CQE Center for Quantitative Economics

Estimating rational stock-market bubbles with sequential Monte Carlo methods

Benedikt Rotermann und Bernd Wilfling [†]

40/2015

[†] Department of Economics, University of Münster, Germany

wissen.leben
WWU Münster

Estimating rational stock-market bubbles with sequential Monte Carlo methods

Benedikt Rotermann [a], Bernd Wilfling [a*]

[a] *Westfälische Wilhelms-Universität Münster, Department of Economics (CeNoS and CQE), Am Stadtgraben 9, 48143 Münster, Germany*

(Date of this version: May 20, 2015)

Abstract

In the context of the present-value stock-price model, we propose a new rational parametric bubble specification that is able to generate periodically recurring and stochastically deflating trajectories. Our bubble model is empirically more plausible than its predecessor variants and has neatly interpretable parameters. We transform our entire stock-price-bubble framework into a nonlinear state-space form and implement a fully-fledged estimation framework, based on sequential Monte Carlo methods. This particle-filtering approach, originally from the engineering literature, enables us (a) to obtain accurate parameter estimates, and (b) to reveal the (unobservable) trajectories of arbitrary rational bubble specifications. We fit our new bubble process to artificial and real-world data and demonstrate the use of parameter estimates to compare important characteristics of historical bubbles which emerged in different stock markets.

Keywords: Present-value model, rational bubble, nonlinear state-space model, particle-filter estimation, EM algorithm.

JEL classification: C15, C32, C58, G10, G12.

*Corresponding author. Tel.: +49 251 83 25040; fax: +49 251 83 25042.
E-mail addresses: 05bero@wiwi.uni-muenster.de (B. Rotermann), bernd.wilfling@wiwi.uni-muenster.de (B. Wilfling).

1 Introduction

Speculative bubbles in asset markets are of substantial relevance to financial theorists, analysts and policy makers. Owing to their impact on both the monetary and real economy, theoretical and empirical work on speculative bubbles abounds in the economics and finance literature. The majority of these articles discuss rational bubbles arising from asset-valuation models, like the well-known present-value stock-price model. However, most of these studies are concerned primarily with the empirical detection of bubbles in artificial and/or real-world financial-market data, predominantly by applying (fixed-sample and sequential) cointegration and unit-roots tests to a dividend-stock-price relationship (see, *inter alia*, West 1987; Diba and Grossman; 1988a, 1988b; Hall et al., 1999; Phillips et al.; 2011; Homm and Breitung, 2012; Phillips et al., 2014). By contrast, only a few attempts have so far been made to estimate parametric specifications of stock-price bubbles from time-series data (see Wu, 1995,1997; Al-Anaswah and Wilfling, 2011, Lammerding et al., 2013). Accordingly, there seems to be a need to establish a rigorous econometric framework for estimating structural forms of rational stock-price bubbles.

A variety of parametric bubble specifications have been proposed in the literature, the most influential being the rational, periodically collapsing bubble model introduced by Evans (1991). Although this stochastic, nonlinear bubble process captures many theoretically justified properties of speculative bubbles, two salient features of Evans' specification appear to be incompatible with empirically well-documented real-word stock-price data. (a) By definition, the Evans bubble necessarily collapses entirely within one period. (b) Whenever it collapses, the Evans bubble falls back to the same non-zero mean value. Allen and Gale (2000) and Kindleberger and Aliber (2005), among others, provide justification for why, after a crash, stock-price adjustments to the fundamental level typically follow a longer-lasting process. Moreover, they

point out that, in the case of recurring bubbles, the respective stock-price adjustment processes are likely to take different periods of time. Following this line of argument, we suggest modelling the burst of a bubble and the subsequent stock-price adjustment as a stochastically deflating process.

Figure 1 about here

Supporting empirical evidence is presented in Figure 1, which displays the NASDAQ stock-market index between January 1990 and October 2013. The two shaded areas represent the deflation processes, starting with the bursting dot-com bubble in March 2000 and the crash in the aftermath of the subprime mortgage crisis in October 2007, respectively. Typical of most deflation processes is that they are caused by bad news to investors operating in a highly uncertain financial-market environment, in which panic trading reactions are likely to accelerate the adjustment process. However, such extreme downturns in asset prices are occasionally interrupted by policy and/or regulatory interventions, which provide room for short-term price recovery.

Overall, this paper pursues two objectives. First, we establish a new nonlinear bubble specification, which combines the periodically collapsing Evans bubble with the incompletely bursting bubble model proposed by Fukuta (1998). Our specification is consistent with rational expectations and capable of generating recurring bubble trajectories with stochastically deflating adjustment processes. Second, we implement an econometric framework based on sequential Monte Carlo methods for estimating our new parametric bubble specification. For this purpose, we set up a nonlinear state-space model that includes stock-price fundamentals and our bubble specification in non-logarithmic terms. In contrast to the previous literature, we explicitly abstain from log-linearizing the basic (present-value) stock-price model. This approach enables us to directly estimate the parameters from the structural economic model and provides

an unambiguous interpretation of (a) the resulting parameter estimates, and (b) the filtered bubble process.

Econometrically speaking, we estimate our nonlinear latent bubble process via the so-called particle filter, a methodology originally from the engineering literature (see Gordon et al., 1993, for a comprehensive presentation). Technical review articles include Doucet et al. (2001) and Creal (2012). Applications to economics and finance include Kim et al. (1998), Fernández-Villaverde and Rubio-Ramirez (2007), Kim and Stoffer (2008), Duan and Fulop (2009) and Pitt et al. (2014). As described below, a prerequisite to applying the particle filter is that all model parameters are known. Since, in practice, appropriate parameter values need to be extracted from the data, we employ the particle-based approach of the Expectation Maximization (EM) algorithm established by Schön et al. (2011) for estimating our nonlinear state-space model. In the empirical part of the paper, we apply our econometric methodology first to artificial data, and then to four real-world stock-price indices, namely the German stock index DAX, the NASDAQ, the S&P 500 and the Hang Seng index. We demonstrate how to exploit the parameter estimates, in order to analyze important quantitative features of the respective stock-price bubble dynamics (such as bubble growth and deflation rates).

The remainder of the paper is organized as follows. Section 2 briefly reviews the basic present-value stock-price model, presents the relevant bubble specifications from the literature and introduces our new stochastically deflating bubble model. Section 3 transforms the entire stock-price model into a nonlinear state-space representation and establishes our estimation framework. Section 4 applies the estimation methodology to artificial and real-world data sets consisting of prices and dividends for the four major stock-price indices. Section 5 concludes.

2 Present-value model and rational bubbles

2.1 Present-value model

In the linear present-value model with rational expectations, the price of a stock at date t, P_t, is given by the Euler equation

$$P_t = \frac{1}{1+r}\left[E_t(P_{t+1}) + E_t(D_{t+1})\right], \tag{1}$$

where D_{t+1} is the stock-dividend payment between t and $t+1$. $E_t(\cdot)$ denotes the conditional expectation operator based on all information available to market participants at time t. r is the *required rate of return* that is just sufficient to compensate investors for the inherent riskiness of the stock (Campbell et al., 1997; Cuthbertson and Nitzsche, 2004). Substituting future prices forward repeatedly, we obtain the following solution of the expectational difference equation (1):

$$P_t = \sum_{i=1}^{\infty}\left(\frac{1}{1+r}\right)^i \cdot E_t(D_{t+i}) + \lim_{n\to\infty}\left(\frac{1}{1+r}\right)^n \cdot E_t(P_{t+n}). \tag{2}$$

In many applications, speculative bubbles are ruled out by assuming validity of the transversality condition

$$\lim_{n\to\infty}\left(\frac{1}{1+r}\right)^n \cdot E_t(P_{t+n}) = 0, \tag{3}$$

yielding the unique fundamental stock price:

$$P_t = P_t^f = \sum_{i=1}^{\infty}\left(\frac{1}{1+r}\right)^i \cdot E_t(D_{t+i}). \tag{4}$$

The basic idea behind a rational bubble is that there are mathematical expressions B_t that (a) are consistent with the limit-term appearing on the right-hand side of Eq. (2), and (b) are chosen such that the stock-price process in Eq. (2) satisfies the

Euler Eq. (1):

$$P_t = P_t^f + B_t = \sum_{i=1}^{\infty} \left(\frac{1}{1+r} \right)^i \cdot E_t(D_{t+i}) + B_t. \tag{5}$$

In view of Eq. (5), we interpret the rational bubble B_t as the deviation of the current stock price P_t from its current fundamental value P_t^f. The entire class of solutions to the Euler Eq. (1) is given by Eq. (5), in which B_t is any random variable satisfying the (discounted) martingale property

$$E_t(B_{t+1}) = (1+r) \cdot B_t \qquad \text{or, equivalently,} \qquad B_t = \frac{1}{1+r} \cdot E_t(B_{t+1}). \tag{6}$$

Any B_t, satisfying the martingale property (6), is called a rational bubble, because its presence in Eq. (5) is consistent with rational expectations.

2.2 Previous rational bubble models

Several alternative rational bubble specifications satisfying the martingale property from Eq. (6) have been suggested in the literature. The structurally simplest rational bubble model may be defined as

$$B_t = (1+r) \cdot B_{t-1} + \omega_t, \tag{7}$$

where ω_t denotes an i.i.d. error term with mean zero (Wu 1995, 1997). However, this specification implies a continuous (stochastic) growth of the bubble leading to infinitely high stock prices at some future date, a pattern inconsistent with real-world data.

The next specification, proposed by Blanchard (1979) and Blanchard and Watson (1982), constitutes a rational bubble with two different states. Given the probability $0 < \pi \leq 1$, it is defined as

$$B_t = \begin{cases} \frac{1}{\pi\psi} B_{t-1} + \varpi_t & , \text{ with probability } \pi \\ \varpi_t & , \text{ with probability } 1-\pi \end{cases}, \tag{8}$$

where $\psi \equiv (1+r)^{-1}$ and ϖ_t is an i.i.d. error term with mean zero. With probability π,

this bubble grows at a faster rate than the required rate of return r (whenever $\pi < 1$). With probability $1 - \pi$, the bubble collapses onto the realization of the stochastic error term ϖ_t. Note that the periodically collapsing bubble (8) suffers from two major drawbacks. (a) In view of Eq. (6), Diba and Grossman (1988b) argue that, in general, rational bubbles cannot start from zero and negative bubbles are ruled out as $t \to \infty$. Owing to the probabilistic nature of the error term ϖ_t, the specification in Eq. (8) does not necessarily preclude either of these two stipulations. (b) In the case of a burst, the bubble (8) collapses entirely within one period, thus precluding an empirically plausibe slow or moderate deflation process.

The most frequently used rational bubble specification, which overcomes Diba and Grossman's (1988b) fundamental critique, is the periodically collapsing bubble from Evans (1991). It is defined as

$$B_t = \begin{cases} \frac{1}{\psi} B_{t-1} u_t & , \text{ if } B_{t-1} \leq \tau \\ [\kappa + \frac{1}{\pi\psi}\Big(B_{t-1} - \kappa\psi\Big)\nu_t] u_t & , \text{ if } B_{t-1} > \tau \end{cases} , \tag{9}$$

where κ and τ are real constants to be chosen such that $0 < \kappa < (1 + r)\tau$ and $\{u_t\}_{t=1}^{\infty}$ is an exogenous process of i.i.d. random variables with $u_t > 0$ and $E_{t-1}(u_t) = 1$ for all t. The variables $\{u_t\}$ are assumed to be lognormally (LN) distributed and scaled to have unit mean, i.e. we assume $u_t = \exp(y_t - \iota^2/2)$ with $\{y_t\}_{t=1}^{\infty}$ being i.i.d. $N(0, \iota^2)$.[1] $\{\nu_t\}_{t=1}^{\infty}$ constitutes an exogenous i.i.d. Bernoulli process independent of $\{u_t\}_{t=1}^{\infty}$ with $\Pr(\nu_t = 1) = \pi$ and $\Pr(\nu_t = 0) = 1 - \pi$ for $0 < \pi \leq 1$. The event $\{\nu_t = 1\}$ signifies that the bubble will continue to grow, whereas the bubble bursts in the case of $\{\nu_t = 0\}$.

Obviously, the Evans bubble (9) features two different rates of growth. For $B_{t-1} \leq \tau$, the bubble grows at the mean rate $\frac{1}{\psi} - 1 = r$. For $B_{t-1} > \tau$, the bubble grows at the faster rate $\frac{1}{\pi\psi} - 1 > r$ (whenever $\pi < 1$), but collapses with probability $1 - \pi$ per period. When the bubble bursts, it falls back to the mean value κ and the process recommences.

