\alpha,\Theta_1,\Omega,\sigma_b^2,\sigma_e^2,S,p_{11},p_{22})$. Since this conditional distribution is nonstandard, we sample Θ_2 using an adaptive Random Walk Metropolis-Hastings algorithm (Roberts and Rosenthal, 2009). In particular, we update each element of Θ_2 at a time. Given the current draw $\Theta_2^{(s)}$, we update the j-th variable by adding a normal random variable centered at zero to obtain the candidate draw Θ_2^*.[6] The candidate is then accepted with probability

$$a(\Theta_2^{(s)};\Theta_2^*) = \min\left\{\frac{f(\Theta_2 = \Theta_2^*|Y,\alpha,\Theta_1,\Omega,\sigma_b^2,\sigma_e^2,S,p_{11},p_{22})}{f(\Theta_2 = \Theta_2^{(s)}|Y,\alpha,\Theta_1,\Omega,\sigma_b^2,\sigma_e^2,S,p_{11},p_{22})},\ 1\right\}. \tag{28}$$

We impose stationarity conditions for expected dividend growth rates and expected returns, i.e., $-1 < \gamma_1 < 1$, $-1 < \delta_1 < 1$.

6. Sample from $f(S|Y,\alpha,\Theta,\Omega,\sigma_b^2,\sigma_e^2,p_{11},p_{22})$. This step can be done using the algorithm proposed by Chib (1996); see also Kim and Nelson (1999). Specifically, we use the following decomposition of the joint conditional density:[7]

$$f(S|Y,\alpha) = f(S_T|Y,\alpha)\prod_{t=1}^{T-1} f(S_t|S_{t+1},Y_{1:t},\alpha_{1:t}), \tag{29}$$

where $Y_{1:t}$ denotes all the data up to time t, and $\alpha_{1:t}$ is similarly defined.
To compute each of these conditional distributions, we first run the Hamilton filter (Hamilton, 1989) to get the filtered distributions $f(S_t|Y_{1:t},\alpha_{1:t})$, $t = 1,2,...,T$. The last iteration of the filter provides $f(S_T|Y,\alpha)$. More specifically, these filtered distributions are defined by

$$f(S_t|Y_{1:t},\alpha_{1:t}) \propto f(y_t|S_t,\alpha_{t-1},y_{t-1})f(S_t|Y_{1:t-1},\alpha_{1:t-1}),$$

where $f(y_t|S_t,\alpha_{t-1},y_{t-1})$ is a multivariate normal distribution defined by the model.
Then, the conditional distribution $f(S_t|S_{t+1},Y_{1:t},\alpha_{1:t})$ can be computed by using:

$$f(S_t|S_{t+1},Y_{1:t},\alpha_{1:t}) \propto f(S_{t+1}|S_t)f(S_t|Y_{1:t},\alpha_{1:t}),$$

[5] The matrix $\Gamma V_{S_t}\Gamma'$ is singular as the first three diagonal elements are zero. To avoid numerical problems, we substitute them with 10^{-8}.

[6] For the first batch of 50 iterations, we update each variable j by adding a $N(0,0.1^2)$ distributed random variable. Then, after the lth batch of 50 iterations, we update the logarithm of the standard deviation of the proposed normal increment $\log(s_j)$, by adding or subtracting an adaptation amount $\delta(l) = \min(0.01, l^{-1/2})$. Specifically, if the fraction of acceptances of variable j was greater than 0.44 on the lth batch, we increase $\log(s_j)$ by $\delta(l)$; otherwise we decrease it by the same amount. Note that Roberts et al. (1997) and Roberts et al. (2001) show that in various one-dimensional settings the optimal acceptance rate is around 0.44.

[7] For notational convenience, in what follows we suppress the dependence on the model parameters.

where $f(S_{t+1}|S_t)$ is the transition probability and $f(S_t|Y_{1:t}, \alpha_{1:t})$ is calculated using the Hamilton filter as described above. Note that the probability $Pr(S_t = 2|S_{t+1}, Y_{1:t}, \alpha_{1:t})$ can be obtained after the normalization:

$$Pr(S_t = 2|S_{t+1}, Y_{1:t}, \alpha_{1:t}) = \frac{f(S_{t+1}|S_t = 2)f(S_t = 2|Y_{1:t}, \alpha_{1:t})}{\sum_{j=1}^{2} f(S_{t+1}|S_t = j)f(S_t = j|Y_{1:t}, \alpha_{1:t})}. \tag{30}$$

Finally, to obtain a draw from $f(S_t|S_{t+1}, Y_{1:t}, \alpha_{1:t})$, we generate a random number from a uniform distribution between 0 and 1. If the generated number is less than $Pr(S_t = 2|S_{t+1}, Y_{1:t}, \alpha_{1:t})$, we set $S_t = 2$; otherwise we set it equal to 1.

7. Sample from $f(p_{11}, p_{22}|Y, \alpha, \Theta, \Omega, \sigma_b^2, \sigma_e^2, S)$.
Conditional on S, the transition probabilities p_{11} and p_{22} are independent of the data y, the state variables α, and other model parameters.
Since we choose conjugate priors for both p_{11} and p_{22}, we only need to calculate the number of switches between the regimes in order to derive the posterior distributions of p_{11} and p_{22}. The posterior distributions are two independent beta distributions:

$$p_{11} \sim Beta(u_{11} + n_{11}, u_{12} + n_{12}), \quad p_{22} \sim Beta(u_{22} + n_{22}, u_{21} + n_{21}),$$

where n_{ij} refers to the transitions from state i to j; $i, j \in \{1, 2\}$.

5. Data and descriptive statistics

We apply our two-regime Markov-switching state space model to artificial as well as real-world datasets. The artificial bubble processes are defined in the sense of Evans (1991), whereas the real-world datasets are drawn from Datastream. For computational reasons we focus our analysis on the last 20 years of monthly data (November 1997–October 2017) for the price index (PI), the dividend yield (DY) and the market value (MV).[8] We use data for five countries: United States (US), United Kingdom (UK), Malaysia (MY), Japan (JP) and Brazil (BR). We choose this set of countries because of their economic relevance and the severe bubble episodes experienced in the past (Kindleberger and Aliber, 2003). Moreover, this allow us to compare our results with those of previous studies. The model can also be used to study bubbles in investment styles such as industry, size and value, and in individual stocks. The model can be applied to stocks not paying dividends, however it needs to use alternative measures of fundamentals such as earnings.

5.1. Simulated datasets

Evans (1991) considers a class of rational bubbles that are positive and periodically collapsing defined as follows:

$$B_{t+1} = \begin{cases} \frac{B_t}{1+r} u_{t+1} & \text{if } B_t \leq \lambda \\ \left(\delta + \frac{1+r}{\pi}\left(B_t - \frac{\delta}{1+r}\right)\theta_{t+1}\right) u_{t+1} & \text{if } B_t > \lambda, \end{cases} \tag{31}$$

where δ and λ are real positive parameters such that $0 < \delta < \lambda(1 + r)$. The variable u_t is assumed to be independent and identically distributed (iid) lognormally with unit mean. Specifically, we assume $u_t = \exp(x_t - \tau^2/2)$ with x_t iid $N(0, \tau^2)$. The process θ_t is an exogenous iid Bernoulli process which assumes the value 1 with probability π and the value 0 with probability $1 - \pi$, $0 < \pi \leq 1$. The parameters δ, λ, and π govern the frequency with which bubbles erupt, the scale of the bubble and the average length of time before collapse.

As long as the bubble process is below λ, the bubble grows at the mean rate $1 + r$. When $B_t > \lambda$ the bubble grows at the faster mean rate $(1 + r)\pi^{-1}$ and it can collapse with a probability $1 - \pi$ in each period. Whenever the bubble collapses, it falls to δ and the process starts again. For the sake of comparison, we use the parameter specifications of the Evans-bubble process adopted by Al-Anaswah and Wilfling (2011) (see Panel A of Table 2).

We simulate expected returns and expected dividend growth rates according to Eqs. (4) and (5) respectively. We use Eqs. (6) and (11) to generate the dividend growth rate and the fundamental price-dividend ratio series. We generate the bubble stock price-dividend ratio by adding the logarithm of the Evans-bubble (31) to the fundamental price-dividend ratio:

$$pd_t = pd_t^f + \log(B_t). \tag{32}$$

Fig. 1 shows a realization for the log-price-dividend ratio and the log-Evans-bubble.

Finally, the time series of returns is built from the series of the dividend growth rate and the price-dividend ratio using approximation (2). In Panel B of Table 2 we present the parameters used for generating the time-series of our state space model. We generate five artificial datasets with either 100 or 200 observations, for space constraint we report only the results for the biggest datasets of 200 observations.

[8] We use monthly data as we aim to also identify bubble episodes with a duration shorter than a year, which could not be identified with annual data. However, we refrain from using data at a higher frequency because they are generally more noisy, making it more difficult to identify the two bubble regimes. Moreover, historical bubble episodes documented by economic historians have typically lasted for at least a few months (Kindleberger and Aliber, 2003).

Table 2
Parameter specification for artificial datasets.

Parameters	Values	
Panel A: Evans-bubble process parameters		
λ	1.000	
τ^2	0.0025	
r	0.0500	
δ	0.5000	
B_0	0.5000	
π	0.5000	
# observations	100 or 200	
Panel B: state space model parameters		
	Regime 1	Regime 2
γ_0	0.0500	0.0100
δ_0	0.0100	0.0500
κ	1.8600	0.1500
ρ	0.5000	1.0500
γ_1	0.6000	0.5000
δ_1	0.9000	0.7000
σ_g^2	0.0010	0.0020
σ_μ^2	0.0020	0.0010
σ_d^2	0.0015	0.0015
$\sigma_{g\mu}$	0.0013	0.0013
σ_{gd}	0.0009	0.0012
$\sigma_{\mu d}$	0.0009	0.0006

Panel A reports the parameter specification for the Evans-bubble in (31). Panel B reports the parameters used for generating the time-series from Eqs. (4)–(6), and (11).

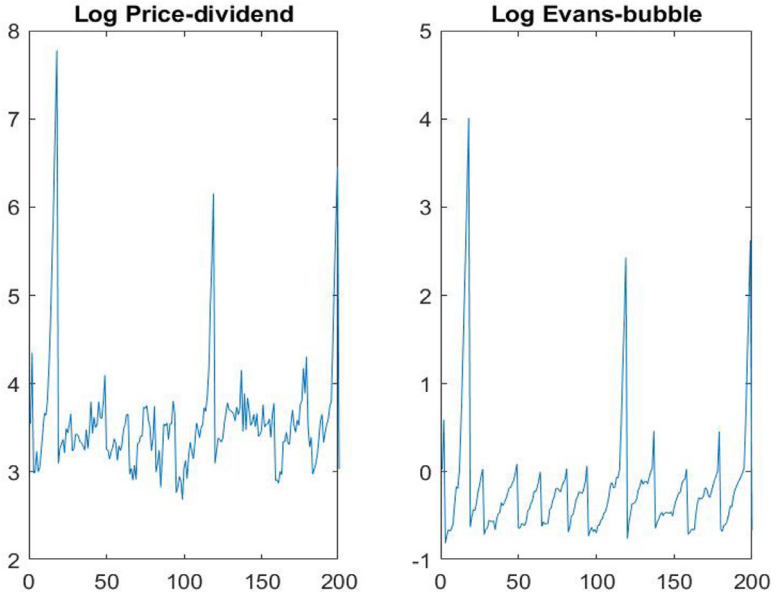

Fig. 1. Artificial dataset 1.