[1] In other words, $\{u_t\}$ represents an i.i.d. distributed lognormal process with $u_t \sim \text{LN}(\frac{-\iota^2}{2}, \iota^2)$.

As mentioned above, the Evans bubble exhibits two empirically unrealistic features. (a) Like the Blanchard bubble (8), the Evans bubble collapses entirely within one period, and (b) the Evans bubble necessarily drops back to the same non-zero mean level κ after bursting.

An alternative theoretical model which establishes a more flexible deflating behavior is the so-called incompletely bursting bubble proposed by Fukuta (1998), the formal specification of which consists of three potential states:

$$B_t = \begin{cases} \frac{1}{\psi}\frac{\alpha_1}{\pi_1}B_{t-1} & , \text{ with probability } \pi_1 \\ \frac{1}{\psi}\frac{\alpha_2}{\pi_2}B_{t-1} & , \text{ with probability } \pi_2 \\ \frac{1}{\psi}\frac{1-\alpha_1-\alpha_2}{1-\pi_1-\pi_2}B_{t-1} & , \text{ with probability } 1-\pi_1-\pi_2 \end{cases}, \tag{10}$$

where it is assumed that $0 < \alpha_1 < 1, 0 < \alpha_2 < 1, 0 < 1 - \alpha_1 - \alpha_2 < 1$, and additionally for the state probabilities, $(1 - \alpha_1 - \alpha_2)/(1 - \pi_1 - \pi_2) < \alpha_2/\pi_2 < \alpha_1/\pi_1$. It is straightforward to verify that in States 1 and 2, we have $B_t > B_{t-1}$, whereas in State 3, the parameter restrictions imply $B_t < B_{t-1}$. It is obvious to characterize State 1 as the "large bubble state", State 2 as the "small bubble state" and State 3 as the "incompletely bursting state". Despite its ability to capture a more realistic deflation process, the major empirical drawback of Fukuta's specification is that, within each state, the bubble is subject to deterministic growth.

2.3 A new rational bubble model

Combining specific features of the Evans bubble (9) with the incomplete-bursting property of the Fukuta model (10), we now introduce a new rational bubble specification, which is strictly positive, periodically explosive (recurring) and stochastically deflating. Using the same notation as for the Evans and the Fukuta bubbles, our specification

consists of two distinct states:

$$B_t = \begin{cases} \frac{\alpha}{\psi\pi} B_{t-1} u_t & , \text{ with probability } \pi \\ \frac{1-\alpha}{\psi(1-\pi)} B_{t-1} u_t & , \text{ with probability } 1-\pi \end{cases} . \tag{11}$$

Via the Bernoulli process $\{\nu_t\}$, this mixture of distributions can be written more compactly in one single equation as

$$B_t = \left[\left(\left[\frac{\alpha}{\psi\pi} - \frac{1-\alpha}{\psi(1-\pi)} \right] \nu_t + \frac{1-\alpha}{\psi(1-\pi)} \right) B_{t-1} \right] u_t, \tag{12}$$

where we assume that $0 < \alpha < 1$. This latter constraint ensures that the bubble never collapses to zero and can thus rebuild.

Additionally, we stipulate that $\frac{\alpha}{\pi} > 1$ and $\frac{1-\alpha}{1-\pi} < \psi$, thus ensuring the following neat interpretation of the two bubble states. In State 1, which occurs with probability π, the bubble grows with the mean factor $\frac{\alpha}{\psi\pi}$ implying the mean growth rate $\frac{\alpha}{\psi\pi} - 1 = \frac{\alpha}{\pi} - 1 + \frac{\alpha}{\pi} \cdot r > r$ (i.e. a faster growth rate than the required rate of return). State 2, occurring with probability $1-\pi$, models the deflation of the bubble with deflation factor $\frac{1-\alpha}{\psi(1-\pi)} < 1$, or equivalently, with negative (imploding) bubble growth rate $\frac{1-\alpha}{\psi(1-\pi)} - 1 < 0$. Hinging on the concrete parameter constellation, the quantitative extent of a specific bubble deflation ranges between a "small/moderate correction" and a "big crash" within one or arbitrarily many periods. In contrast to Fukuta's specification (10), our model allows for a stochastic and thus more realistic bubble growth/deflation within each state. In addition to its empirically more plausible trajectories, compared to the Evans and Fukuta models, our specification (12) is also more parsimonious.

It remains to be shown that our bubble specification is rational, by verifying the martingale property (6). Using (a) the stochastic independence of the processes $\{u_t\}$ and $\{\nu_t\}$, (b) the conditional unit-mean assumption for all u_t, and (c) the Bernoulli

distribution for all ν_t, we obtain

$$\begin{aligned}
E_t(B_{t+1}) &= E_t\left\{\left[\left(\left[\frac{\alpha}{\psi\pi}-\frac{1-\alpha}{\psi(1-\pi)}\right]\nu_{t+1}+\frac{1-\alpha}{\psi(1-\pi)}\right)B_t\right]u_{t+1}\right\} \\
&= E_t\left\{\left(\left[\frac{\alpha}{\psi\pi}-\frac{1-\alpha}{\psi(1-\pi)}\right]\nu_{t+1}+\frac{1-\alpha}{\psi(1-\pi)}\right)B_t\right\}\cdot E_t(u_{t+1}) \\
&= E_t\left\{\left(\left[\frac{\alpha}{\psi\pi}-\frac{1-\alpha}{\psi(1-\pi)}\right]\nu_{t+1}+\frac{1-\alpha}{\psi(1-\pi)}\right)B_t\right\} \\
&= \pi\cdot\left(\frac{\alpha}{\psi\pi}-\frac{1-\alpha}{\psi(1-\pi)}+\frac{1-\alpha}{\psi(1-\pi)}\right)B_t+(1-\pi)\cdot\frac{1-\alpha}{\psi(1-\pi)}\cdot B_t \\
&= \frac{1}{\psi}B_t.
\end{aligned}$$

Figure 2 about here

Figure 2 displays four simulated trajectories of our stochastically deflating bubble process (12), where we set $\psi = 0.9804$ in each simulation, but choose distinct combinations of the parameters ι^2, α and π. We initiate all bubble processes with the starting value $B_0 = 0.5$. The trajectories consist of 250 observations representing a time span of approximately 21 years, based on monthly data. All trajectories exhibit two or three major bubbles, all differing from each other (a) in their respective stochastically explosive growth rates during the emerging phases, and (b) in their stochastic deflation rates during the adjustment process.

3 Specification issues and nonlinear state-space estimation

In this section, we first express our present-value stock-price model in nonlinear state-space form and then establish the estimation framework using sequential Monte Carlo methods.

3.1 Nonlinear state-space representation

In order to set up our state-space model, we reconsider Eq. (5), in which the stock price P_t equals the present value of expected future dividends, i.e. the fundamental value P_t^f, plus the bubble term B_t. In order to obtain an estimable solution to the stock-price P_t, we need an assumption on the dynamics of future dividend payments. Many suggestions on how to model dividend dynamics are available in the literature, for example, in the classical Gordon (1959) growth model, where dividends grow at a constant rate. Alternatively, several consumption-based asset-pricing models (*inter alia* Lucas, 1978; Barro, 2006) assume that dividends follow a simple random walk or a random walk with drift. A common feature of all these dividend specifications is that the resulting fundamental stock-price solution to Eq. (5), P_t^f, can be written as a constant ϕ multiplied by the current level of dividend payments, D_t.[2] In line with these theoretical models, we specify the fundamental stock-price at date t as

$$P_t^f = \phi \cdot D_t + \varepsilon_t,$$

where ε_t constitutes a Gaussian white-noise error term (with mean zero and variance σ_ε^2) capturing the impact of all other factors on the fundamental stock price.

In our present-value stock-price equation,

$$P_t = P_t^f + B_t = \phi \cdot D_t + \varepsilon_t + B_t, \tag{13}$$

we specify B_t as our stochastically deflating rational bubble model from Eq. (12). Thus, Eqs. (13) and (12) constitute our complete econometric framework, which we view as a nonlinear two-equation state-space model. In particular, since stock prices (P_t) and dividends (D_t) are observable variables, whereas the bubble component (B_t) is latent,

[2] In the Gordon growth model, the constant ϕ is affected by the (real) rate of return and the dividend growth rate, whereas in the asset-pricing models, ϕ includes the first two moments of the dividend process, the time preference rate and the degree of relative risk aversion.

Eq. (13) represents our observation equation, and Eq. (12) our state equation.

At this stage, it is worth mentioning an advantage of our nonlinear state-space formulation over the log-linear approximation to the standard present-value model, as proposed by Campbell and Shiller (1988). Our direct modeling (avoiding logs) preserves the clear-cut economic relationship between the stock price P_t, its fundamental value P_t^f, and the bubble component B_t, which follows directly from the original Euler Eq. (1). By contrast, the log-linear approximation of the present-value model blurs this relationship and the bubble component B_t no longer constitutes the deviation of the stock price from its fundamental value, but rather a fundamental value-to-price ratio, which is less obvious to interpret.

3.2 The particle filter

In the following subsections, we introduce the so-called particle-filter and particle-smoother approaches, which we apply in our empirical analysis in Section 4. These sequential Monte Carlo methods enable us to estimate and filter the unobserved state variable (the bubble compenent) from our nonlinear state-space framework. Our exposition closely follows those in Schön et al. (2011) and Creal (2012).

Adopting the notation from the standard literature, we consider the general nonlinear state-space model consisting of the observation equation

$$y_t = m_t(x_t, \mu_t, \zeta_t), \tag{14}$$

and the state equation

$$x_t = h_t(x_{t-1}, \lambda_t, \eta_t), \tag{15}$$

in which y_t and x_t represent the observable and the state variable for $t = 1, \ldots, T$. In our empirical application in Section 4, y_t and x_t take on the roles of the stock price P_t and the bubble component B_t. The random variables μ_t and λ_t in Eqs. (14)

and (15) constitute observable input variables, where μ_t, in particular, assumes the role of the dividend payments D_t in Section 4. Since μ_t and λ_t are not essential to establishing our econometric methodology, we formally fade out both variables in the subsequent notation for ease of readability. Moreover, $\{\zeta_t\}_{t=1}^T$ and $\{\eta_t\}_{t=1}^T$ are assumed to be mutually independent i.i.d. processes with known probability density functions (pdfs), while m_t and h_t are deterministic nonlinear functions of given form. In line with the literature, we formally write the pdfs of y_t and x_t as $p(y_t|x_t;\boldsymbol{\theta})$ and $p(x_t|x_{t-1};\boldsymbol{\theta})$, respectively. We emphasize that in general, both pdfs differ in their functional forms, although we denote each by the unified outer symbol $p(\cdot|\cdot)$ for ease of notation. Both pdfs depend on the parameter vector $\boldsymbol{\theta}$, which has to be estimated from a series of observations $\mathbf{y}_{1:T} \equiv (y_1, \ldots, y_T)$. However, in our first step, we assume the parameter vector $\boldsymbol{\theta}$ to be known and defer the estimation of $\boldsymbol{\theta}$ to Section 3.4.

Since the states $\mathbf{x}_{0:T} \equiv (x_0, x_1, ..., x_T)$ are unobservable, we also estimate them (a) by using the observable data $\mathbf{y}_{1:T} = (y_1, y_2, ..., y_T)$, and (b) by exploiting the formal structure of the underlying state-space model. To this end, we consider the joint conditional pdf $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\theta})$ (that is, the joint probability distribution of the unobservable states given the observed data) and the associated (conditional) expectation vector of the states $\mathbb{E}(\mathbf{x}_{0:T}) \equiv \mathbb{E}(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\theta})$. Adopting the compact notation from Creal (2012), we may write

$$\mathbb{E}(\mathbf{x}_{0:T}) = \int \mathbf{x}_{0:T}\, p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\theta})\, \mathrm{d}\mathbf{x}_{0:T}. \tag{16}$$

If the pdf $p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\theta})$ were known, we could simulate the N trajectories $\mathbf{x}_{0:T}^{(1)}, \ldots, \mathbf{x}_{0:T}^{(N)}$ from this target distribution and a natural simulation-based estimate of the expected states would be given by $\widehat{\mathbb{E}(\mathbf{x}_{0:T})} = (1/N)\sum_{i=1}^N \mathbf{x}_{0:T}^{(i)}$.

However, since, in our application, the pdf of the target distribution is unknown, we cannot simulate from it and therefore have to resort to the concept of importance

sampling. The basic idea behind importance sampling is to choose an appropriate proposal distribution $g_{0:T}(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\varphi})$ (with $\boldsymbol{\varphi}$ denoting the parameter vector) from which it is possible to simulate. To illustrate, we can rewrite Eq. (16) as

$$\mathbb{E}(\mathbf{x}_{0:T}) = \int \mathbf{x}_{0:T} \frac{p(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\theta})}{g_{0:T}(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\varphi})} g_{0:T}(\mathbf{x}_{0:T}|\mathbf{y}_{1:T};\boldsymbol{\varphi})\, \mathrm{d}\mathbf{x}_{0:T}. \tag{17}$$

In view of Eq. (17), we can estimate the expectation vector by the following weighted average of the N trajectories $\mathbf{x}_{0:T}^{(1)}, \ldots, \mathbf{x}_{0:T}^{(N)}$ simulated from our proposal distribution:

$$\widehat{\mathbb{E}(\mathbf{x}_{0:T})} = \sum_{i=1}^{N} \mathbf{x}_{0:T}^{(i)} \frac{w^{(i)}}{\sum_{j=1}^{N} w^{(j)}}, \tag{18}$$

where the weights

$$w^{(i)} \propto \frac{p(\mathbf{x}_{0:T}^{(i)}|\mathbf{y}_{1:T};\boldsymbol{\theta})}{g_{0:T}(\mathbf{x}_{0:T}^{(i)}|\mathbf{y}_{1:T};\boldsymbol{\varphi})} \tag{19}$$

are appropriately chosen so as (a) to capture the relationship between the target and the proposal distribution, and (b) to correct for simulating from the "wrong" distribution.

In our empirical application, it turns out to be advantageous to modify the above-described procedure slightly, so that we can draw sequentially from the proposal distribution. More explicitly, we factor our proposal distribution into two conditional distributions by writing for any date t

$$g_{0:t}(\mathbf{x}_{0:t}|\mathbf{y}_{1:t};\boldsymbol{\varphi}) = g_t(x_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t};\boldsymbol{\varphi}) g_{0:t-1}(\mathbf{x}_{0:t-1}|\mathbf{y}_{1:t-1};\boldsymbol{\varphi}). \tag{20}$$

Thus, at any date t, we only draw a new set of N values $x_t^{(1)}, \ldots, x_t^{(N)}$ from the first pdf $g_t(x_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t};\boldsymbol{\varphi})$ of the factored proposal distribution in Eq. (20) and append these N values to the corresponding N trajectories realized up to date $t-1$, $\mathbf{x}_{0:t-1}^{(1)}, \ldots, \mathbf{x}_{0:t-1}^{(N)}$. Overall, the full trajectories at date t can thus be represented as $\mathbf{x}_{0:t}^{(i)} = \left(\mathbf{x}_{0:t-1}^{(i)}, x_t^{(i)}\right)$ for $i = 1, \ldots, N$.[3]

[3]Note that the function g_t may change over time.

Writing the joint conditional probability distribution $p(\mathbf{x}_{0:T}^{(i)}|\mathbf{y}_{1:T};\boldsymbol{\theta})$ in recursive form, we obtain the weight for path i at date t as

$$\begin{aligned} w_t^{(i)} &= \frac{p(y_t|x_t^{(i)};\boldsymbol{\theta})p(x_t^{(i)}|x_{t-1}^{(i)};\boldsymbol{\theta})p(\mathbf{x}_{0:t-1}^{(i)}|\mathbf{y}_{1:t-1};\boldsymbol{\theta})}{p(y_t|\mathbf{y}_{1:t-1};\boldsymbol{\theta})g_t(x_t^{(i)}|\mathbf{x}_{0:t-1}^{(i)},\mathbf{y}_{1:t};\boldsymbol{\varphi})g_{0:t-1}(\mathbf{x}_{0:t-1}^{(i)}|\mathbf{y}_{1:t-1};\boldsymbol{\varphi})} \\ &\propto w_{t-1}^{(i)}\frac{p(y_t|x_t^{(i)};\boldsymbol{\theta})p(x_t^{(i)}|x_{t-1}^{(i)};\boldsymbol{\theta})}{g_t(x_t^{(i)}|\mathbf{x}_{0:t-1}^{(i)},\mathbf{y}_{1:t};\boldsymbol{\varphi})} \\ &= w_{t-1}^{(i)}\hat{w}_t^{(i)}, \end{aligned} \tag{21}$$

where

$$\hat{w}_t^{(i)} \equiv \frac{p(y_t|x_t^{(i)};\boldsymbol{\theta})p(x_t^{(i)}|x_{t-1}^{(i)};\boldsymbol{\theta})}{g_t(x_t^{(i)}|\mathbf{x}_{0:t-1}^{(i)},\mathbf{y}_{1:t};\boldsymbol{\varphi})}. \tag{22}$$

The major advantage of this sequential procedure is that we only have to calculate the ratio from Eq. (22) at each date t, while we can use Eq. (21) for updating previous weights. Thus, instead of re-computing the entire expressions from Eqs. (18) and (19), we simply update our estimation when the new observation y_{t+1} becomes available. Overall, at date t, we obtain (a) N paths of the state variable each of length t, and (b) N corresponding weights, i.e. $\left(\mathbf{x}_{0:t-1}^{(i)}, x_t^{(i)}, w_t^{(i)}\right)$ for $i = 1,\ldots,N$. This approach is referred to as sequential importance sampling and the draws are called particles.

A frequently encountered problem in sequential importance sampling is that with time elapsing, the normalized weight of one particle converges to 1, while the remaining weights converge to 0. In this case, the estimate of the unobserved state consists of a single draw. In practice, this phenomenon (known as weight degeneracy) is circumvented by executing an additional resampling step at each date t. In our susequent application, we follow Gordon et al. (1993) and draw (at each date t) a new set of values $\left(\tilde{x}_t^{(1)},\ldots\tilde{x}_t^{(N)}\right)$ from the existing sample $\left(x_t^{(1)},\ldots,x_t^{(N)}\right)$ proportional to the normalized weights $w_t^{(i)}/\sum_{j=1}^N w_t^{(j)}, i = 1,\ldots,N$. After resampling, we set all N weights to $w_t^{(i)} = 1/N$. In general, the specific combination of sequential importance sampling

and resampling constitutes the basic particle-filtering approach.[4]

Overall, the above particle-filtering algorithm yields the following approximation to the filtered density $p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta})$:

$$p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta}) \approx \sum_{i=1}^{N} \tilde{w}_t^{(i)} \mathbf{1}(x_t = x_t^{(i)}), \tag{23}$$

where $\mathbf{1}(\cdot)$ denotes the indicator function (which is equal to 1 when its argument is true and 0 otherwise) and $\tilde{w}_t^{(i)} \equiv w_t^{(i)} / \sum_{j=1}^{N} w_t^{(j)}$. By analogy with Eq. (18), a simulation-based estimator of the expected unobserved state at date t,

$$\mathbb{E}(x_t) \equiv \mathbb{E}(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta}) = \int x_t p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta})\, \mathrm{d}x_t,$$

is given by

$$\widehat{\mathbb{E}(x_t)} = \sum_{i=1}^{N} x_t^{(i)} \tilde{w}_t^{(i)}, \tag{24}$$

which, after the resampling step, can be written as

$$\widehat{\mathbb{E}(x_t)} = \frac{1}{N} \sum_{i=1}^{N} \tilde{x}_t^{(i)}. \tag{25}$$

Finally, we need an appropriate proposal distribution $g_t(x_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t};\boldsymbol{\varphi})$ for implementing the particle filter. A straightforward candidate is the known density of the state equation, $p(x_t|x_{t-1};\boldsymbol{\theta})$, from which we can easily sample. The virtue of using this density is that the weights in Eq. (22) are then only defined through the known density of the observation equation, $p(y_t|x_t;\boldsymbol{\theta})$. This specific combination of proposal distribution, plus resampling strategy, constitutes the original particle filter as established by Gordon et al. (1993).

[4]Douc and Cappé (2005) provide an in-depth comparison of alternative resampling algorithms. Creal (2012) presents distinct particle-filtering approaches, which only differ in their choices of the proposal distribution $g_t(x_t|\mathbf{x}_{0:t-1}, \mathbf{y}_{1:t};\boldsymbol{\varphi})$ and the resampling algorithm used.

3.3 The particle smoother

Based on the full set of data collected up to date T, we now aim at improving all state estimates by approximating the so-called smoothed density $p(x_t|\mathbf{y}_{1:T};\boldsymbol{\theta})$ via an appropriate particle smoother. Our smoothing method closely follows Schön et al. (2011) and is equivalent to the reweighting particle smoother proposed by Hürzeler and Künsch (1998) and Doucet et al. (2000).

Applying the *Theorem of Total Probabilities*, we first express the requested smoothed density as

$$p(x_t|\mathbf{y}_{1:T};\boldsymbol{\theta}) = \int p(x_t|x_{t+1},\mathbf{y}_{1:t};\boldsymbol{\theta})p(x_{t+1}|\mathbf{y}_{1:T};\boldsymbol{\theta})\,\mathrm{d}x_{t+1}, \tag{26}$$

which, via *Bayes' Formula*, we rewrite as

$$p(x_t|\mathbf{y}_{1:T};\boldsymbol{\theta}) = p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta}) \int \frac{p(x_{t+1}|x_t;\boldsymbol{\theta})p(x_{t+1}|\mathbf{y}_{1:T};\boldsymbol{\theta})}{p(x_{t+1}|\mathbf{y}_{1:t};\boldsymbol{\theta})}\,\mathrm{d}x_{t+1}. \tag{27}$$

Invoking the *Theorem of Total Probabilities* once more, we represent the denominator of the integrand as

$$p(x_{t+1}|\mathbf{y}_{1:t};\boldsymbol{\theta}) = \int p(x_{t+1}|x_t;\boldsymbol{\theta})p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta})\,\mathrm{d}x_t. \tag{28}$$

Overall, we can write the smoothed density in terms of (a) the filtered density $p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta})$, (b) the density of the state equation $p(x_{t+1}|x_t;\boldsymbol{\theta})$, and (c) the smoothed density $p(x_{t+1}|\mathbf{y}_{1:T};\boldsymbol{\theta})$:

$$p(x_t|\mathbf{y}_{1:T};\boldsymbol{\theta}) = p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta}) \int \frac{p(x_{t+1}|x_t;\boldsymbol{\theta})p(x_{t+1}|\mathbf{y}_{1:T};\boldsymbol{\theta})}{\int p(x_{t+1}|x_t;\boldsymbol{\theta})p(x_t|\mathbf{y}_{1:t};\boldsymbol{\theta})\,\mathrm{d}x_t}\,\mathrm{d}x_{t+1}. \tag{29}$$

By analogy with Eq. (23), we approximate the smoothed density from Eq. (29) by

$$p(x_t|\mathbf{y}_{1:T};\boldsymbol{\theta}) \approx \sum_{i=1}^{N} \tilde{w}_{t|1:T}^{(i)} \mathbf{1}(x_t = x_t^{(i)}), \tag{30}$$

where

$$\tilde{w}_{t|1:T}^{(i)} = \tilde{w}_t^{(i)} \sum_{k=1}^{N} \frac{\tilde{w}_{t+1|1:T}^{(k)} p(x_{t+1}^{(k)}|x_t^{(i)};\boldsymbol{\theta})}{\sum_{j=1}^{N} \tilde{w}_t^{(j)} p(x_{t+1}^{(k)}|x_t^{(j)};\boldsymbol{\theta})}. \tag{31}$$

In order to implement this smoothing algorithm, we need the particles and the corresponding weights $\left(x_t^{(i)}, w_t^{(i)}\right)$ from the sequential importance sampling for every date t. For $t = T$, we have $\tilde{w}_{T|1:T}^{(i)} = \tilde{w}_T^{(i)}$ and the smoothed density $p(x_{t+1}|\mathbf{y}_{1:T};\boldsymbol{\theta})$ can be obtained recursively (see Schön et al., 2011). Based on the N particles $x_t^{(i)}$ from the sequential importance sampling and the N smoothed weights $\tilde{w}_{t|1:T}^{(i)}$, we estimate the (expected) smoothed states for any other date t (given all observations up to date T) by

$$\widehat{\mathbb{E}(x_{t|1:T})} = \sum_{i=1}^{N} x_t^{(i)} \tilde{w}_{t|1:T}^{(i)}. \tag{32}$$

3.4 Maximum likelihood estimation of parameters with the EM algorithm

We now address the estimation of the parameter vector $\boldsymbol{\theta}$ of our nonlinear state-space model. While maximum likelihood estimators of $\boldsymbol{\theta}$ are straightforward to implement in many state-space models, a non-trivial problem arises in our nonlinear framework. Owing to the resampling procedure, the likelihood function approximated by our particle-filtering approach becomes a discontinuous function of $\boldsymbol{\theta}$ (Hürzeler and Künsch, 2001; Creal, 2012) rendering invalid the use of gradient-based optimizers for maximizing the likelihood function (Pitt, 2002; Kantas et al., 2009). To overcome this discontinuity problem, we apply a variant of the *Expectation Maximization* (EM) algorithm established by Schön et al. (2011).