5.2. Real-world datasets

In line with the literature, we use monthly data on the price index (PI), the dividend yield (DY) and the market value (MV) for the Datastream country indices for: United States (US), United Kingdom (UK), Malaysia (MY), Japan (JP), and Brazil (BR).[9] We consider data from November 1997 to October 2017, a total of 20 years of monthly data. All the data are sourced from Datastream (see Appendix Appendix A for further details on the data used).

Table 3 reports some descriptive statistics for the log-price-dividend ratio of the country indices. The US, the UK and Japan have an average log-price-dividend ratio of about 6.50, while the average for Malaysia is 5.08 and 4.26 for Brazil,

[9] Datastream country indices include a representative list of stocks for each country. The sample covers a minimum of 75 - 80% of total market capitalisation. Suitability for inclusion is determined by market value and availability of data.

Table 3
Descriptive Statistics of the Log-Price-Dividend ratio.

	United States	United Kingdom	Malaysia	Japan	Brazil
Panel A: Summary statistics					
Mean	6.5562	6.5585	5.0764	6.6452	4.2627
Std	0.3225	0.2980	0.4924	0.4847	0.8850
Range	1.9831	1.6619	2.8625	2.4947	3.4539
Num.obs.	240	240	240	240	240
Panel B: Correlation matrix					
US	1.0000				
UK	0.7563	1.0000			
MY	0.3726	0.4361	1.0000		
JP	0.5911	0.5195	0.1743	1.0000	
BR	0.0257	0.3348	0.5490	0.3843	1.0000

This table presents summary statistics and correlations for the price-dividend ratio (in log) of the United States, United Kingdom, Malaysia, Japan and Brazil.

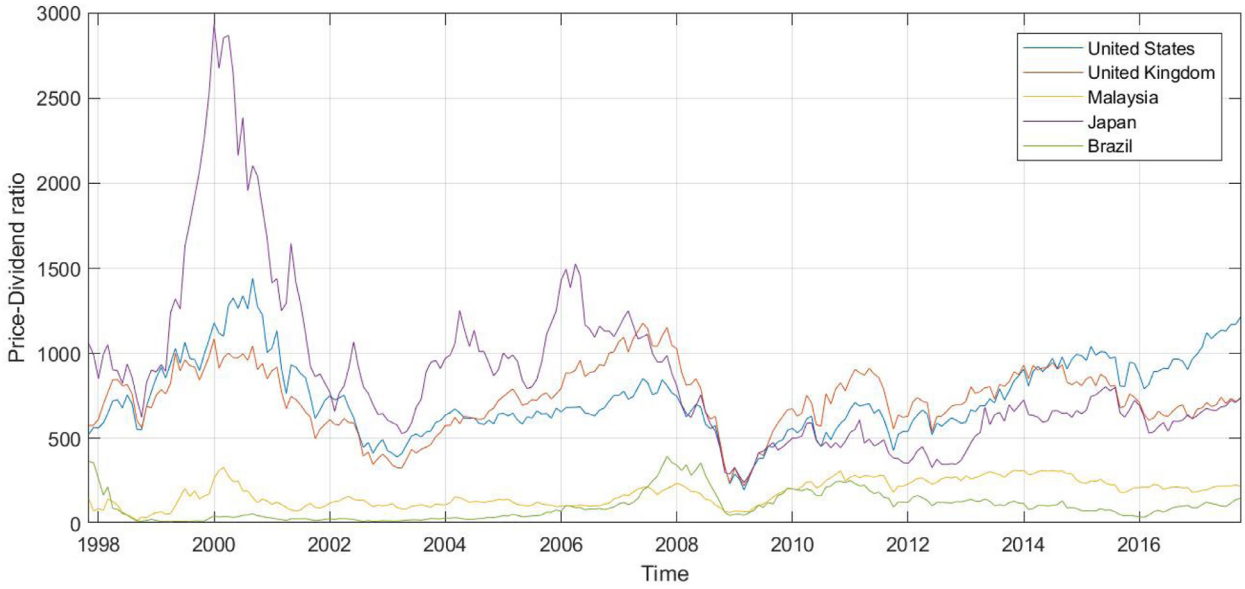

Fig. 2. Price-dividend ratio by country.

with the latter being more volatile. The correlations among the developed countries (US, UK and Japan) are quite high, and they range from 0.75 between the US and the UK, to 0.52 between Japan and the UK. The price-dividend ratio of Malaysia is positively correlated with the other countries with the exception for Japan where the correlation is negative. The price-dividend ratio of Brazil is negatively correlated to Japan and it shows almost no correlation with the US.

Fig. 2 shows the time series plot of the price-dividend ratio for the five country indices. The series for Japan hits a maximum in year 2000, then it decreases and it maintains levels comparable to those of the US and the UK. The price-dividend ratio for Malaysia and Brazil is lower compared to those of the other economies, and they are highly correlated. We can observe that all the countries have experienced a sharp drop in the price-dividend ratio in 2008 in correspondence with the Global Financial Crises.

6. Results

We estimate the present-value model in Eqs. (16) to (18) using the MCMC method described in Section 4. We run a total of 1,000,000 iterations, and discard the first 100,000 as burn-in. To improve the efficiency of the sampler, we perform thinning every 100 draws.[10] Using the sample of the posterior draws, we report the sample means as point estimates. 95% credible intervals are constructed using the 2.50% and 97.50% sample quantiles.

The parameter ρ is the main variable of interest, indeed $1/\rho$ is the autoregressive parameter of the bubble process. A value of ρ less than one implies that the bubble is explosive, while a value greater than one means that the bubble process

[10] Appendix B discusses the efficiency of the Metropolis-within-Gibbs sampler.

Table 4
Parameter estimates – Artificial datasets.

	Dataset 1		Dataset 2		Dataset 3		Dataset 4		Dataset 5	
Regimes	1	2	1	2	1	2	1	2	1	2
γ_0	0.0469	0.0360	0.0027	0.0039	0.0027	0.0021	0.0305	0.0567	0.0566	0.0454
δ_0	0.0272	0.0159	0.0169	0.0180	0.0175	0.0178	0.0112	0.0369	0.0368	0.0258
ρ	**0.7673**	**1.3224**	**0.7619**	**1.3132**	**0.7920**	**1.3296**	**0.7865**	**1.2557**	**0.7716**	**1.3216**
γ_1	0.5793	0.4947	0.5772	0.5249	0.6058	0.5075	0.6062	0.4943	0.5909	0.4868
δ_1	0.8014	0.8552	0.7811	0.8025	0.7546	0.8583	0.8055	0.8316	0.7960	0.8491
p_{11}, p_{22}	**0.8972**	**0.6750**	**0.8779**	**0.6903**	**0.9170**	**0.7179**	**0.9095**	**0.6784**	**0.8960**	**0.6341**
σ_g^2	0.0012	0.0019	0.0012	0.0016	0.0014	0.0025	0.0013	0.0020	0.0012	0.0018
σ_μ^2	0.0017	0.0032	0.0014	0.0037	0.0026	0.0036	0.0015	0.0028	0.0015	0.0030
σ_d^2	0.0039	0.0067	0.0031	0.0073	0.0043	0.0061	0.0043	0.0159	0.0063	0.0101
σ_b^2	0.0115	0.8209	0.0069	0.1200	0.0084	0.0348	0.0090	0.5212	0.0106	0.7752
$\sigma_{g\mu}$	0.0001	0.0001	0.0002	0.0000	0.0002	0.0001	0.0001	0.0000	0.0001	0.0001
σ_{gd}	0.0005	0.0006	0.0004	0.0006	0.0006	0.0006	0.0005	0.0006	0.0005	0.0008
$\sigma_{\mu d}$	0.0003	0.0002	0.0001	0.0003	0.0000	0.0001	0.0000	0.0001	0.0001	0.0004
σ_e^2	0.0018	0.0014	0.0009	0.0012	0.0010	0.0018	0.0014	0.0042	0.0016	0.0026

We present the estimation results of the present-value model in Eqs. (16) to (18) for the five artificial datasets. The model is estimated according to the procedure described in Section 4. Note that $1/\rho$ is the autoregressive parameter of the bubble process. When ρ is less than one the bubble is explosive, when it is greater than one the bubble is stable. $p_{11} = P(S_t = 1|S_{t-1} = 1)$ and $p_{22} = P(S_t = 2|S_{t-1} = 2)$ are the transition probabilities of the two-regime Markov-switching model.

Table 5
95% credible intervals - Artificial datasets.

	Dataset 1		Dataset 2		Dataset 3		Dataset 4		Dataset 5	
ρ_1	0.6966	0.8274	0.6710	0.8478	0.6776	0.9310	0.6761	0.8936	0.6979	0.8468
ρ_2	1.2157	1.4234	1.1453	1.4439	1.1516	1.4618	1.0544	1.4300	1.2249	1.4331

We present the 95% credible intervals for the bubble parameter ρ in Regime 1 (ρ_1) and Regime 2 (ρ_2). The 95% credible intervals consists of the 2.50% and 97.50% quantiles of the posterior distribution of ρ_1 and ρ_2.

is stable. $p_{11} = P(S_t = 1|S_{t-1} = 1)$ and $p_{22} = P(S_t = 2|S_{t-1} = 2)$ are the transition probabilities of the two-regime Markov process. If $p_{11} > p_{22}$, then the bubble surviving regime is more persistent than the bubble collapsing regime.

Table 4 reports the posterior estimates of the parameters for our state space model for the artificial datasets. The main variables are presented in bold characters. Table 5 shows the bubble 95% credible intervals for the parameters ρ_1 and ρ_2, we can observe that the two parameters are statistically significant and significantly different from each other. In particular, ρ_1 is significantly less than one, while ρ_2 is significantly greater than one, meaning that we correctly find significant regime-switch in our artificial datasets. Moreover, in Regime 1 the bubble is explosive while in Regime 2 it collapses. The transition probability p_{11} is always greater than p_{22}, suggesting that the surviving bubble regime is more persistent than the collapsing bubble regime.