Consider the observabe data $\mathbf{y}_{1:T}$ and let $\ell(\boldsymbol{\theta}|\mathbf{y}_{1:T}) \equiv \log[p(\mathbf{y}_{1:T};\boldsymbol{\theta})]$ denote the log-likelihood function with respect to $\boldsymbol{\theta}$ given all observations. The basic idea behind the EM algorithm is that in each iteration k, the algorithm determines an estimate $\boldsymbol{\theta}_k$ of the unknown parameter vector $\boldsymbol{\theta}$, such that the observed-data log-likelihood value exceeds

the log-likelihood value from the preceding iteration, that is $\ell(\boldsymbol{\theta}_k|\mathbf{y}_{1:T}) > \ell(\boldsymbol{\theta}_{k-1}|\mathbf{y}_{1:T})$. We now additionally assume that the data set may be incomplete. Following Schön (2011), we denote the missing data by $\mathbf{x}_{1:T}$ (whose role is taken on by the state vector in our empirical application) and consider the complete-data log-likelihood function

$$\ell(\boldsymbol{\theta}|\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) \equiv \log[p(\mathbf{y}_{1:T}, \mathbf{x}_{1:T}; \boldsymbol{\theta})]. \tag{33}$$

In general, this complete-data log-likelihood function has a convenient shape and maximization becomes feasible (McLachlan and Krishnan, 2007).

The fact that the missing data $\mathbf{x}_{1:T}$ are unobservable requires that, at the beginning of each iteration k, the complete-data log-likelihood function from Eq. (33) has to be approximated by its expectation conditional on the observed values $\mathbf{y}_{1:T}$ and the given initial guess $\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1}$ for the true parameter vector. Then, the following objective function is defined:

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k) \equiv \int \ell(\boldsymbol{\theta}|\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) p(\mathbf{x}_{1:T}|\mathbf{y}_{1:T}; \boldsymbol{\theta}_k) \, \mathrm{d}\mathbf{x}_{1:T}. \tag{34}$$

The crucial step in iteration k is then to choose the parameter vector $\boldsymbol{\theta}$ so that $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k) > Q(\boldsymbol{\theta}_k, \boldsymbol{\theta}_k)$. This specific $\boldsymbol{\theta}$ then replaces the initial guess $\boldsymbol{\theta}_k = \boldsymbol{\theta}_{k-1}$ and iteration k is terminated. Overall, after starting with $k = 0$ and a starting vector $\boldsymbol{\theta}_0$, the EM algorithm becomes a two-step procedure in each iteration k. (a) In the expectation step, we determine the objective function $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ given the initial vector $\boldsymbol{\theta}_k$. (b) In the maximization step, we maximize $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ with respect to $\boldsymbol{\theta}$. We continue iterating in this two-step fashion until the difference between the Q-values of two successive iterations becomes smaller than some prespecified convergence criterion. Schön et al. (2011) and McLachlan and Krishnan (2007) show that each new iteration increases the value of the log-likelihood function, that is $\ell(\boldsymbol{\theta}|\mathbf{y}_{1:T}) > \ell(\boldsymbol{\theta}_k|\mathbf{y}_{1:T})$. This justifies considering the vector $\boldsymbol{\theta}$ from the final iteration as our desired maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$.

What remains is to obtain an operable form of the objective function $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ for our nonlinear state-space framework, as specified by Eqs. (14) and (15).[5] To this end, we first apply *Bayes' Formula* and exploit the Markov nature of the state equation (15), so as to write the complete-data log-likelihood function as

$$\begin{aligned} \ell(\boldsymbol{\theta}|\mathbf{y}_{1:T}, \mathbf{x}_{1:T}) &= \log[p(\mathbf{y}_{1:T}|\mathbf{x}_{1:T}; \boldsymbol{\theta})] + \log[p(\mathbf{x}_{1:T}; \boldsymbol{\theta})] \\ &= \log[p(x_1; \boldsymbol{\theta})] + \sum_{t=1}^{T-1} \log[p(x_{t+1}|x_t; \boldsymbol{\theta})] + \sum_{t=1}^{T} \log[p(y_t|x_t; \boldsymbol{\theta})]. \quad (35) \end{aligned}$$

Via the definition of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ in Eq. (34), we obtain the objective function as

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k) &= \int \log[p(x_1; \boldsymbol{\theta})] p(x_1|\mathbf{y}_{1:T}; \boldsymbol{\theta}_k) \, \mathrm{d}x_1 \\ &\quad + \sum_{t=1}^{T-1} \int \int \log[p(x_{t+1}|x_t; \boldsymbol{\theta})] p(x_{t+1}, x_t|\mathbf{y}_{1:T}; \boldsymbol{\theta}_k) \, \mathrm{d}x_t \, \mathrm{d}x_{t+1} \\ &\quad + \sum_{t=1}^{T} \int \log[p(y_t|x_t; \boldsymbol{\theta})] p(x_t|\mathbf{y}_{1:T}; \boldsymbol{\theta}_k) \, \mathrm{d}x_t. \quad (36) \end{aligned}$$

Ultimately, for determining $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$, we need the densities of the observation and state equation plus the densities $p(x_t|\mathbf{y}_{1:T}; \boldsymbol{\theta}_k)$ and $p(x_{t+1}, x_t|\mathbf{y}_{1:T}; \boldsymbol{\theta}_k)$, approximations of which we receive as by-products of the particle smoother, as described in Section 3.3. Being now equipped with these, we implement the EM algorithm as follows. In the expectation step of iteration k, we use the results from Eqs. (30) and (31) in conjunction with Eq. (23) and, following Schön et al. (2011), approximate the Q-function by

$$\begin{aligned} \widehat{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}_k) &= \sum_{i=1}^{N} \tilde{w}_{1|1:T}^{(i)} \log[p(x_1^{(i)}; \boldsymbol{\theta})] \\ &\quad + \sum_{t=1}^{T-1} \sum_{i=1}^{N} \sum_{j=1}^{N} \tilde{w}_{t|1:T}^{(ij)} \log[p(x_{t+1}^{(j)}|x_t^{(i)}; \boldsymbol{\theta})] \\ &\quad + \sum_{t=1}^{T} \sum_{i=1}^{N} \tilde{w}_{t|1:T}^{(i)} \log[p(y_t|x_t^{(i)}; \boldsymbol{\theta})], \quad (37) \end{aligned}$$

[5] For a detailed derivation of the following steps, see Schön et al. (2011).

where

$$\tilde{w}_{t|1:T}^{(ij)} = \frac{\tilde{w}_t^{(i)} \tilde{w}_{t+1|1:T}^{(j)} p(x_{t+1}^{(j)}|x_t^{(i)}; \boldsymbol{\theta}_k)}{\sum_{l=1}^{N} \tilde{w}_t^{(l)} p(x_{t+1}^{(j)}|x_t^{(l)}; \boldsymbol{\theta}_k)}, \tag{38}$$

and the $x_t^{(i)}, i = 1, \ldots, N$, are the particles from the sequential importance sampling. In the maximization step, we maximize the function $\widehat{Q}$ with respect to $\boldsymbol{\theta}$ to obtain the initial $\boldsymbol{\theta}_{k+1}$. We continue iterating until $\widehat{Q}(\boldsymbol{\theta}_{k+1}, \boldsymbol{\theta}_k) - \widehat{Q}(\boldsymbol{\theta}_k, \boldsymbol{\theta}_k)$ becomes smaller than some prespecified threshold value $c > 0$ and consider $\boldsymbol{\theta}_k$ from the terminal iteration as our maximum likelihood estimate $\widehat{\boldsymbol{\theta}}$. We note that the convergence properties of this particle-filtering-based EM algorithm correspond to those of general EM methods. In particular, Hu et al. (2008) and Schön et al. (2011) show that $\widehat{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ is an arbitrarily accurate approximation to $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$.

3.5 Standard errors

One drawback of the EM algorithm is that it does not automatically provide estimates of the covariance matrix of the ML estimators. However, various approaches to estimating the information matrix $\mathbb{I}(\boldsymbol{\theta})$ under the EM algorithm have been proposed in the literature, so that we can obtain asymptotic standard errors from the inverse of its estimate, $[\widehat{\mathbb{I}(\boldsymbol{\theta})}]^{-1}$. In our application below, we use the numerically stable estimator of the information matrix, as established by Duan and Fulop (2011).

As the starting point of presentation, we recall that the variance-covariance matrix of the score vector of the observed-data log-likelihood function, $\mathbf{s}(\boldsymbol{\theta}) \equiv \partial \ell(\boldsymbol{\theta}|\mathbf{y}_{1:T})/\partial \boldsymbol{\theta}$, equals the information matrix:

$$\mathbb{V}[\mathbf{s}(\boldsymbol{\theta})] = \mathbb{I}(\boldsymbol{\theta}). \tag{39}$$

Instead of using the scores of the observed-data log-likelihood function $\ell(\boldsymbol{\theta}|\mathbf{y}_{1:T})$, Duan and Fulop (2011) suggest approximating the covariance matrix $\mathbb{V}[\mathbf{s}(\boldsymbol{\theta})]$ by the smoothed scores of the complete-data log-likelihood function $\ell(\boldsymbol{\theta}|\mathbf{y}_{1:T}, \mathbf{x}_{1:T})$, as given in Eq. (35).[6]

[6]We leave the formal derivation aside and straightforwardly adapt the Duan and Fulop (2011)

Since the smoothed scores do not constitute martingale differences, we additionally apply the Newey-West (1987) estimator with lag truncation parameter l. Overall, we derive the Duan-Fulop (2011) asymptotic standard errors of our maximum likelihood estimates $\widehat{\boldsymbol{\theta}}$ from the matrix

$$[\widehat{\mathbb{I}(\boldsymbol{\theta})}]^{-1} = \left[\mathbf{A}_0 + \sum_{j=1}^{l}\left(1 - \frac{j}{l+1}\right)(\mathbf{A}_j + \mathbf{A}_j')\right]^{-1}, \tag{40}$$

where

$$\mathbf{A}_j = \sum_{t=1}^{T-j} \mathbf{a}_t(\widehat{\boldsymbol{\theta}})\mathbf{a}_{t+j}(\widehat{\boldsymbol{\theta}})',$$

with

$$\begin{aligned}\mathbf{a}_t(\widehat{\boldsymbol{\theta}}) &= \sum_{i=1}^{N}\sum_{j=1}^{N} \tilde{w}_{t|1:T}^{(ij)} \left[\frac{\partial \log\left[p(x_t^{(j)}|x_{t-1}^{(i)};\boldsymbol{\theta})\right]}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right] \\ &\quad + \sum_{i=1}^{N} \tilde{w}_{t|1:T}^{(i)} \left[\frac{\partial \log\left[p(y_t|x_t^{(i)};\boldsymbol{\theta})\right]}{\partial \boldsymbol{\theta}}\bigg|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}}\right].\end{aligned}$$

4 Empirical analysis

In this section, we estimate our stochastically deflating bubble specification from Section 2.3 via the particle-filtering approach for various artificial and real-world data sets. We first recall that our stock-price equation (13) and bubble specification (12) represent the observation equation (14) and the state equation (15) of the general nonlinear state-space framework, the explicit distributions of which are required for implementing the particle-filtering methodology. Since the error term ε_t in the observation equation (13) is assumed to be normally distributed, the observation equation is also Gaussian, with the pdf given by

$$p(P_t|B_t, D_t; \boldsymbol{\theta}) = p(y_t|x_t, \mu_t; \boldsymbol{\theta}) = \mathrm{N}(\phi \cdot D_t + B_t, \sigma_\varepsilon^2), \tag{41}$$

approximation procedure to our state-space framework. Technical details are available upon request.

where $\mathrm{N}(\cdot,\cdot)$ denotes the pdf of the normal distribution. The distribution of our state equation is completely characterized via the distribution of the stochastically deflating bubble model from Eqs. (11) and (12). Obviously, our bubble specification constitutes a mixture of two lognormal distributions with mixing weights π and $1-\pi$. Thus, the pdf of the state equation can be written as

$$\begin{aligned} p(B_t|B_{t-1};\boldsymbol{\theta}) &= p(x_t|x_{t-1},\lambda_t;\boldsymbol{\theta}) \\ &= \pi\cdot\mathrm{LN}\left(\frac{-\iota^2}{2}+\log\left[\frac{\alpha}{\pi\psi}B_{t-1}\right],\iota^2\right) \\ &\quad +(1-\pi)\cdot\mathrm{LN}\left(\frac{-\iota^2}{2}+\log\left[\frac{1-\alpha}{(1-\pi)\psi}B_{t-1}\right],\iota^2\right), \end{aligned} \tag{42}$$

where $\mathrm{LN}(\cdot,\cdot)$ denotes the pdf of the lognormal distribution. We collect all model parameters in Eqs. (41) and (42) in the vector $\boldsymbol{\theta}=(\phi,\sigma_\varepsilon^2,\psi,\iota^2,\pi,\alpha)'$.

Figure 3 about here

4.1 Artificial data

In order to assess the overall reliability of the estimation procedure, we first apply the approach to artificial data. For this purpose, we chose the parameter vector $\boldsymbol{\theta}=(\phi,\sigma_\varepsilon^2,\psi,\iota^2,\pi,\alpha)'=(50,1.5,0.9804,0.02,0.87,0.91)'$ and—assuming that dividends follow a simple random walk—simulated the stock-price and the stochastically deflating bubble series via Eqs. (13) and (12), respectively. Figure 3 displays both trajectories, each consisting of 250 observations.

Table 1 about here

Owing to the computational burden inherent in the EM algorithm, we only used $N=300$ particles in the estimation procedure. As the convergence criterion for termi-

nating the EM algorithm, we chose the threshold value $c = 1/N \approx 0.0033$.[7] We initialized the EM algorithm with the starting vector $\boldsymbol{\theta}_0 = (30, 0.5, 0.85, 0.01, 0.7, 0.75)'$ and numerically maximized the $\widehat{Q}$-function from Eqs. (37) and (38) by using the FMINCON module in MATLAB. The EM algorithm converged after 285 iterations. Finally, we computed Duan-Fulop (2011) standard errors according to Eq. (40), for which we used $N = 500$ particles and $l = 15$ lags. The estimation results are shown in Table 1. Obviously, all parameters appear to be accurately estimated.

Figure 4 about here

As described in Sections 3.2 and 3.3, an estimation of the states (that is the unobservable bubble series) can be achieved by applying the particle filter and the particle smoother. Figure 4 displays the true bubble trajectory, compared to the bubble series estimated via (a) the particle-filter (upper panel), and (b) the particle-smoother (lower panel), where we used $N = 500$ particles in each approach. Obviously, both techniques yield estimates of the bubble series that are almost indistinguishable from the true bubble data.

Figure 5 about here

4.2 Real-world data

We now apply our estimation procedure to the following four major stock-price indices: the German stock index (DAX), the Standard and Poor's index (S&P 500), the National Association of Securities Dealers Automated Quotations index (NASDAQ composite) and the Hang Seng index (HSI). All time series cover monthly data between January

[7]The convergence of the EM algorithm generally hinges on the accuracy of the approximation to $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$. As shown in Schön et al. (2011), $\widehat{Q}(\boldsymbol{\theta}, \boldsymbol{\theta}_k)$ constitutes an arbitrarily accurate approximation as $N \to \infty$. Owing to the relatively low number of $N = 300$ particles, we chose c as dependent on N, thus accepting a less accurate approximation. However, Schön et al. (2011) demonstrate that even when using only a small number of particles, the procedure generates reasonably accurate estimates.

1981 and February 2014 (398 observations). We chose 1981 as the beginning of our sampling period for all four indices, because at that time, price movements appeared to be moderate and did not indicate the existence of any (substantial) bubbles. We obtained the dividend time series by multiplying the respective stock-price index with the corresponding dividend yield. We converted the nominal data to real data, using the respective consumer-price indices. Except for the S&P 500 data, which we compiled from Robert Shiller's website,[8] all other data are from *Datastream*. To achieve numerical stability of the EM algorithm, we divided the DAX, NASDAQ and HSI time series by 10 and the S&P 500 time series by 20. The data are displayed in Figure 5.

Table 2 about here

4.2.1 Estimation results

Table 2 displays the estimation results for the four stock-price indices. We ran the parameter estimation procedure with $N = 300$ particles and used the threshold value $c = 1/300$ as our convergence criterion. For the DAX and the HSI, the EM algorithm converged after 256 and 219 iterations, respectively. For the NASDAQ and S&P 500, the parameter values stopped changing substantially after about 250 iterations. However, the EM algorithm did not converge according to our (conservative) threshold value, so that we terminated the EM algorithm after 500 iterations for these two indices. As in the artificial-data case, standard errors were computed according to Eq. (40) with $N = 500$ particles and $l = 15$ lags.

All parameter estimates in Block (1) of Table 2 exhibit economically plausible values. It is notable that we estimated all parameters freely, without imposing any inequality restrictions. Block (3) of Table 2 reveals that the (free) parameter estimates of all indices satisfy the important theoretical bubble constraints $\frac{\alpha}{\pi} > 1$ and $\frac{1-\alpha}{1-\pi} < \psi$

[8] See http://www.econ.yale.edu/∼shiller/data.htm.

discussed in Section 2.3. This ensures that in State 1, the index-specific stock-price bubbles grow at a faster mean rate than the corresponding required rates of return (cf. rows "r (in %)" in Block (2) and "Growth rate (in %) " in Block (4)), while the bubbles actually deflate in State 2 at the negative rates displayed in row "Deflation rate (in %)" in Block (5).

The estimated probabilities of being in the deflationary State 2 (row "$1 - \pi$" in Block (2)) are higher for the DAX (0.0516) and the NASDAQ (0.0405) than for the S&P 500 (0.0074) and HSI (0.0135). On the other hand, Block (5) of Table 2 indicates that the deflation rates of the S&P 500 (-43.7394 %) and the HSI (-88.4126 %) appear to substantially exceed those of the DAX (-22.1834 %) and the NASDAQ (-18.4483 %). We conclude that the DAX and NASDAQ bubbles are likely to run through deflation periods more frequently than the S&P 500 and HSI bubbles. However, once the indices have entered the deflationary state, the S&P 500 and HSI bubbles deflate at significantly faster rates than their DAX and NASDAQ counterparts, thus reverting far more rapidly to fundamentally justified levels.

Figure 6 about here

Based on the parameter estimates, we now disentangle the latent bubble process from stock-prices and dividends via the filtering methods described in Sections 3.2 and 3.3. Figure 6 displays the four estimated bubble components, as extracted via the particle smoother with $N = 500$ particles, along with the corresponding fundamental processes (computed as the difference between the stock-price index and the estimated bubble values).

All bubble trajectories exhibit clear-cut deflating periods during the new-economy crisis between 2000 and 2003 and the subprime-mortgage crisis in 2007/2008. Except for the NASDAQ, we also find moderate deflation phases in the aftermath of the Black

Monday (October 1987). Although all bubble trajectories undergo several bursts, they rarely deflate completely within only a few periods and typically begin to rebuild from different base values (as opposed to the theoretical properties of the Evans bubble). In contrast to the DAX, NASDAQ and S&P 500 bubbles, the HSI bubble exhibits several pronounced peaks, with relatively short-lived deflation periods (typically lasting only a few months).

Figure 7 about here

Figure 7 displays the bubble-index ratios, which we interpret as the speculation share in the four analyzed stock markets. In January 1981, about 40% of the DAX value could be ascribed to the bubble component, whereas for the NASDAQ and S&P 500 the speculative share started near zero, but grew to about 70 % and 44% in the run-up to Black Monday in October 1987. For these three indices, the respective ratios peaked at the beginning of the new-economy crisis, in particular for the DAX in February 2000 (84 %), for the NASDAQ in September 2000 (94 %) and for the S&P 500 in August 2000 (74 %). For the HSI, the ratio remained comparably moderate, ranging between 2 % and 58% during the entire sampling period.

4.3 Model selection and specification diagnostics

Prior to evaluating the goodness-of-fit of our stochastically deflating bubble model (12), some comments of the model comparison between the Evans bubble (9) and our specification are in order. We also attempted to fit the Evans bubble to our real-world time series, using the particle-filter framework. However, in so doing, we ran directly into an identification problem that prevented us from obtaining explicit parameter estimates for the Evans bubble. To analyze this problem, we fitted the Evans specification (9) to a plethora of artificial data sets and found that the EM

algorithm appears to be incapable of globally identifying the threshold parameter τ, thus rendering it infeasible to fit the Evans bubble to our real-world stock-price indices.[9]

By contrast, the estimation of our stochastically deflating bubble specification (12) does not raise any technical problems. All that remains is to investigate its goodness-of-fit, for which we invoke the test suggested by Diebold et al. (1998). This test is based on the Rosenblatt (1952) probability integral transform and assesses the model fit by evaluating a sequence of one-step-ahead density forecasts. To outline the procedure, let $\{y_t\}_{t=1}^T$ be a series of realizations from the sequence of densities $\{f_t(y_t|\Omega_{t-1})\}_{t=1}^T$ conditioned on an information set Ω_{t-1}. If the model $\{p_t(y_t|\Omega_{t-1})\}_{t=1}^T$ is correctly specified (i.e. if $\{p_t(y_t|\Omega_{t-1})\}_{t=1}^T = \{f_t(y_t|\Omega_{t-1})\}_{t=1}^T$), then the sequence of probability integral transforms (PITs)

$$z_t = \int_{-\infty}^{y_t} p_t(u|\Omega_{t-1})\,\mathrm{d}u \tag{43}$$

should be i.i.d. uniformly distributed over the interval $[0,1]$ (that $\{z_t\}$ should be i.i.d. $U(0,1)$).

Although originally established for out-of sample evaluations, we use this test in-sample to assess specification adequacy. The advantage of this testing procedure is that we do not compare our model with any reference model, but test directly how well our model fits our given data. Any deviation from the i.i.d. $U(0,1)$ distribution may indicate incorrect distributional assumptions about the underlying data-generating process and/or insufficiently captured conditional dynamics (see Tay and Wallis, 2000, p. 250). To implement the goodness-of-fit test, we need the one-step-ahead forecast densities. In our nonlinear state-space framework, the forecast densities $p(y_t|\mathbf{y}_{1:t-1};\boldsymbol{\theta})$ coincide with the contributions of the observed-data log-likelihood, as given by the prediction-error decomposition (Creal, 2012). Generally, an appropriate estimate of the forecast density at time t may be derived as a byproduct of the particle filter. In

[9]Detailed results of our simulation analysis are available upon request.

our case, in which we use a resampling step in each period of the particle filter, we approximate the forecast density by

$$p(y_t|\mathbf{y}_{1:t-1};\boldsymbol{\theta}) \approx \frac{1}{N}\sum_{i=1}^{N} \hat{w}_t^{(i)}, \tag{44}$$

with $\hat{w}_t^{(i)}$ defined as in Eq. (22), which coincides with the density of the observation equation given in Eq. (41).

Figure 8 about here

Several methods of evaluating the i.i.d. uniformity of the sequence of PITs $\{z_t\}$ have been suggested in the literature (Diebold et al., 1998; Tay and Wallis, 2000). As a graphical device, we plot the empirical cumulative distribution functions (cdfs) of our four PIT sequences and analyze their deviations from the diagonal (that is, from the theoretical cdf of the $U(0,1)$ distribution). Figure 8 displays the respective cdfs along with confidence intervals around the diagonal. We compute the interval bounds by using the critical value of the Kolmogorov-Smirnov goodness-of-fit test at the 5% level. Except for the HSI, all other PIT sequences appear to remain within their confidence bands, thus revealing no substantial deviation from the $U(0,1)$ distribution.