Table 6 reports the posterior estimates of the real-world datasets, while Table 7 shows the 95% credible intervals for the autoregressive bubble parameter ρ. Again, we observe ρ to be significantly different in the two regimes for all the real-world datasets except Malaysia. Moreover, we find that ρ_1 is not significantly greater than one and ρ_2 is significantly greater than one at the 5%-level for the US, the UK, Japan and Brazil. Instead, in Malaysia ρ_1 does not appear to be significantly smaller than 1 at 5%-level, meaning that we do not find statistical significance of an explosive behaviour of the bubble process in this country. Consistent with Al-Anaswah and Wilfling (2011), and Lammerding et al. (2013), we find that the transition probability p_{11} is always greater than p_{22}, meaning that the surviving bubble regime is more persistent than the collapsing bubble regime in all the real-world datasets. In particular, the probability of remaining in Regime 1 (p_{11}) is higher than 99% in the US, the UK and Japan, while in Malaysia it is 98.40% and it equals 97.23% in Brazil. The estimates of the transition probability p_{22}, instead, vary between 80.29% (Brazil) and 92.89% (Japan). These results also imply that the expected duration of the surviving regime $1/(1 - p_{11})$ is higher than the expected duration of the collapsing regime $1/(1 - p_{22})$. Concerning the remaining model parameters, we find that they are not significantly different across the two regimes. The unconditional expected log dividend growth rate (γ_0) and the unconditional expected log return (δ_0) are not significantly different from zero. Consistently with Fama and French (1988), Campbell and Cochrane (1999), Ferson et al. (2003), Pástor and Stambaugh (2009), Binsbergen and Koijen (2010) and others, we find expected returns to be highly persistent with a monthly persistence coefficient (δ_1) of above 0.75 for all the datasets. The estimated persistence of expected dividend growth rates (γ_1) instead is generally lower, it ranges between 0.49 and 0.56. Furthermore, shocks to expected dividend growth rates and expected returns are generally positively correlated.

We next examine the identification of speculative bubbles, the prediction of returns and dividend growth rates, and the variance decomposition of the price-dividend ratio and unexpected returns in Sections 6.1–6.3, respectively.

Table 6
Parameter estimates - Real-world datasets.

	United States		United Kingdom		Malaysia		Japan		Brazil	
Regimes	1	2	1	2	1	2	1	2	1	2
γ_0	0.0025	0.0007	0.0031	0.0021	0.0011	0.0000	0.0007	0.0003	0.0032	0.0018
δ_0	0.0167	0.0193	0.0161	0.0192	0.0192	0.0195	0.0188	0.0195	0.0170	0.0182
ρ	**0.9331**	**1.2458**	**0.9236**	**1.2385**	**0.9990**	**1.2487**	**0.9407**	**1.2503**	**0.8732**	**1.2182**
γ_1	0.5661	0.4911	0.5573	0.4921	0.5536	0.5133	0.5650	0.5007	0.5382	0.5065
δ_1	0.7723	0.7870	0.7527	0.7761	0.8649	0.7926	0.7949	0.7989	0.8331	0.7485
p_{11}, p_{22}	**0.9915**	**0.9159**	**0.9905**	**0.9061**	**0.9840**	**0.8777**	**0.9948**	**0.9289**	**0.9723**	**0.8029**
σ_g^2	0.0004	0.0069	0.0004	0.0059	0.0004	0.0025	0.0004	0.0085	0.0006	0.0024
σ_μ^2	0.0004	0.0072	0.0004	0.0067	0.0004	0.0034	0.0004	0.0098	0.0006	0.0037
σ_d^2	0.0030	0.0115	0.0031	0.0173	0.0044	0.0391	0.0040	0.0117	0.0067	0.1130
σ_b^2	0.0035	0.0106	0.0039	0.0116	0.0047	0.0314	0.0047	0.0103	0.0081	0.0162
$\sigma_{g\mu}$	0.0001	0.0000	0.0002	0.0005	0.0001	0.0001	0.0001	0.0005	0.0002	0.0002
σ_{gd}	0.0001	0.0000	0.0000	0.0004	0.0000	0.0004	0.0000	0.0002	0.0000	0.0001
$\sigma_{\mu d}$	0.0002	0.0000	0.0002	0.0005	0.0001	0.0002	0.0002	0.0013	0.0003	0.0007
σ_e^2	0.0009	0.0011	0.0009	0.0012	0.0012	0.0019	0.0012	0.0010	0.0056	0.0958

We present the estimation results of the present-value model in Eqs. (16) to (18) for the five real-world datasets. The model is estimated according to the procedure described in Section 4. Note that $1/\rho$ is the autoregressive parameter of the bubble process. When ρ is less than one the bubble is explosive, when it is greater than one the bubble is stable. $p_{11} = P(S_t = 1|S_{t-1} = 1)$ and $p_{22} = P(S_t = 2|S_{t-1} = 2)$ are the transition probabilities of the two-regime Markov-switching model.

Table 7
95% credible intervals – Real-world datasets.

	United States		United Kingdom		Malaysia		Japan		Brazil	
ρ_1	0.8254	0.9977	0.8199	0.9918	0.7881	1.2441	0.8591	0.9999	0.7756	0.9960
ρ_2	1.1096	1.3766	1.1010	1.3746	1.0930	1.4092	1.1456	1.3571	1.0323	1.3682

We present the 95% credible intervals for the bubble parameter ρ in Regime 1 (ρ_1) and Regime 2 (ρ_2). The 95% credible intervals consists of the 2.50% and 97.50% quantiles of the posterior distribution of ρ_1 and ρ_2.

6.1. Bubble identification

In this section, we analyze the smoothed surviving-probabilities $P(S_t = 1|\Psi_T)$ that the bubble process has been in Regime 1 at time t, $(t = 1, ..., T)$, in order to distinguish between dates on which either the surviving bubble Regime 1 or the collapsing bubble Regime 2 has been in force. The business-cycle literature generally suggests using the following decision rule: Regime 1 has been in force if $P(S_t = 1|\Psi_T) > 0.50$, while Regime 2 has been in force if $P(S_t = 1|\Psi_T) \leq 0.50$ (see for instance Goodwin, 1993). However, we expect the bubble process to be in the surviving Regime 1 most of the time, and in the collapsing Regime 2 only for few short periods. This intuition is confirmed by our results which show that the expected duration of the surviving Regime 1 is higher than the expected duration of the collapsing Regime 2, moreover it is confirmed by the findings of Al-Anaswah and Wilfling (2011), and Lammerding et al. (2013). Hence, the threshold value of 0.50 may fail to correctly identify regime switch dates. We follow Al-Anaswah and Wilfling (2011) who suggest to take into consideration the first two moments of $P(S_t = 1|\Psi_T)$. They adopt the threshold value $m - 2sd$, where m and sd are respectively the sample mean and sample standard deviation of $\{P(S_t = 1|\Psi_T)\}_{(t=1,...,T)}$.

Figs. 3 and 4 display the time series of the smoothed surviving-probabilities and the log-price-dividend for the artificial datasets, shaded areas denote bubble collapsing regime dates. Fig. 3 refers to the artificial datasets 1–4. For datasets 1 and 2 our procedure is able to identify eleven out of twelve bubbles. In dataset 3 (third panel Fig. 3) we identify twelve out of fourteen bubbles. The bottom panel of Fig. 3 refers to dataset 4, in this case all the bubbles are detected correctly, and in dataset 5 (Fig. 4) we can identify twelve out of thirteen bubbles.

Summing up, as can be seen from the figures above, our methodology correctly identifies 57 out of 62 (92.27%) of all the bubble collapsing dates in the five artificial datasets. We also observe that our procedure may fail to recognize bubbles of smaller size, in particular we may fail to identify those bubbles which emerge after a bubble of a bigger size. Further, our procedure never signals a bubble which has no counterpart in the price process.

We now turn our attention to the results for the real-world datasets. Fig. 5 graphs the smoothed surviving-probabilities and the log-price-dividend ratio for the US. We identify only one collapsing period from October 2008, that is after the collapse of the investment bank Lehman Brothers, until May 2009. The smoothed surviving-probabilities slightly decrease in two episodes at the beginning of the series and again after the 2008 collapse, however the decrease is not sharp enough to be signaled as a bubble collapse. Consistent with Balke and Wohar (2009), we cannot say that the decline in stock prices in the 2000s has been caused by a bubble bursting. Similar comments apply to the United Kingdom (Fig. 6) for which the series of the price-dividend ratio is strongly correlated with that of the US.

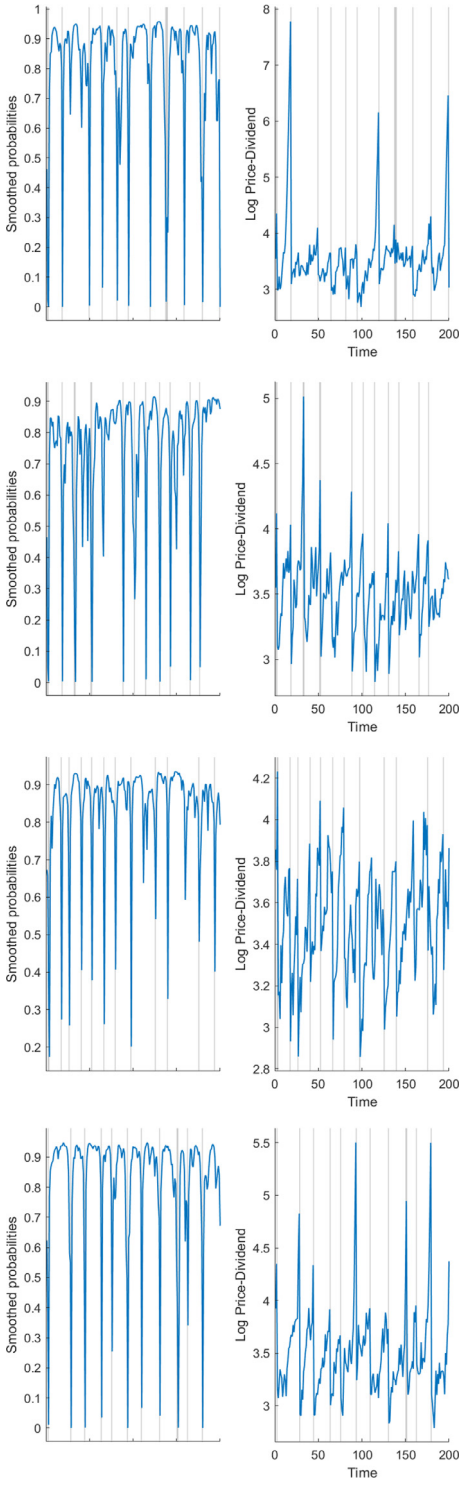

Fig. 3. Smoothed Surviving-probabilities and Log-price-dividend - Artificial datasets 1 to 4.