Table 3 about here

In a more formal analysis, we also employ a battery of well-known goodness-of-fit tests, namely the Kolmogorov-Smirnov, the Cramer-von Mises and Anderson-Darling tests. Table 3 contains the values of the respective test statistics along with the p-values, where, for each test, the null hypothesis states that the PIT sequence was randomly sampled from the $U(0,1)$ distribution. Evidently, the results appear to be robust against the distinct testing procedures. For the PIT sequences of the DAX and

the NASDAQ the null hypothesis of having been randomly sampled from the $U(0,1)$ distribution cannot be rejected at any conventional level, thus indicating a satisfactory model fit. By contrast, for the S&P 500 and the HSI, all tests reject the null hypothesis at the 5% and 1% levels, respectively, leaving room for specification improvement for these latter indices.

5 Conclusion

In this paper, we propose a new parametric specification of a rational bubble that is able to generate periodically recurring and stochastically deflating trajectories. Our theoretical bubble process is closely related to economic theory, empirically plausible and combines all important features of rational bubble specifications previously proposed in the literature. The second objective of the paper consists of establishing a rigorous econometric framework for estimating our nonlinear and unobservable bubble specification. To this end, we plug our parametric bubble process into the standard present-value stock-price model, which we then transform into a nonlinear state-space model. In order to estimate our entire state-space framework, we adapt a particle-filtering approach originally from the engineering literature.

We successfully apply the estimation methodology to artificial and real-world stock-price data and demonstrate the advantage of our new parametric bubble model over previous specifications with respect to estimation feasibility and data fit. It is worth mentioning that our framework avoids the typical log-linearization of the standard present-value model. As a result, all parameter estimates (and in particular, all estimates of our bubble parameters) can be interpreted unambiguously in direct correspondance with the underlying economic theory. This enables us to quantify several important features characterizing the dynamics of historical bubbles, like growth rates during the emerging phase and deflation rates during the downturn. In view of this, a

plethora of interesting applications appear to be conceivable in future empirical studies, such as comparisons of bubble dynamics in different stock-markets, and evaluations of the impact of monetary policy measures on bubble dynamics.

The diagnostic tests in Section 4.3 for the S&P 500 and the Hang Seng indices suggest that there is still room for bubble-specification improvement. In line with the suggestions of Diebold et al. (1998) and Tay and Wallis (2000), one straightforward attempt to improve on the conditional bubble dynamics consists of including additional bubble lags (e.g. B_{t-2}) on the right-hand side of Eq. (12). It is far from obvious, however, how this could be accomplished technically (a) without violating the martingale property (6), and (b) ensuring that the resulting bubble specification leads to a stock-price solution that satisfies the Euler equation (1). In view of this, a useful direction for future research may entail the development of alternative rational bubble specifications providing more flexible conditional dynamics.

References

Al-Anaswah, N. and B. Wilfling (2011). Identification of speculative bubbles using state-space models with Markov-switching. Journal of Banking and Finance 35 (5), 1073-1086.

Allen, F. and D. Gale (2000). Bubbles and crises. The Economic Journal 110 (460), 236-255.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. The Quarterly Journal of Economics 121 (3), 823-866.

Blanchard, O. J. (1979). Speculative bubbles, crashes and rational expectations. Economics Letters 3 (4), 387-389.

Blanchard, O. J. and M. W. Watson (1982). Bubbles, rational expectations and financial markets. In P. Wachtel (Ed.), Crises in the Economic and Financial Structure, pp. 295-315. Lexington Books, Lexington.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). The Econometrics of Financial Markets. Princeton University Press, Princeton.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. Review of Financial Studies 1 (3), 195-228.

Creal, D. (2012). A survey of sequential Monte Carlo methods for economics and finance. Econometric Reviews 31 (3), 245-296.

Cuthbertson, K. and D. Nitzsche (2004). Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange. Wiley, New York.

Diba, B. T. and H. I. Grossman (1988a). Explosive rational bubbles in stock prices? The American Economic Review 78 (3), 520-530.

Diba, B. T. and H. I. Grossman (1988b). The theory of rational bubbles in stock prices. The Economic Journal 98 (392), 746-754.

Diebold, F. X., T. A. Gunther, and A. S. Tay (1998). Evaluating density forecasts, with applications to financial risk management. International Economic Review 39 (4), 863-883.

Douc, R. and O. Cappé (2005). Comparison of resampling schemes for particle filtering. In 4th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 64-69.

Doucet, A., N. de Freitas, and N. J. Gordon (2001). Sequential Monte Carlo Methods in Practice. Springer Verlag, New York.

Doucet, A., S. Godsill, and C. Andrieu (2000). On sequential Monte Carlo sampling methods for Bayesian filtering. Statistics and Computing 10 (3), 197-208.

Duan, J.-C. and A. Fulop (2009). Estimating the structural credit risk model when equity prices are contaminated by trading noises. Journal of Econometrics 150 (2), 288-296.

Duan, J. C. and A. Fulop (2011). A stable estimator of the information matrix under EM for dependent data. Statistics and Computing 21 (1), 83-91.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. The American Economic Review 81 (4), 922-930.

Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2007). Estimating macroeconomic models: a likelihood approach. The Review of Economic Studies 74 (4), 1059-1087.

Fukuta, Y. (1998). A simple discrete-time approximation of continuous-time bubbles. Journal of Economic Dynamics and Control 22 (6), 937-954.

Gordon, M. J. (1959). Dividends, earnings, and stock prices. The Review of Economics and Statistics 41 (2), 99-105.

Gordon, N. J., D. J. Salmond, and A. F. M. Smith (1993). A novel approach to nonlinear and non-Gaussian Bayesian state estimation. IEE Proceedings. Part F: Radar and Signal Processing 140 (2), 107-113.

Hall, S. G., Z. Psaradakis, and M. Sola (1999). Detecting periodically collapsing bubbles: a Markov-switching unit root test. Journal of Applied Econometrics 14 (2), 143-154.

Homm, U. and J. Breitung (2012). Testing for speculative bubbles in stock markets: a comparison of alternative methods. Journal of Financial Econometrics 10 (1), 198-231.

Hu, X.-L., T. B. Schön, and L. Ljung (2008). A basic convergence result for particle filtering. IEEE Transactions on Signal Processing 56 (4), 1337-1348.

Hürzeler, M. and H. R. Künsch (1998). Monte Carlo approximations for general state-space models. Journal of Computational and Graphical Statistics 7 (2), 175-193.

Hürzeler, M. and H. R. Künsch (2001). Approximating and maximising the likelihood for a general state space model. In A. Doucet, N. de Freitas, and N. J. Gordon (Eds.), Sequential Monte Carlo Methods in Practice, pp. 159-175. Springer Verlag, New York.

Kantas, N., A. Doucet, S. S. Singh, and J. M. Maciejowski (2009). An overview of sequential Monte Carlo methods for parameter estimation in general statespace models. In 15th IFAC Symposium on System Identification, Volume 15, pp. 774-785.

Kim, J. and D. S. Stoffer (2008). Fitting stochastic volatility models in the presence of irregular sampling via particle methods and the EM algorithm. Journal of Time Series Analysis 29 (5), 811-833.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. The Review of Economic Studies 65 (3), 361-393.

Kindleberger, C. P., and R. Z. Aliber (2005). Manias, Panics and Crashes: A History of Financial Crises. John Wiley and Sons, Hoboken.

Lammerding, M., P. Stephan, M. Trede, and B. Wilfling (2013). Speculative bubbles in recent oil price dynamics: Evidence from a Bayesian Markov-switching state-space approach. Energy Economics 36, 491-502.

Lucas, R. E. (1978). Asset prices in an exchange economy. Econometrica 46 (6), 1429-1445.

McLachlan, G. and T. Krishnan (2007). The EM Algorithm and Extensions. John Wiley and Sons, Hoboken.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3), 703-708.

Phillips, P. C., S. Shi, and J. Yu (2014). Specification sensivity in right-tailed unit root testing fir explosive behaviour. Oxford Bulletin of Economics and Statistics 76 (3), 315-333.

Phillips, P. C., Y. Wu, and J. Yu (2011). Explosive behavior in the 1990s NASDAQ: when did exuberance escalate asset values? International Economic Review 52 (1), 201-226.

Pitt, M. K. (2002). Smooth particle filters for likelihood evaluation and maximisation. Unpublished Manuscript, Department of Economics, Warwick University.

Pitt, M. K., S. Malik, and A. Doucet (2014). Simulated likelihood inference for stochastic volatility models using continuous particle filtering. Annals of the Institute of Statistical Mathematics 66 (3), 527-552.

Rosenblatt, M. (1952). Remarks on a multivariate transformation. The Annals of Mathematical Statistics 23 (3), 470-472.

Schön, T. B., A. Wills, and B. Ninness (2011). System identification of nonlinear state-space models. Automatica 47 (1), 39-49.

Tay, A. and K. F. Wallis (2000). Density forecasting: a survey. Journal of Forecasting 19 (4), 235-254.

West, K.D. (1987). A specification test for speculative bubbles. Quarterly Journal of Economics 102, 553-580.

Wu, Y. (1995). Are there rational bubbles in foreign exchange markets? Evidence from an alternative test. Journal of International Money and Finance 14, 27-46.

Wu, Y. (1997). Rational bubbles in the stock market: accounting for the US stock-price volatility. Economic Inquiry 35 (2), 309-319.

Tables and Figures

Table 1

Particle-filter parameter estimates using the EM algorithm

Parameter	True value	Estimate	Standard error
ϕ	50.0000	49.8220	0.5112
σ_ε^2	1.5000	1.6839	0.2072
ψ	0.9804	0.9698	1.6239×10^{-4}
ι^2	0.0200	0.0231	6.2637×10^{-5}
π	0.8700	0.8917	0.0073
α	0.9100	0.9255	0.0044

Note: Standard errors were computed from the Duan and Fulop (2011) estimator given in Eq. (40) with $N = 500$ particles and $l = 15$ lags.

Table 2
Estimation results for the DAX, NASDAQ, S&P 500, and the HSI

Parameter	DAX	NASDAQ	S&P 500	HSI
(1) EM estimates:				
ϕ	13.3782	29.9531	20.2057	23.0731
	(0.0308)	(0.0324)	(0.0037)	(0.0101)
σ_ε^2	0.1277	0.4476	1.3530	3.1976
	(3.3094×10^{-4})	(0.0014)	(0.0033)	(0.0109)
ψ	0.9912	0.9840	0.9848	0.9589
	(2.0210×10^{-5})	(1.7480×10^{-5})	(9.8682×10^{-6})	(1.8464×10^{-4})
ι^2	0.0047	0.0061	0.0036	0.0414
	(2.2588×10^{-7})	(7.6905×10^{-7})	(6.7574×10^{-8})	(5.5004×10^{-5})
π	0.9484	0.9595	0.9926	0.9865
	(2.2129×10^{-4})	(0.0011)	(3.7926×10^{-5})	(6.2247×10^{-5})
α	0.9602	0.9675	0.9959	0.9985
	(1.4712×10^{-4})	(8.6242×10^{-4})	(1.3766×10^{-5})	(9.0434×10^{-7})
(2) Other estimates:				
r (in %)	0.8878	1.6260	1.5435	4.2862
$1-\pi$	0.0516	0.0405	0.0074	0.0135
(3) Bubble constraints:				
$\frac{\alpha}{\pi}$	1.0124	1.0083	1.0033	1.0122
$\frac{1-\alpha}{1-\pi}-\psi$	−0.2199	−0.1815	−0.4307	−0.8478
(4) State 1:				
$\frac{\alpha}{\psi\pi}$	1.0214	1.0247	1.0188	1.0555
Growth rate (in %)	2.1431	2.4733	1.8811	5.5547
(5) State 2:				
$\frac{1-\alpha}{\psi(1-\pi)}$	0.7782	0.8155	0.5626	0.1159
Deflation rate (in %)	−22.1834	−18.4483	−43.7394	−88.4126

Note: Standard errors are in parantheses and were computed from the Duan and Fulop (2011) estimator given in Eq. (40) with $N = 500$ particles and $l = 15$ lags.

Table 3
Goodness-of-fit tests for the DAX, NASDAQ, S&P 500, and the HSI

Test	DAX	NASDAQ	S&P 500	HSI
Kolmogorov-Smirnov	0.0394	0.0559	0.0706**	0.0994***
	(0.5586)	(0.1617)	(0.0361)	(0.0007)
Cramer-von Mises	0.1154	0.3128	0.5646**	0.8200***
	(0.5149)	(0.1244)	(0.0273)	(0.0065)
Anderson-Darling	0.8025	1.9191	3.3592**	5.4010***
	(0.4792)	(0.1018)	(0.0180)	(0.0019)

Note: p-values are in parantheses. $***$, $**$, $*$ denote statistical significance at the 1%, 5%, and 10% levels, respectively.