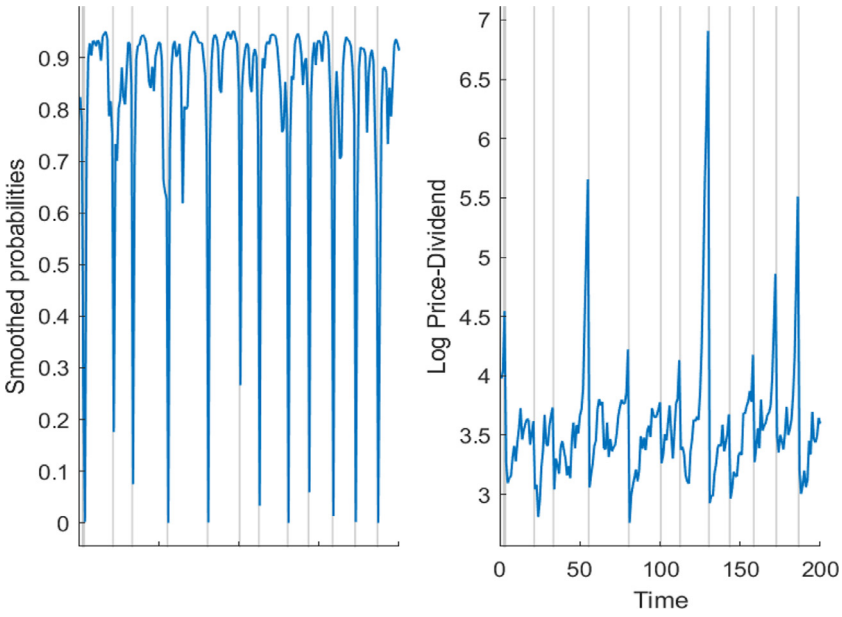

Fig. 4. Smoothed Surviving-probabilities and Log-price-dividend - Artificial dataset 5.

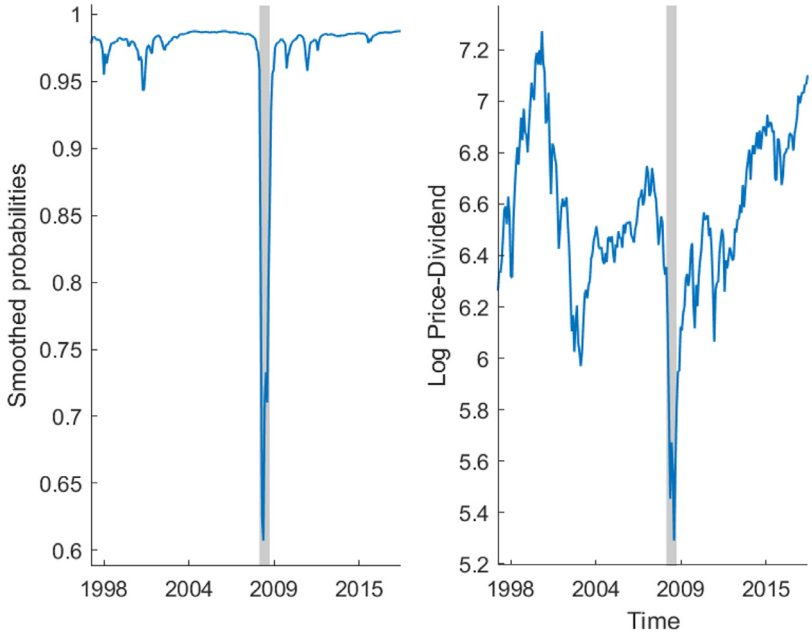

Fig. 5. Smoothed Surviving-probabilities and Log-price-dividend - United States.

For Malaysia (Fig. 7), we observe three clusters of smoothed probabilities indicating collapsing regimes. The first lasts from November 1997 until December 1998, the second from May 1999 until October 1999, and the third is in January 2000 which reflect the Asian financial crisis. By end of 1997, Malaysian ratings fell from investment grade to junk. In January 1998, the Malaysian currency (the ringgit) had already lost 50% of its value to the US dollar. The economy started to recover in 1999.

Japan (Fig. 8) was affected less significantly by the Asian financial crises. The smoothed surviving-probabilities signal collapsing regimes in November 1997 and April 1999. Further, like the US and the United Kingdom, Japan experiences a collapse from October to November 2008 in correspondence of the 2008 global financial crises.

Fig. 9 display the results for Brazil. The smoothed probabilities exhibit some clusters signaling a collapsing regime from November 1997 until April 1998, from July 1998 until September 1998, and from December 1999 until January 2000. These

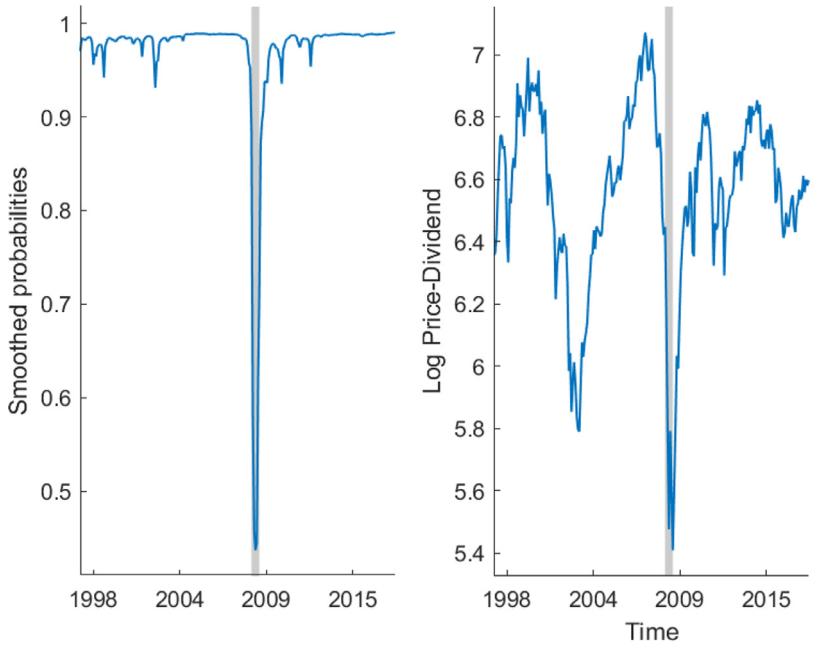

Fig. 6. Smoothed Surviving-probabilities and Log-price-dividend - United Kingdom.

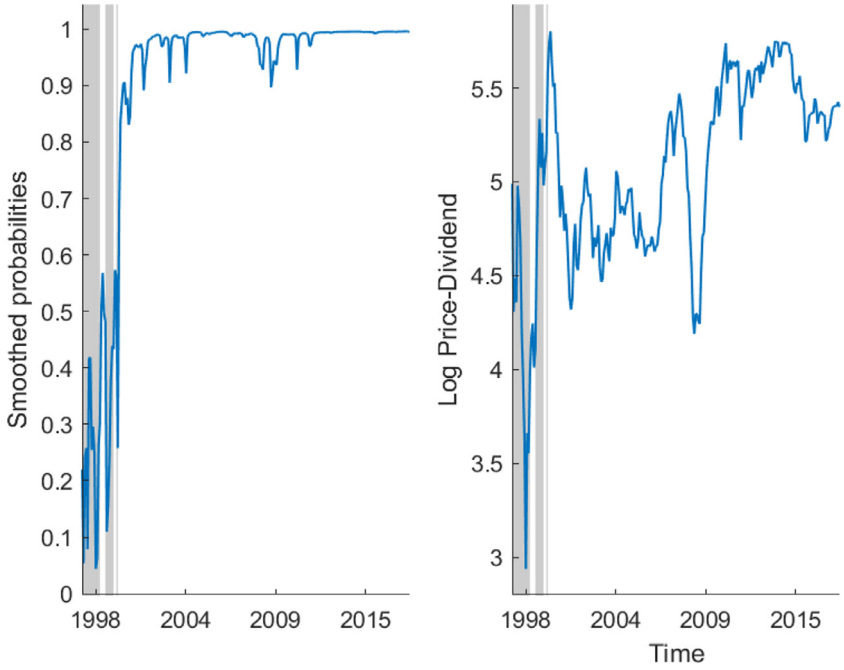

Fig. 7. Smoothed Surviving-probabilities and Log-price-dividend - Malaysia.

episodes reflect the Brazilian stock-market and currency crisis which culminated in a sharp devaluation of the Brazilian currency (the real) to the US dollar in 1999. The effect was caused by the 1997 Asian financial crisis which led Brazil to increase interest rates and to institute spending cuts and tax increases in an attempt to maintain the value of its currency. The devaluation also precipitated fears that the ongoing economic crisis in Asia would spread to South America, as many South American countries were heavily dependent on industrial exports from Brazil. We register also a drop in the smoothed surviving-probabilities in May 2001 in correspondence of the Brazil energy crises, and in November 2008, one month after the US, the United Kingdom and Japan collapse for the 2008 global financial crises. However, we do not detect a collapsing regime associated with the 2014 Brazilian economic crises.

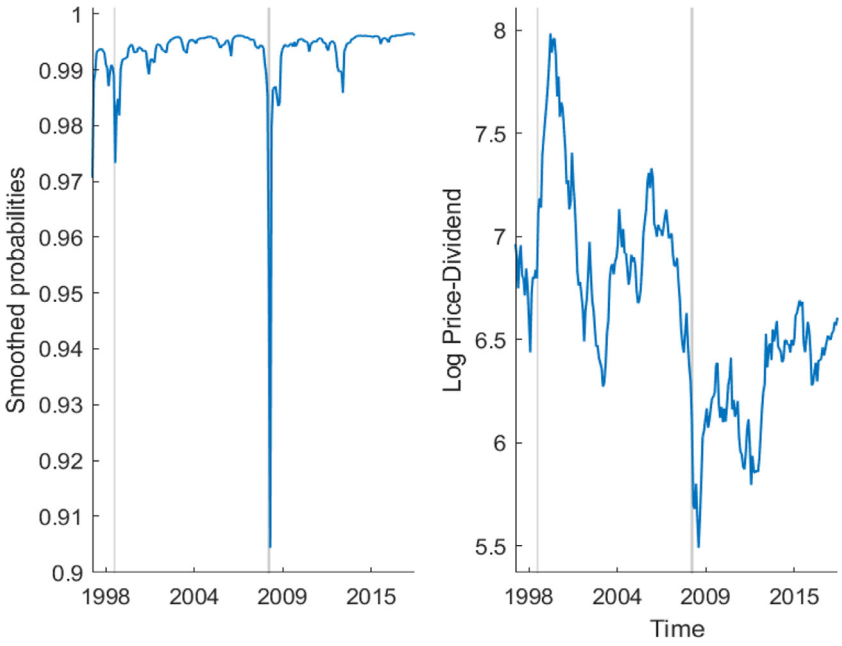

Fig. 8. Smoothed Surviving-probabilities and Log-price-dividend - Japan.

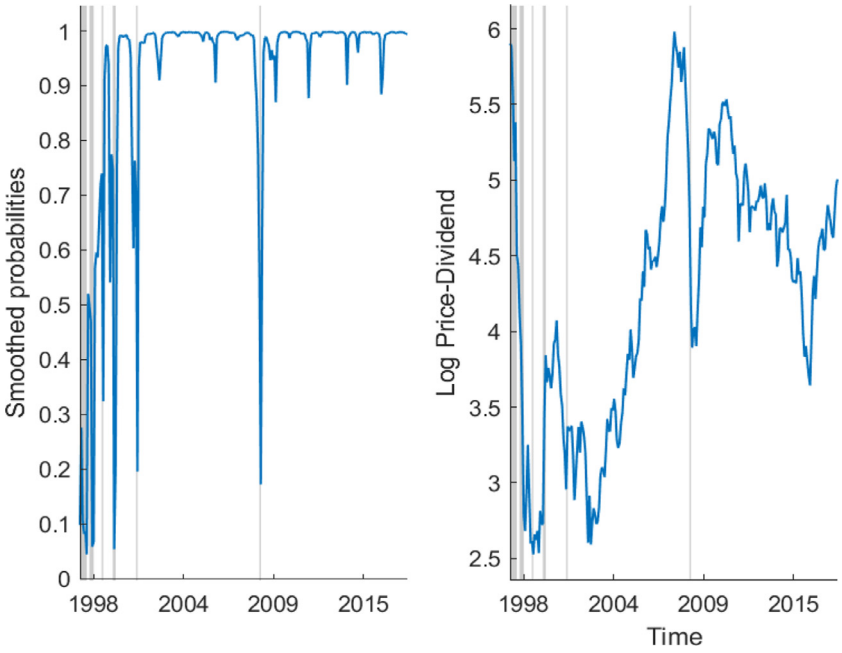

Fig. 9. Smoothed Surviving-probabilities and Log-price-dividend - Brazil

6.2. Prediction of returns and dividend growth rates

We now investigate the in-sample predictability of our state space model. We use the estimated series of expected dividend growth rates ($\hat{\bar{g}}_t$) and expected returns ($\hat{\bar{\mu}}_t$) as if they were observables and we regress them on realized dividend growth rates and realized returns.

$$\begin{aligned} \Delta d_{t+1} &= \alpha_d + \beta_d \hat{\bar{g}}_t + \epsilon^{d*}_{t+1}, \\ r_{t+1} &= \alpha_r + \beta_r \hat{\bar{\mu}}_t + \epsilon^{r*}_{t+1}. \end{aligned} \tag{33}$$

Table 8
Regression results - Artificial datasets.

	Dataset 1		Dataset 2		Dataset 3		Dataset 4		Dataset 5	
	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}
Constant	0.0473**	0.1948**	0.001	0.1415**	0.0021	0.1148**	0.0381**	0.0777**	0.0586**	0.0571
	(0.0027)	(0.0612)	(0.0029)	(0.0221)	(0.0031)	(0.0179)	(0.0028)	(0.0303)	(0.0031)	(0.0464)
$\hat{\bar{g}}_t$	2.0699**		2.0522**		1.8021**		1.8918**		2.0773**	
	(0.0785)		(0.0763)		(0.0686)		(0.0695)		(0.0873)	
$\hat{\bar{\mu}}_t$		6.8698**		6.5605**		3.6952**		3.5828**		4.4319**
		(1.5246)		(0.9837)		(0.5139)		(1.1197)		(1.5355)
R^2	**0.7783**	**0.093**	**0.7853**	**0.1834**	**0.7771**	**0.2071**	**0.7889**	**0.0492**	**0.7407**	**0.0404**
Adj.R^2	0.7772	0.0884	0.7842	0.1793	0.776	0.2031	0.7879	0.0444	0.7394	0.0355

We report regression results for respectively dividend growth rates and returns on their estimated expected values represented by $\hat{\bar{g}}_t$ and $\hat{\bar{\mu}}_t$. Standard errors are reported in parenthesis. Note:** $p \leq 0.05$, * $p \leq 0.1$.

Table 9
Regression results - Real-world datasets.

	United States		United Kingdom		Malaysia		Japan		Brazil	
	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}	Δd_{t+1}	r_{t+1}
Constant	0.0041**	0.0062**	0.0069**	0.0032	0.0021	0.0099*	0.0004	0.0026	0.0136**	0.024**
	(0.0017)	(0.0031)	(0.002)	(0.0033)	(0.0033)	(0.0052)	(0.002)	(0.0034)	(0.0057)	(0.0066)
$\hat{\bar{g}}_t$	3.8308**		3.9518**		4.9302**		4.4305**		6.4918**	
	(0.1606)		(0.1851)		(0.2481)		(0.1896)		(0.4284)	
$\hat{\bar{\mu}}_t$		0.2511		0.2636		-0.0834		0.0901		0.5914**
		(0.1722)		(0.1784)		(0.3721)		(0.1468)		(0.2136)
R^2	**0.7049**	**0.0089**	**0.6569**	**0.0091**	**0.6239**	**0.0002**	**0.6964**	**0.0016**	**0.4910**	**0.0312**
Adj.R^2	0.7037	0.0047	0.6554	0.0049	0.6223	-0.004	0.6951	-0.0026	0.4889	0.0271

We report regression results for respectively dividend growth rates and returns on their estimated expected values represented by $\hat{\bar{g}}_t$ and $\hat{\bar{\mu}}_t$. Standard errors are reported in parenthesis. Note:** $p \leq 0.05$, * $p \leq 0.1$.

Table 8 shows the results for the artificial datasets, for dividend growth rates the R^2 ranges from 74.07% to 78.89% while for returns it is between 4.04% and 20.71%.

Expected dividend growth rates explains a large fraction of actual dividend growth, while the fraction of explained return variability is lower. Also, given that the goodness-of-fit measures of returns vary substantially in the five datasets, we argue that the return predictability features of this model are less robust with respect to the dividend predictability features.

Figs. 10, and 11 plot the realized and estimated expected dividend growth rates on the left, and the realized and estimated expected returns on the right for the five artificial datasets. The realized and expected dividend growth series are strongly correlated, however the latter series is more stable. Concerning returns, their expectation is less volatile than the series of realized returns. Further, the correlation between realized and expected returns is lower than that for dividend growth rates.

Table 9 reports the results for the real-world datasets. The R^2 values for dividend growth rates are quite high in all the datasets, the highest value is recorded for the US where it is equal to 70.49% while the lowest value is registered for Brazil where it is equal to 49.10%.[11] Instead, the coefficient β_r in the regressions for returns is significant only for Brazil and the goodness-of-fit measure is equal to 3.12%. In the other countries the R^2 values for returns are less than 1% meaning that the estimated series of expected returns ($\hat{\bar{\mu}}_t$) do not help in predicting stock returns.

Visual inspection of Figs. 12–16 confirms what we have observed in the regression results.

6.3. Variance decomposition

In this section we derive the variance decomposition of both the price-dividend ratio and unexpected returns. The variance decomposition of the price-dividend ratio is given by

$$\begin{aligned} Var(pd_t) \quad &= B_2^2 Var(g_t) + B_1^2 Var(\mu_t) + Var(b_t) - 2B_1 B_2 Cov(g_t \mu_t) - 2B_1 Cov(b_t \mu_t) + \\ &\quad 2B_2 Cov(g_t b_t). \end{aligned} \tag{34}$$

The first term, $B_2^2 Var(g_t)$, represents the variation in the price-dividend ratio due to expected dividend growth rate variation. The second term, $B_1^2 Var(\mu_t)$, measures the variation in the price-dividend ratio due to expected return variation. The third

[11] Since the dividend growth predictability we find may be driven by the seasonality in dividend payments, we apply a stable seasonal filter to estimate the seasonal component in the series of dividend growth rates and expected dividend growth rates. Then we estimate the regression equation using the deseasonalized series. The dividend growth predictability is confirmed and the results are available upon request.

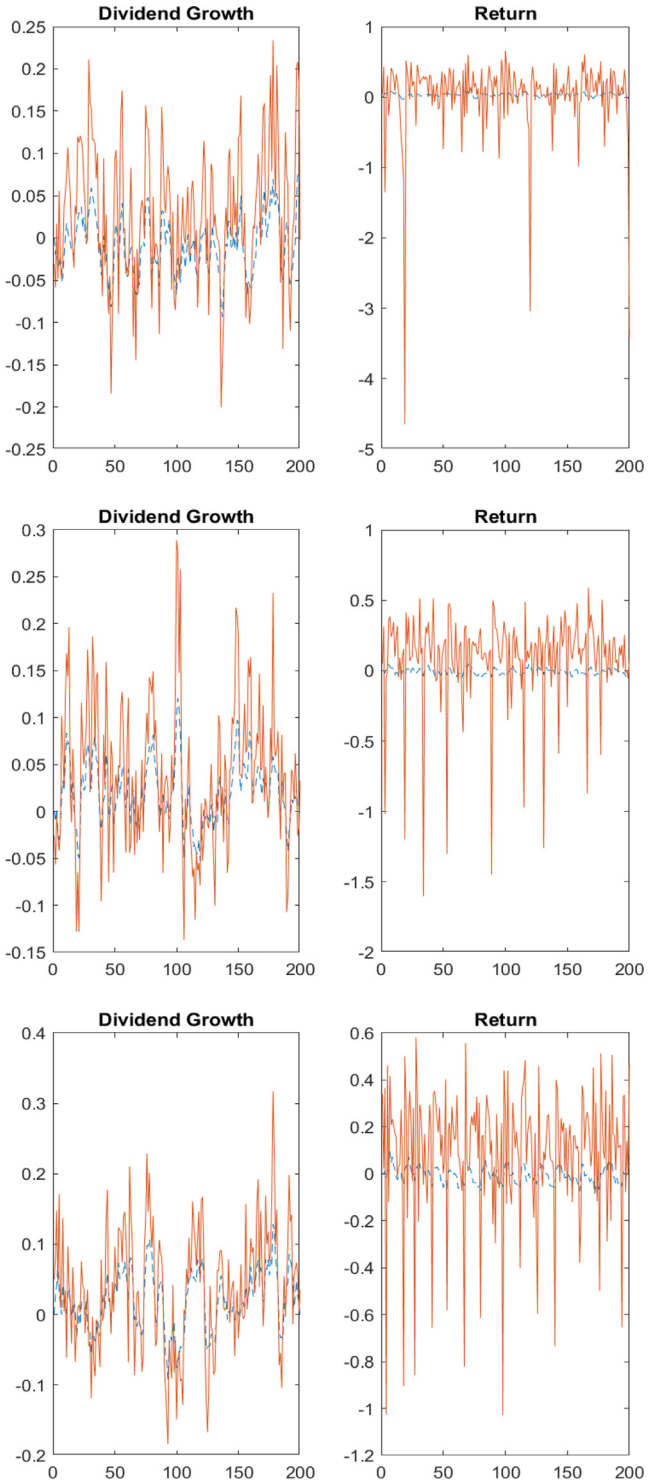

Fig. 10. Realized and Expected series Artificial datasets 1 to 3. The solid lines represent the realized time series of dividend growth rates (left) and returns (right), while the dashed lines represent the estimated series of expected dividend growth rates (left) and expected returns (right).