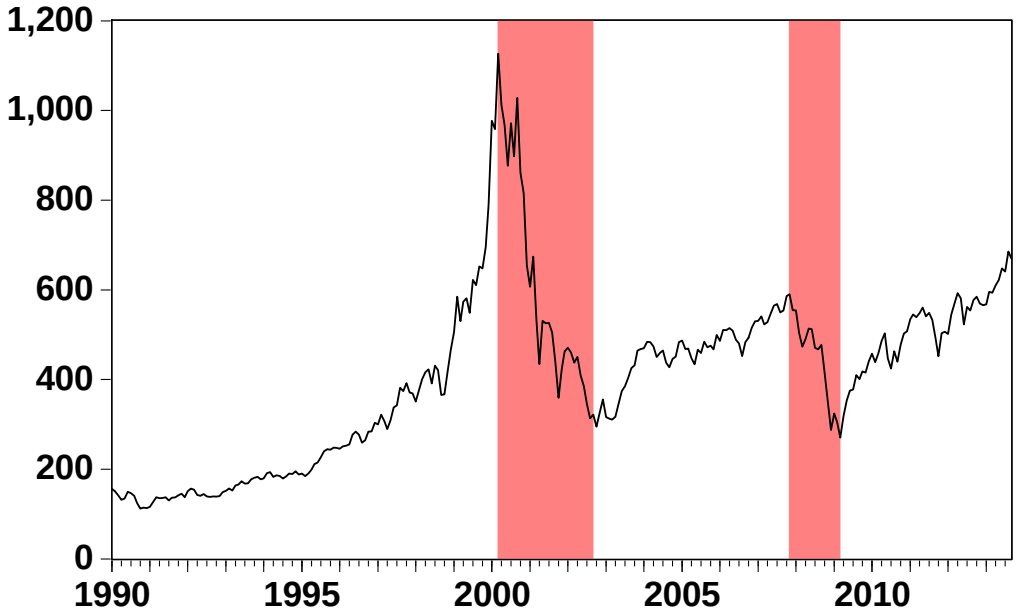

Figure 1. NASDAQ stock-market index, January 1990 – October 2013

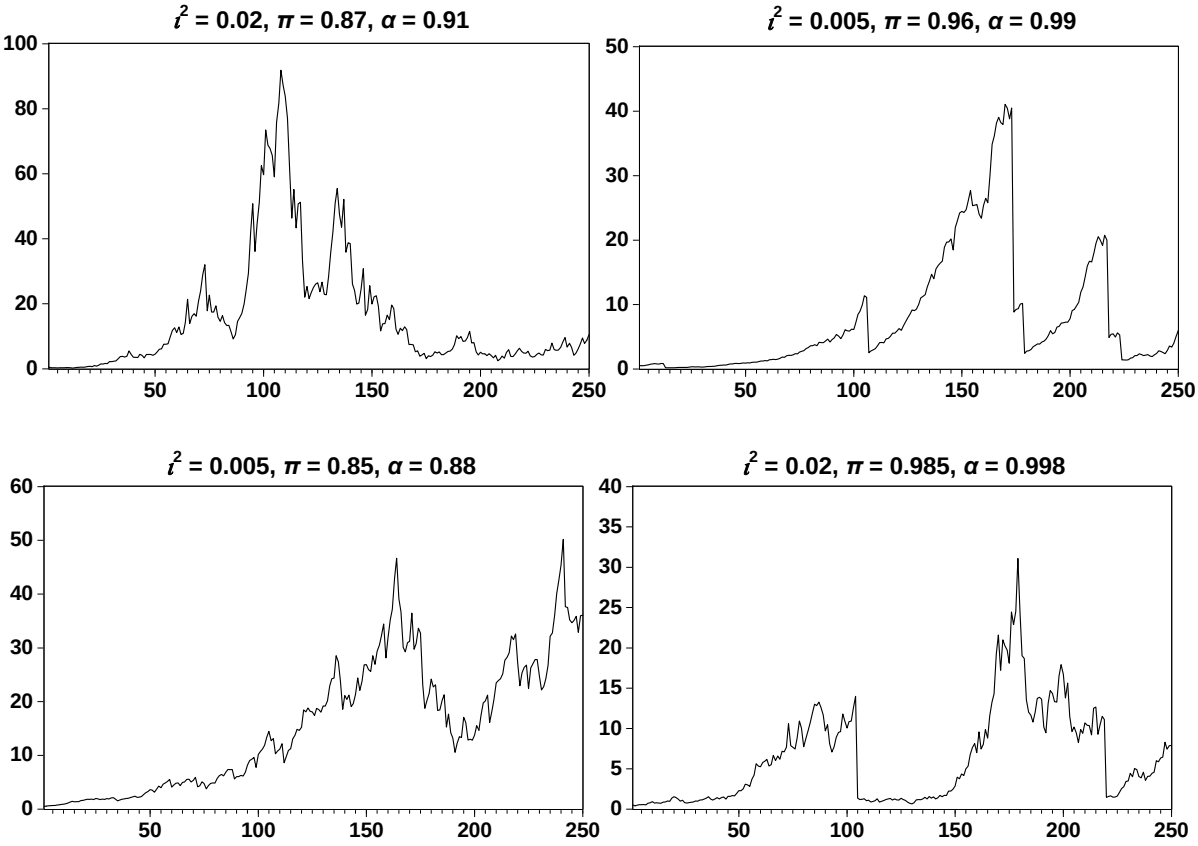

Figure 2. Bubble trajectories simulated according to Eq. (12)

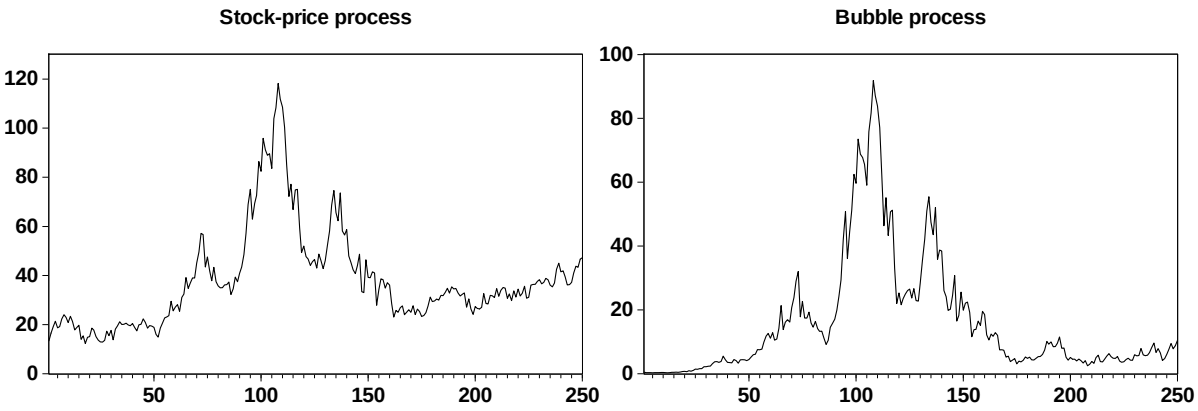

Figure 3. Simulated stock-price and included bubble process

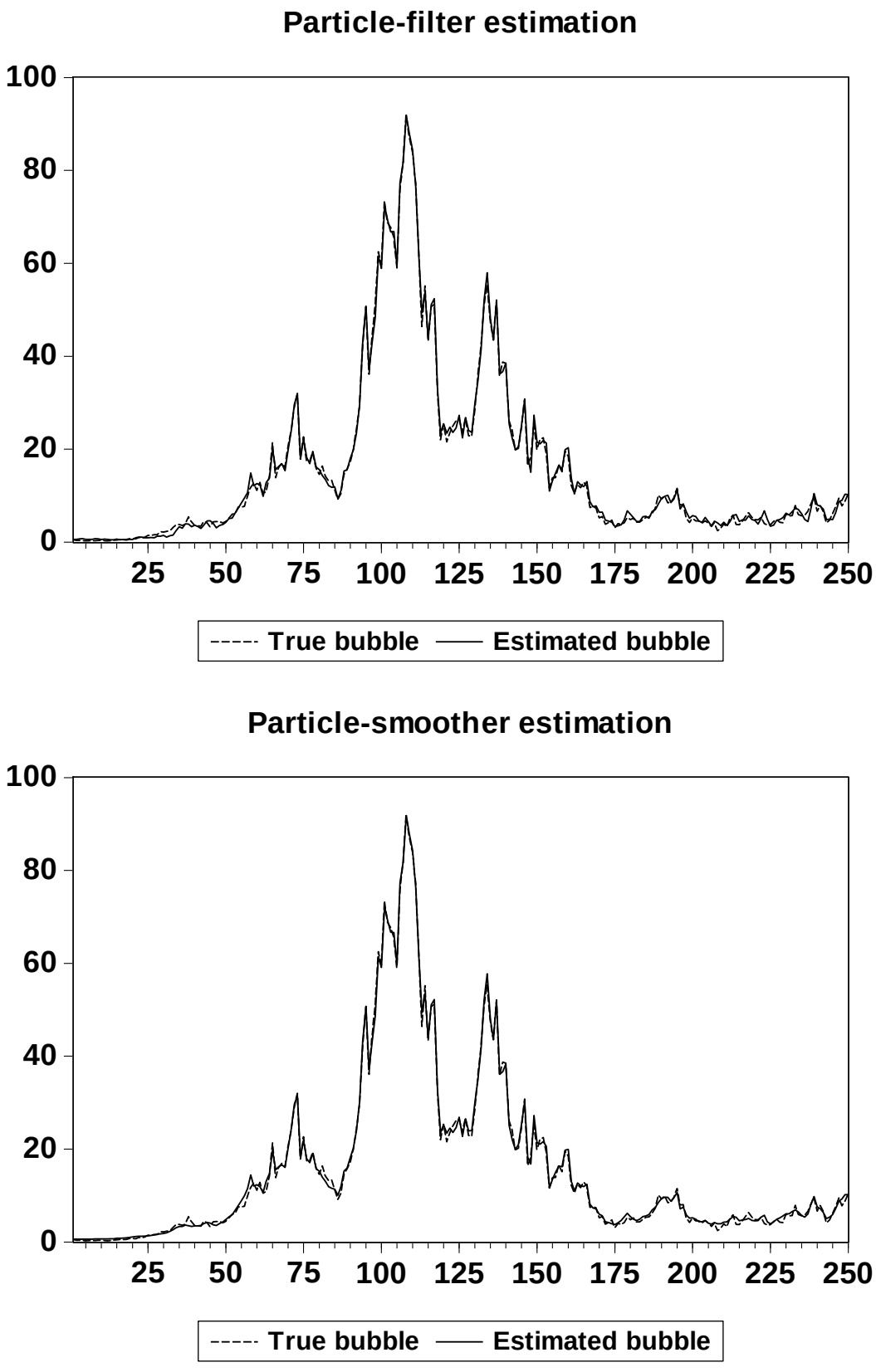

Figure 4. True versus estimated bubble processes

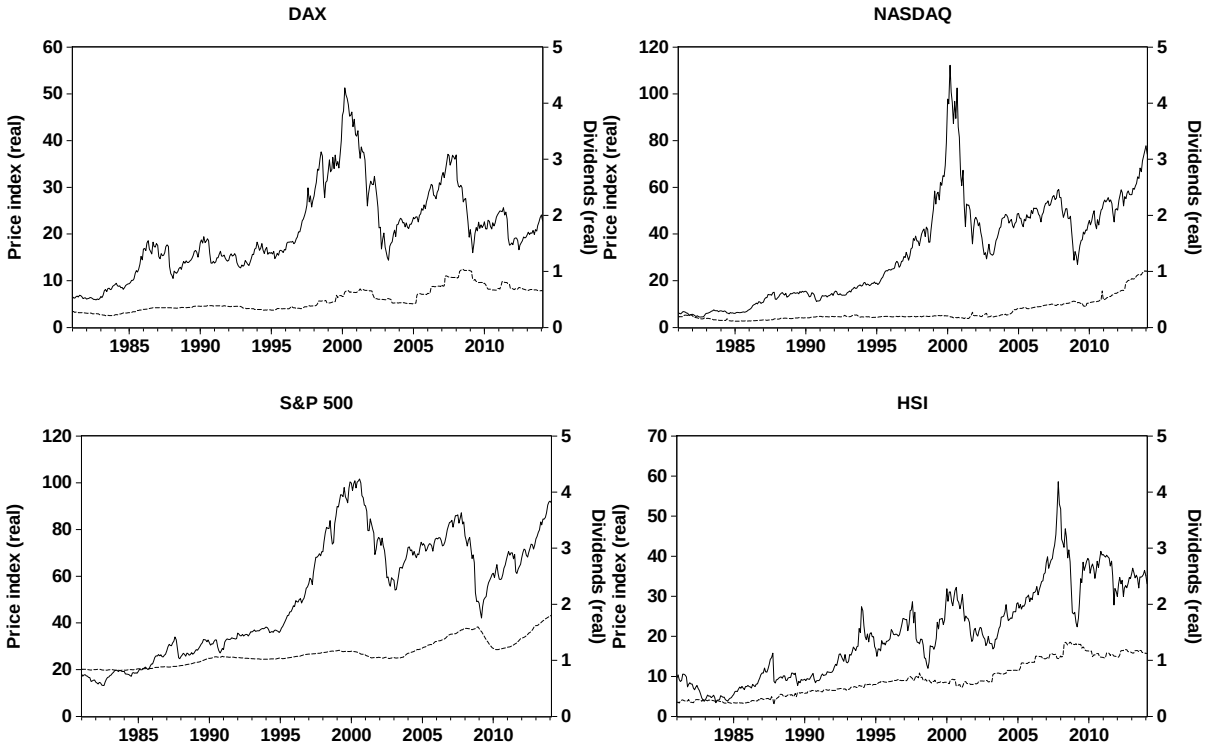

Figure 5. DAX, NASDAQ, S&P 500 and HSI (solid lines) and dividends (dashed lines)