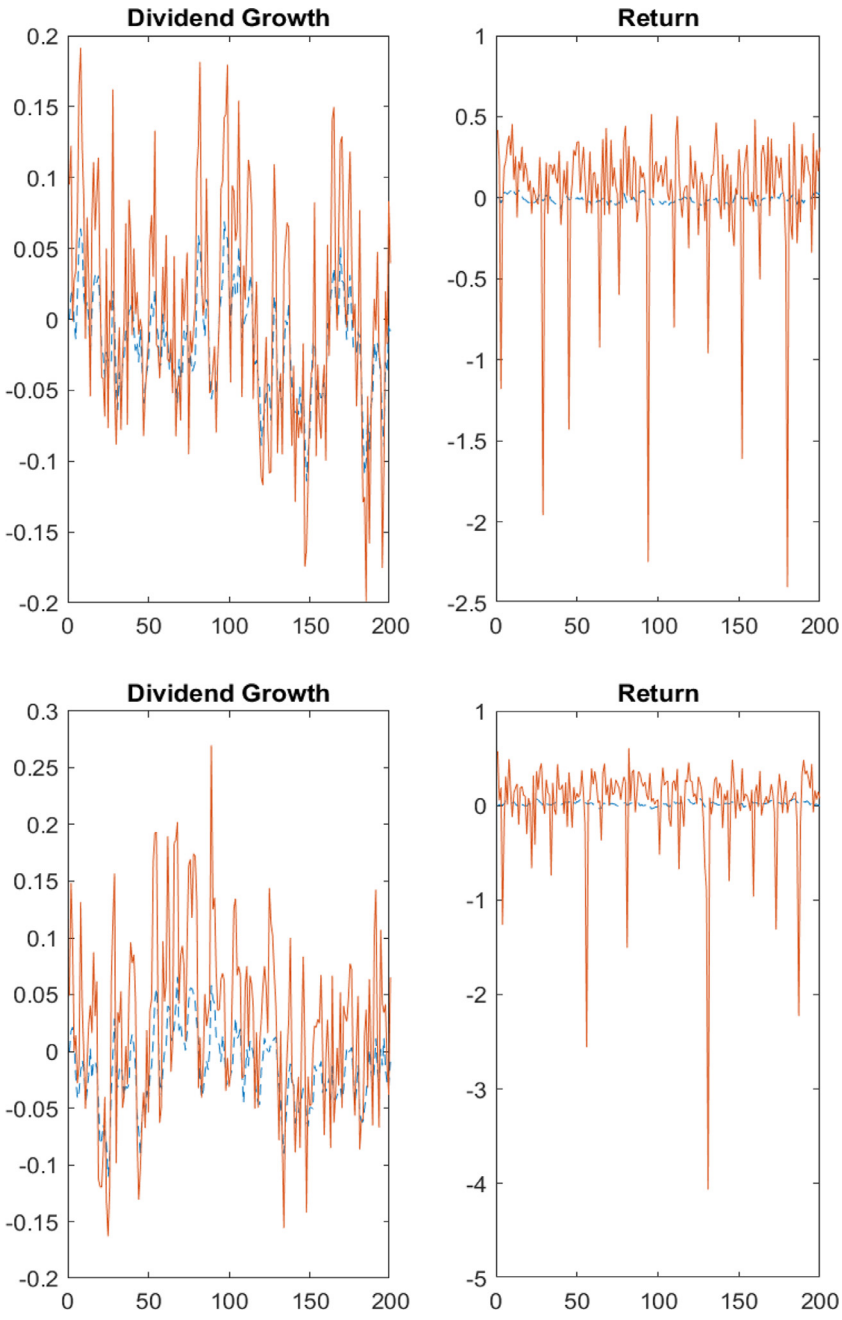

Fig. 11. Realized and Expected series - Artificial datasets 4 to 5. The solid lines represent the realized time series of dividend growth rates (left) and returns (right), while the dashed lines represent the estimated series of expected dividend growth rates (left) and expected returns (right).

term, $Var(b_t)$, accounts for the variation in the price-dividend ratio due to bubble variation. The remaining terms represent the covariation among these three components. The unexpected returns can be written as

$$r_{t+1} - \mu_t = \rho B_2 \epsilon^g_{t+1} - \rho B_1 \epsilon^\mu_{t+1} + \rho \epsilon^b_{t+1} + \epsilon^d_{t+1}. \tag{35}$$

As before, we decompose the unexpected return variation into the influence of dividend growth variation, discount rate variation, bubble variation and the covariance among these terms.

Table 10 summarizes the results for the variance decomposition of the price dividend ratio and unexpected returns for the real-world datasets. We use sample covariances and we standardize all terms so that the sum is equal to 100%. We find that in the surviving bubble regime, most of the variation in price-dividend ratio is related to the bubble variation. Specifically, bubble variation accounts for more than 50% of the price-dividend variation in all the countries under study with the

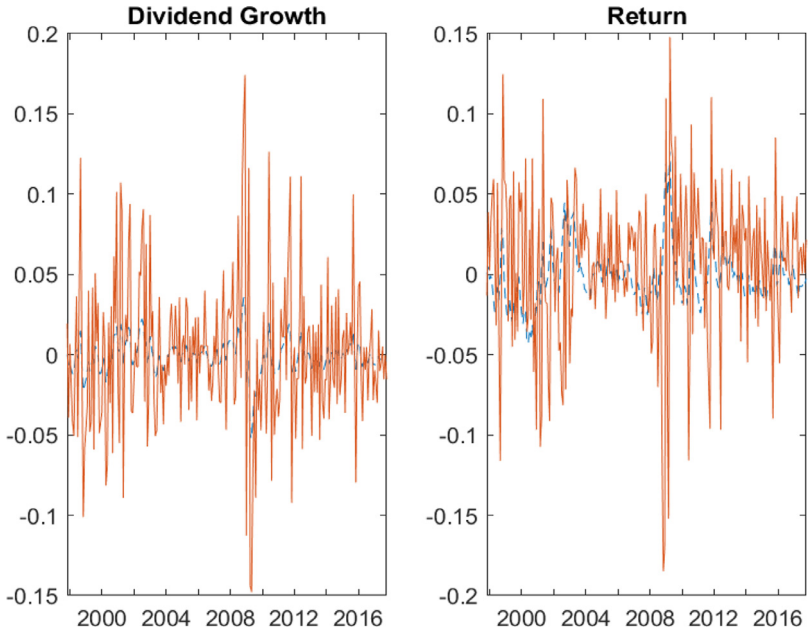

Fig. 12. Realized and Expected series - United States. The solid lines represent the realized time series of dividend growth rates (left) and returns (right), while the dashed lines represent the estimated series of expected dividend growth rates (left) and expected returns (right).

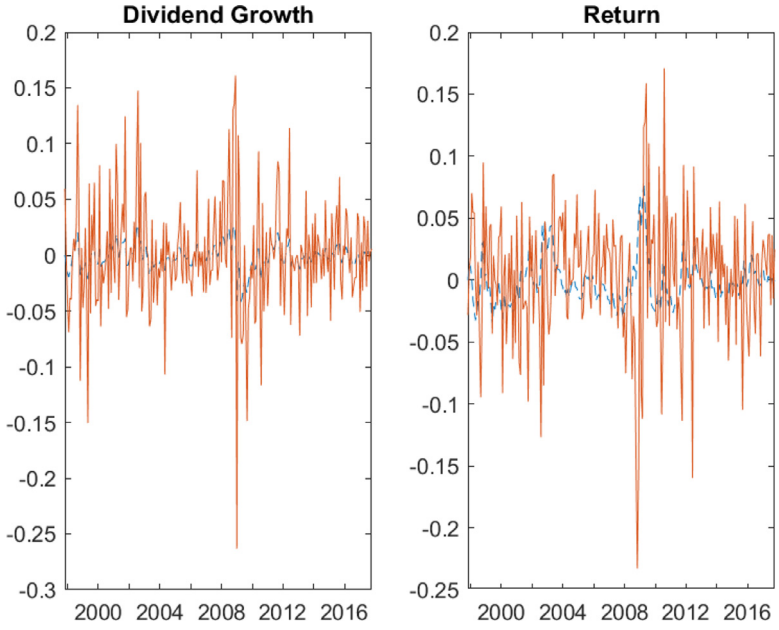

Fig. 13. Realized and Expected series - United Kingdom

exception of Brazil where it accounts for about 36%. Instead, consistent with Binsbergen and Koijen (2010), in the collapsing bubble regime discount rate variation accounts for most of the variation in the price-dividend ratio. Again, consistent with Binsbergen and Koijen (2010) we document that dividend growth plays a major role in explaining unexpected returns variation in the surviving bubble regime, while discount rate variation accounts most of the variation in unexpected returns in the collapsing bubble regime with the exception for Brazil. Also, bubble variation explains a large share of unexpected return variation in the surviving bubble regime. These results are consistent with Balke and Wohar (2009) which find that the bubble component is substantially important in explaining fluctuations in the log price-dividend ratio when there are no permanent components in market fundamentals.

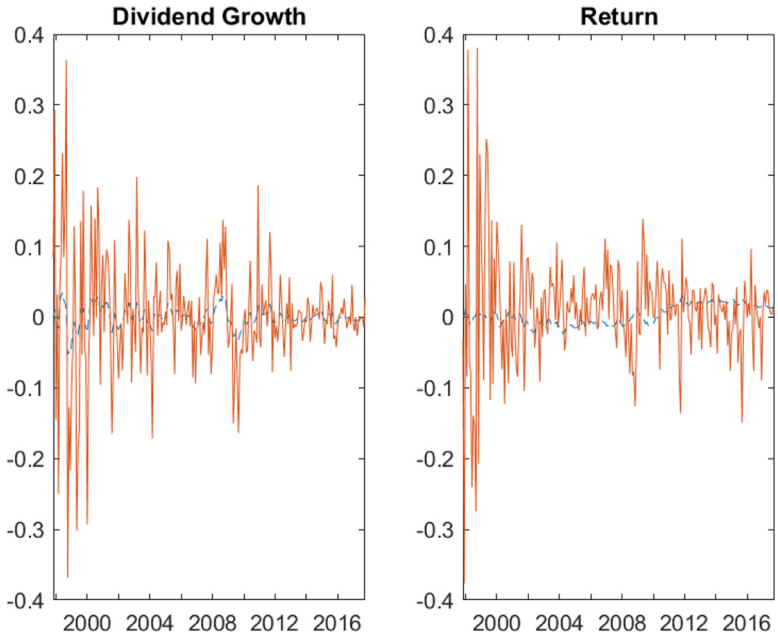

Fig. 14. Realized and Expected series - Malaysia

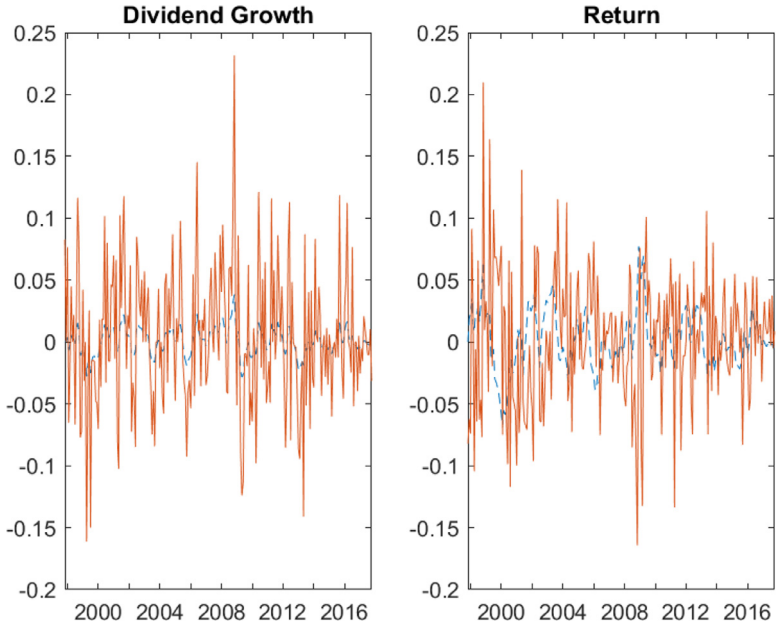

Fig. 15. Realized and Expected series - Japan The solid lines represent the realized time series of dividend growth rates (left) and returns (right), while the dashed lines represent the estimated series of expected dividend growth rates (left) and expected returns (right).