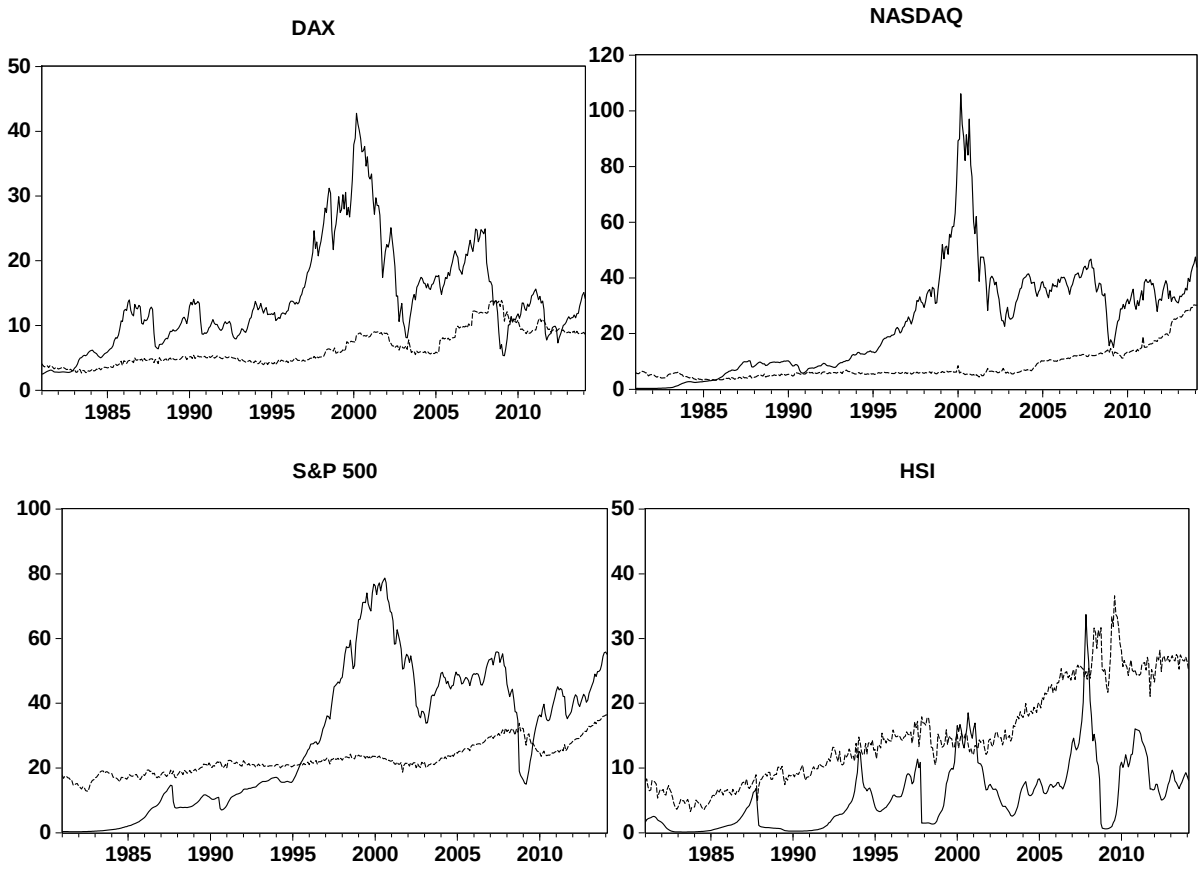

Figure 6. Estimated bubble processes (solid lines) and fundamental processes (dashed lines) of the DAX, NASDAQ, S&P 500 and HSI

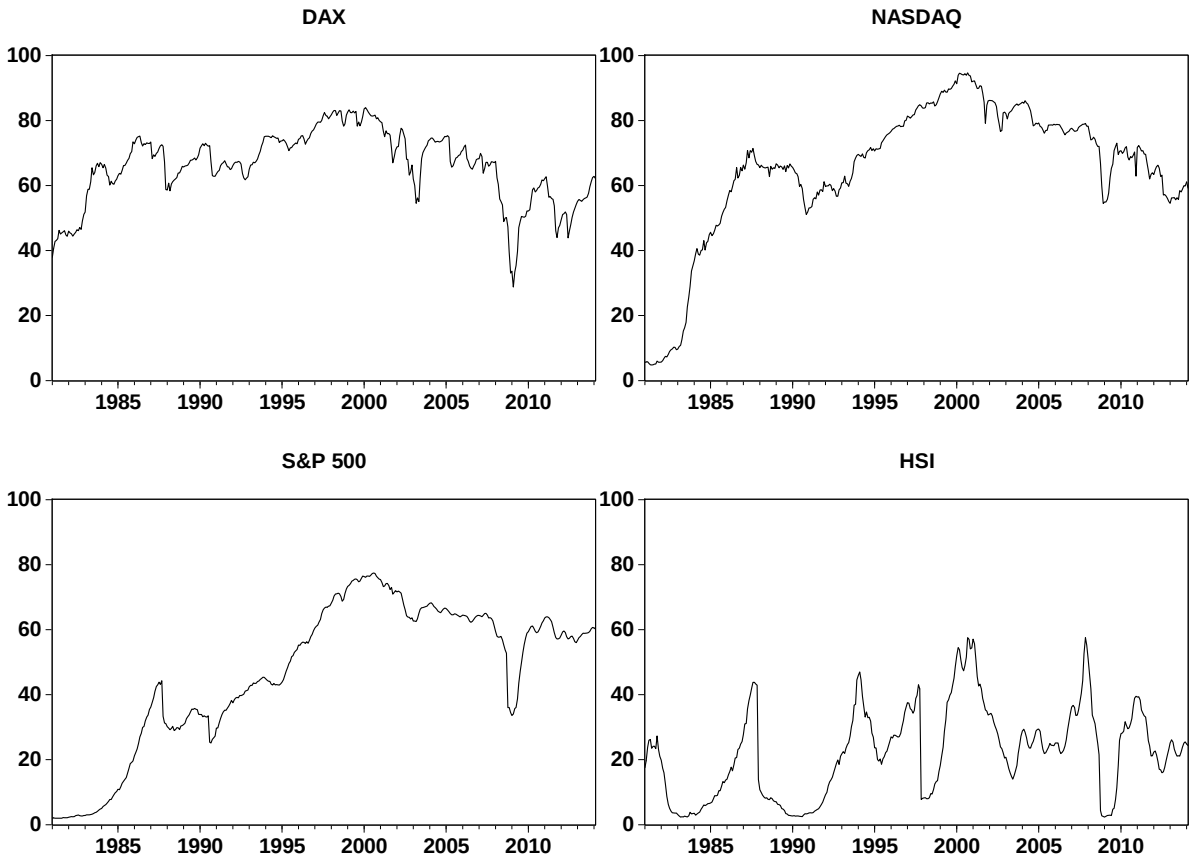

Figure 7. Bubble-stock-price ratios of the DAX, NASDAQ, S&P 500 and HSI

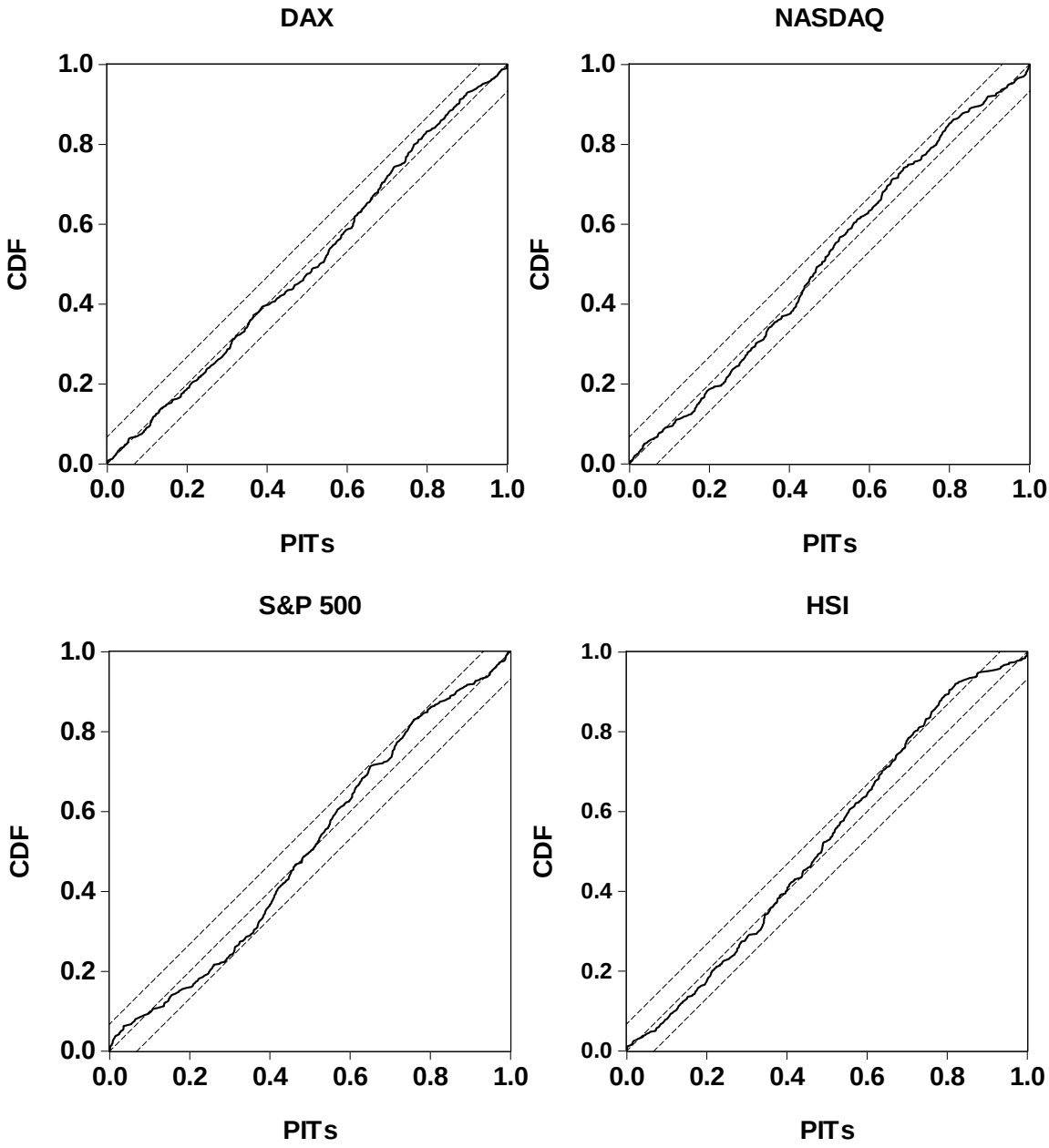

Figure 8. Sample cumulative distribution functions of the PITs for the DAX, NASDAQ, S&P 500, and HSI