7. Conclusions

We have shown that in the surviving bubble regime, most of the variation in the price-dividend ratio is related to the bubble variation. Specifically, bubble variation accounts for more than 50% of the price-dividend variation in all the countries under study with the exception of Brazil where it accounts for about 36%. Further, bubble variation explains also a large share of unexpected return variation in the surviving bubble regime.

These results suggest that present-value models should not ignore the bubble component of stock prices. This paper proposes to incorporate a speculative bubble subject to a surviving and a collapsing regime into the present-value model by

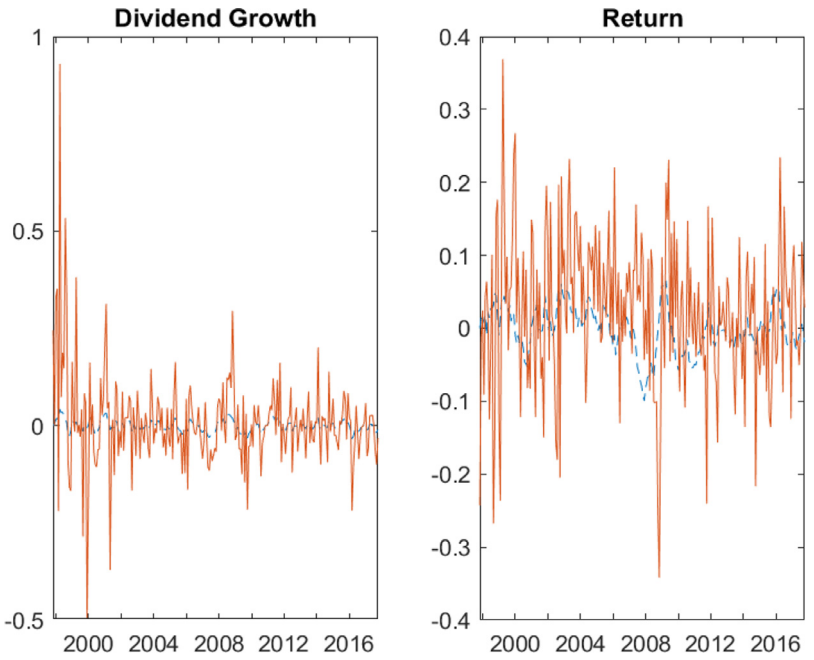

Fig. 16. Realized and Expected series - Brazil The solid lines represent the realized time series of dividend growth rates (left) and returns (right), while the dashed lines represent the estimated series of expected dividend growth rates (left) and expected returns (right).

Table 10
Variance decomposition – Real-world datasets.

	United States		United Kingdom		Malaysia		Japan		Brazil	
Regimes	1	2	1	2	1	2	1	2	1	2
Panel A: Price-dividend ratio										
Discount rate	7.5399	96.2946	7.8771	88.2220	45.2701	84.2113	10.7564	99.9907	18.2560	50.4675
Div. growth	0.8436	0.4624	1.0325	0.9127	2.6420	1.0683	0.5748	0.0028	0.8314	2.3222
Bubble	63.4978	0.8138	52.7685	3.2424	56.4073	10.1861	51.2994	0.0064	35.9414	27.2425
Covariances	28.1188	2.4292	38.3219	7.6229	4.3193	4.5343	37.3695	0.0002	44.9711	19.9677
Panel B: Unexpected returns										
Discount rate	124.1429	100.6786	82.1027	121.5784	80.4225	93.3425	111.9314	99.9597	79.6270	234.0628
Div. growth	171.6528	0.7922	150.0561	12.2131	267.8176	4.5416	160.3030	0.0015	201.7372	424.8113
Bubble	94.6137	0.4122	92.0662	1.2156	103.1748	1.8223	95.5578	0.0016	58.0041	13.6928
Covariances	−290.4094	−1.8830	−224.2250	−35.0071	−351.4148	0.2935	−267.7922	0.0372	−239.3683	−572.5669

We report the decomposition of the price-dividend variation and the unexpected returns variation into discount rate variation, dividend growth variation, bubble variation and the covariances among these three terms.

Table 11
Datastream Country Indices.

Country	Time span
United States	February 1973–October 2017
United Kingdom	January 1965–October 2017
Malaysia	February 1986–October 2017
Japan	January 1973–October 2017
Brazil	August 1994 - October 2017

This Table reports the time span of the country indices available on Datastream.

Binsbergen and Koijen (2010), who adopts a latent variables approach to estimate expected returns and expected dividend growth rates. Further, we suggest an econometrically robust Bayesian MCMC methodology to estimate our model.

This study applies our two-regime Markov-switching state space model to artificial as well as real-world datasets. The artificial bubble processes are defined in the sense of Evans (1991), we show that our bubble-detection methodology is able to identify 92.27% of all the bubble collapsing dates in the artificial datasets, moreover it never signals a bubble which has no counterpart in the price process. These results represent an improvement with respect to the approach discussed in Al-Anaswah and Wilfling (2011) which correctly identifies around 50% of all the bubble episodes. The real-world datasets

Table 12
Autocorrelation of Posterior Draws - No Thinning.

	$i+1$		$i+10$		$i+100$		Inefficiency Factors	
Regimes	1	2	1	2	1	2	1	2
γ_0	0.3319	0.2981	0.1597	0.1507	0.0969	0.0882	339.8204	281.9305
δ_0	0.3341	0.3010	0.1634	0.1532	0.0983	0.0889	373.7996	300.7588
ρ	0.9331	0.9451	0.5648	0.6362	0.0992	0.1366	329.7122	649.0314
γ_1	0.7376	0.7431	0.0875	0.1381	0.0208	0.0301	20.0032	19.4034
δ_1	0.8884	0.9740	0.4278	0.8062	0.1347	0.3214	357.6982	1160.8113
σ_g^2	0.0782	0.0398	0.0034	0.0087	0.0005	0.0021	1.1804	1.0529
σ_μ^2	0.5578	0.0975	0.2621	0.0194	0.0160	0.0027	18.5462	1.4797
σ_d^2	0.5842	0.1167	0.2849	0.0500	0.1949	0.0293	832.6511	32.4101
σ_b^2	0.8850	0.7013	0.3447	0.2434	0.0037	0.0191	19.5314	18.1775
$\sigma_{g\mu}$	0.0570	0.0297	0.0006	0.0008	0.0002	0.0005	1.0668	1.0275
σ_{gd}	0.0784	0.0375	0.0008	0.0017	0.0004	0.0005	1.1569	1.0532
$\sigma_{\mu d}$	0.2826	0.0592	0.0263	0.0079	0.0024	0.0021	2.6347	1.1085
σ_e^2	0.9284	0.8122	0.5055	0.3544	0.0344	0.0388	38.8663	37.3296

We present the average autocorrelation of the posterior draws of the model in Eqs. (16) to (18) for the artificial datasets. The model is estimated according to the procedure described in Section 4. The last two columns report the inefficiency factors.

Table 13
Autocorrelation of Posterior Draws - Thinning every 100 draws.

	$i+1$		$i+10$		$i+100$		Inefficiency factors	
Regimes	1	2	1	2	1	2	1	2
γ_0	0.0930	0.0946	0.0433	0.0385	0.0010	0.0065	4.4284	3.7236
δ_0	0.0945	0.1010	0.0488	0.0364	0.0102	0.0025	4.6193	3.7973
ρ	0.1024	0.1385	0.0412	0.0374	0.0020	0.0204	4.1054	6.6608
γ_1	0.0219	0.0226	0.0063	0.0097	0.0024	0.0140	1.1690	1.1190
δ_1	0.1327	0.3178	0.0416	0.0577	0.0085	0.0282	4.3552	11.8271
σ_g^2	0.0009	0.0012	0.0068	0.0162	0.0017	0.0043	1.0000	1.0000
σ_μ^2	0.0188	0.0005	0.0023	0.0000	0.0086	0.0003	1.0228	1.0000
σ_d^2	0.1876	0.0396	0.0709	0.0157	0.0135	0.0001	9.2677	1.4590
σ_b^2	0.0034	0.0202	0.0011	0.0097	0.0006	0.0073	1.0000	1.0245
$\sigma_{g\mu}$	0.0082	0.0009	0.0027	0.0045	0.0076	0.0017	1.0000	1.0000
σ_{gd}	0.0120	0.0051	0.0027	0.0071	0.0040	0.0009	1.0000	1.0000
$\sigma_{\mu d}$	0.0012	0.0055	0.0055	0.0062	0.0047	0.0028	1.0000	1.0000
σ_e^2	0.0268	0.0364	0.0117	0.0071	0.0094	0.0102	1.0794	1.1165

We present the average autocorrelation of the posterior draws with thinning every 100 draws of the model in Eqs. (16) to (18) for the artificial datasets. The model is estimated according to the procedure described in Section 4. The last two columns report the inefficiency factors.

consist of the monthly time series data for the price index, the dividend yield and the market value for the United States, United Kingdom, Malaysia, Japan and Brazil. We find that our framework is able to identify most of the bubble periods classified as such by Kindleberger and Aliber (2003).

In line with Al-Anaswah and Wilfling (2011) and Lammerding et al. (2013), we document the existence of statistically significant Markov-switching in the data-generating process of real-world stock price bubbles. Furthermore, our methodology is also able to predict dividend growth rates as well as returns with R^2 values ranging from 74.07% to 78.89% for dividend growth rates and 4.04% and 20.71% for returns in the artificial datasets. In the real-world datasets, we find that dividend growth rates are predictable with R^2 values ranging from 70.49% for the US to 49.10% for Brazil. However, the R^2 values for returns are less than 1% with the exception of Brazil where it is above 3%.

A common drawback of the bubble literature is that rejection of the present-value model that are interpreted as evidence of the presence of bubbles can still be explained by alternative structures of the fundamentals. In this paper we mitigate this issue in two ways. First, we restrict our analysis to rational bubbles which impose fairly strong restrictions on the dynamics of the bubble component. Hence these restrictions can help us to identify the non-fundamental component in the data. Second, we use a less restrictive fundamentals model, indeed our econometric procedure allows us to analyse a more complex model with time-varying discount rates and regime-switching in fundamentals and the bubble. Doing so, we allow the fundamentals part to fit better the data, leaving less room for a bubble.

In sum, our setup allows to model jointly expected dividend growth rates, expected returns and the bubble component of stock prices. As such it may improve conventional methods for the detection of real-time stock-price bubbles allowing an early detection of future bubbles. Moreover, this methodology allows for hypothesis testing of some features of expected dividend growth rates and expected returns such as their persistence.

Declaration of Competing Interest

None.

Acknowledgement

We thank the editor and two anonymous referees for their valuable suggestions and comments. We would like also to thank Diego Amaya, Fulvio Corsi, Tim Kroencke, Mathijs van Dijk, the participants to the 2019 European and Latin America meetings of the Financial Management Association, and to the 2018 Simposio de Analisis Economico. The usual disclaimer applies.

Appendix A. Datastream data

Datastream country indices include a representative list of stocks for each country. The sample covers a minimum of 75–80% of total market capitalisation. Suitability for inclusion is determined by market value and availability of data. Table 11 reports the time span available on Datastream for the country indices used. The aggregate price index (PI), dividend yield (DY), and market value (MV) for each country index are calculated as follows:

$$\begin{aligned} PI_t &= PI_{t-1} * \frac{\sum_{i=1}^{M} P_{i,t} N_{i,t}}{\sum_{i=1}^{M} P_{i,t-1} N_{i,t} f}, \; PI_0 = 100, \\ DY_t &= \frac{\sum_{i=1}^{M} D_{i,t} N_{i,t}}{\sum_{i=1}^{M} P_{i,t} N_{i,t}} * 100, \\ MV_t &= \sum_{i=1}^{M} P_{i,t} N_{i,t}, \end{aligned} \tag{36}$$

where $P_{i,t}$ is the unadjusted price of asset i in month t, N_t is the number of shares in issue on month t, f adjustment factor for capital actions, $D_{i,t}$ is dividend per share of asset i in month t, and M is the number of constituents in index. We use the above variables to compute the log price-dividend ratio (pd_t), dividend growth rate (Δd_t) and returns (r_t).

Appendix B. Efficiency of the metropolis-within-Gibbs sampler

Since the Metropolis-within-Gibbs sampler used to simulate from the joint posterior distribution may produce parameter draws that are highly autocorrelated, in Table 12 we report the average autocorrelation of the ith parameter draw with the $(i+1)$th draw, $(i+10)$th draw, and $(i+100)$th draw. To further assess the efficiency of the sampler, in the last two columns we report the average inefficiency factors of the posterior draws which are defined as follow

$$1 + 2\sum_{l=1}^{L} Corr(\theta_i, \theta_{i+l}), \tag{37}$$

where $Corr(\theta_i, \theta_{i+l})$ is the sample autocorrelation of parameter θ at lag length l, L is the maximum lag and it is chosen to be large enough so that the autocorrelation tapers off. The inefficiency factor measures the number of extra draws needed to obtain results equivalent to the ideal case of independent draws. Consider for example an inefficiency factor of 50, this means that around 5000 posterior draws are needed to have the same information of 100 independent draws. The inefficiency factor of independent draws is one.

From Table 12 we can observe that some of the parameters show high autocorrelations, with large inefficiency factors as well. The maximum inefficiency factor is reported for the parameter δ_1 in regime 2 and it is equal to 1160.8113.

To improve the efficiency of the sampler we perform thinning of the posterior draws and keep only every 100th draws. Table 13 reports the average autocorrelation of the posterior draws and the inefficiency factors after performing thinning of the draws. We can see that thinning significantly reduce the inefficiency factors.

References

Al-Anaswah, N., Wilfling, B., 2011. Identification of speculative bubbles using state-space models with Markov-switching. J. Bank. Financ. 35 (5), 1073–1086.
Asako, Y., Funaki, Y., Ueda, K., Uto, N., 2020. (a) symmetric information bubbles: experimental evidence. J. Econ. Dyn. Control 110, 103744.
Balke, N.S., Wohar, M.E., 2009. Market fundamentals versus rational bubbles in stock prices: a Bayesian perspective. J. Appl. Econ. 24 (1), 35–75.
Binsbergen Van Jules, H., Koijen, R.S., 2010. Predictive regressions: a present-value approach. J. Financ. 65 (4), 1439–1471.
Bohl, M.T., 2003. Periodically collapsing bubbles in the us stock market? Int. Rev. Econ. Financ. 12 (3), 385–397.
Bohl, M.T., Siklos, P.L., 2004. The present value model of US stock prices redux: a new testing strategy and some evidence. Q. Rev. Econ. Financ. 44 (2), 208–223.
Brooks, C., Katsaris, A., 2005. A three-regime model of speculative behaviour: modelling the evolution of the s&p 500 composite index. Econ. J. 115 (505), 767–797.
Campbell, J.Y., Cochrane, J.H., 1999. By force of habit: a consumption-based explanation of aggregate stock market behavior. J. Polit. Econ. 107 (2), 205–251.
Campbell, J.Y., Shiller, R.J., 1988. The dividend-price ratio and expectations of future dividends and discount factors. Rev. Financ. Stud. 1 (3), 195–228.
Cerqueti, R., Costantini, M., 2011. Testing for rational bubbles in the presence of structural breaks: evidence from nonstationary panels. J. Bank. Financ. 35 (10), 2598–2605.
Chan, J.C., 2013. Moving average stochastic volatility models with application to inflation forecast. J. Econ. 176 (2), 162–172.
Chan, J.C., Jeliazkov, I., 2009. Efficient simulation and integrated likelihood estimation in state space models. Int. J. Math. Model. Numer. Optim. 1 (1–2), 101–120.

Check, A., 2014. A New Test for Asset Bubbles. Technical Report. Working Paper.
Chen, S.-W., Hsu, C.-S., Xie, Z., 2016. Are there periodically collapsing bubbles in the stock markets? New international evidence. Econ. Model. 52, 442–451.
Chib, S., 1996. Calculating posterior distributions and modal estimates in Markov mixture models. J. Econ. 75 (1), 79–97.
Choi, K.H., Kim, C.-j., Park, C., 2017. Regime shifts in price-dividend ratios and expected stock returns: a present-value approach. J. Money Credit Bank. 49 (2–3), 417–441.
Diba, B.T., Grossman, H.I., 1988. Explosive rational bubbles in stock prices? Am. Econ. Rev. 78 (3), 520–530.
Driffill, J., Sola, M., 1998. Intrinsic bubbles and regime-switching. J. Monet. Econ. 42 (2), 357–373.
Evans, G.W., 1991. Pitfalls in testing for explosive bubbles in asset prices. Am. Econ. Rev. 81 (4), 922–930.
Fama, E.F., French, K.R., 1988. Dividend yields and expected stock returns. J. Financ. Econ. 22 (1), 3–25.
Ferson, W.E., Sarkissian, S., Simin, T.T., 2003. Spurious regressions in financial economics? J. Financ. 58 (4), 1393–1413.
Froot, K.A., Obstfeld, M., 1991. Intrinsic bubbles: the case of stock prices. Am. Econ. Rev. 81 (5), 1189–1214.
Fulop, A., Yu, J., 2017. Bayesian analysis of bubbles in asset prices. Econometrics 5 (4), 47.
Goodwin, T.H., 1993. Business-cycle analysis with a Markov-switching model. J. Bus. Econ. Stat. 11 (3), 331–339.
Gürkaynak, R.S., 2008. Econometric tests of asset price bubbles: taking stock. J. Econ. Surv. 22 (1), 166–186.
Hall, S.G., Psaradakis, Z., Sola, M., 1999. Detecting periodically collapsing bubbles: a Markov-switching unit root test. J. Appl. Econ. 14 (2), 143–154.
Hamilton, J.D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. Econometrica 57 (2), 357–384.
Hansen, L.P., Sargent, T.J., 1980. Formulating and estimating dynamic linear rational expectations models. J. Econ. Dyn. Control 2, 7–46.
Jiang, X., Lee, B.-S., 2007. Stock returns, dividend yield, and book-to-market ratio. J. Bank. Financ. 31 (2), 455–475.
Kanas, A., 2005. Nonlinearity in the stock price–dividend relation. J. Int. Money Financ. 24 (4), 583–606.
Kim, C.-J., Nelson, C.R., 1999. State-space Models With Regime Switching: Classical and Gibbs-sampling Approaches With Applications. The MIT press.
Kindleberger, C.P., Aliber, R., 2003. Manias, Panics, and Crashes: a History of Financial Crises. Springer.
Lammerding, M., Stephan, P., Trede, M., Wilfling, B., 2013. Speculative bubbles in recent oil price dynamics: evidence from a Bayesian Markov-switching state-space approach. Energy Econ. 36, 491–502.
Li, C.W., Xue, H., 2009. A Bayesian's bubble. J. Financ. 64 (6), 2665–2701.
McCausland, W.J., Miller, S., Pelletier, D., 2011. Simulation smoothing for state-space models: acomputational efficiency analysis. Comput. Stat. Data Anal. 55 (1), 199–212.
McMillan, D.G., 2007. Bubbles in the dividend–price ratio? evidence from an asymmetric exponential smooth-transition model. J. Bank. Financ. 31 (3), 787–804.
Miao, J., Wang, P., Xu, Z., 2015. A Bayesian DSGE model of stock market bubbles and business cycles. Quant. Econ..
Pástor, L., Stambaugh, R.F., 2009. Predictive systems: living with imperfect predictors. J. Financ. 64 (4), 1583–1628.
Phillips, P.C., Wu, Y., Yu, J., 2011. Explosive behavior in the 1990s nasdaq: when did exuberance escalate asset values? Int. Econ. Rev. 52 (1), 201–226.
Piatti, I., Trojani, F., 2017. Predictable risks and predictive regression in present-value models.
Roberts, G.O., Gelman, A., Gilks, W.R., et al., 1997. Weak convergence and optimal scaling of random walk metropolis algorithms. Ann. Appl. Probab. 7 (1), 110–120.
Roberts, G.O., Rosenthal, J.S., 2009. Examples of adaptive MCMC. J. Comput. Graph. Stat. 18 (2), 349–367.
Roberts, G.O., Rosenthal, J.S., et al., 2001. Optimal scaling for various metropolis-hastings algorithms. Stat. Sci. 16 (4), 351–367.
Sarno, L., Taylor, M.P., 2003. An empirical investigation of asset price bubbles in latin american emerging financial markets. Appl. Financ. Econ. 13 (9), 635–643.
Shi, S., Song, Y., 2014. Identifying speculative bubbles using an infinite hidden Markov model. J. Financ. Econ. 14 (1), 159–184.
Shi, S.-P., 2013. Specification sensitivities in the Markov-switching unit root test for bubbles. Empir. Econ. 45 (2), 697–713.
Shiller, R. J., 1980. Do stock prices move too much to be justified by subsequent changes in dividends?
West, K.D., 1987. A specification test for speculative bubbles. Q. J. Econ. 102 (3), 553–580.
Wu, Y., 1997. Rational bubbles in the stock market: accounting for the us stock-price volatility. Econ. Inq. 35 (2), 309.
Zheng, H., 2020. Coordinated bubbles and crashes. J. Econ. Dyn. Control 103974.