Econometric Reviews
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597248

Volatility, Jumps, and Predictability of Returns: A Sequential Analysis

Davide Raggi[a]; Silvano Bordignon[b]
[a] Department of Economics, University of Bologna, Bologna, Italy [b] Department of Statistics, University of Padova, Padova, Italy

Online publication date: 09 June 2011

To cite this Article Raggi, Davide and Bordignon, Silvano(2011) 'Volatility, Jumps, and Predictability of Returns: A Sequential Analysis', Econometric Reviews, 30: 6, 669 — 695
To link to this Article: DOI: 10.1080/07474938.2011.553570
URL: http://dx.doi.org/10.1080/07474938.2011.553570

Econometric Reviews, 30(6):669–695, 2011

ISSN: 0747-4938 print/1532-4168 online
DOI: 10.1080/07474938.2011.553570

Taylor & Francis
Taylor & Francis Group

VOLATILITY, JUMPS, AND PREDICTABILITY OF RETURNS: A SEQUENTIAL ANALYSIS

Davide Raggi[1] and Silvano Bordignon[2]

[1]*Department of Economics, University of Bologna, Bologna, Italy*
[2]*Department of Statistics, University of Padova, Padova, Italy*

□ *In this article we propose a Monte Carlo algorithm for sequential parameter learning for a stochastic volatility model with leverage, nonconstant conditional mean and jumps. We are interested in estimating the time invariant parameters and the nonobservable dynamics involved in the model. Our simple but effective idea relies on the auxiliary particle filter algorithm mixed together with the Markov Chain Monte Carlo (MCMC) methodology. Adding an MCMC step to the auxiliary particle filter prevents numerical degeneracies in the sequential algorithm and allows sequential evaluation of the fixed parameters and the latent processes. Empirical evaluation on simulated and real data is presented to assess the performance of the algorithm. A numerical comparison with a full MCMC procedure is also provided. We also extend our methodology to superposition models in which volatility is obtained by a linear combination of independent processes.*

Keywords Auxiliary particle filters; Bayesian estimation; Leverage; MCMC; Return's predictability; Stochastic volatility with jumps.

JEL Classification C11; C15; C32; C58.

1. INTRODUCTION

In this article we propose a methodology to analyze the sequential parameter learning problem for a stochastic volatility model with jumps and a predictable conditional mean. We aim at updating the estimates of the parameters of interest together with the states continuously, following the flow of information arriving in the markets. There are various reasons, both from a practical and a theoretical point of view why we think sequential methods are appealing. Sequential procedures seem suitable when we are interested in real time applications where we need to update

Address correspondence to Davide Raggi, Department of Economics, University of Bologna, Piazza Scaravilli, no. 2, Bologna 40126, Italy; E-mail: davide.raggi@unibo.it

our estimates regularly. For example, economic agents need to produce estimates and forecasts in real time, meaning that they need to adapt their estimates every time a new observation is available. One of the most compelling advantages of sequential Monte Carlo methods is their reduced computational burden compared with other Monte Carlo procedures such as the Markov Chain Monte Carlo (MCMC), which requires us to restart the inferential procedure from scratch.

Our procedure builds on the particle filtering algorithm of Liu and West (2001) in which we include an MCMC step to prevent the algorithm from degenerating after a number of iterations. A direct application of the Liu and West (2001) algorithm to our model and long time series, in fact, evidences some occasional pitfalls, since a null posterior covariance matrix of the parameters is estimated. The use of MCMC together with particle filters has been proposed in Gilks and Berzuini (2001) and has proved to be an effective combination between the computational advantages of sequential algorithms and the statistical efficiency of MCMC methods. The introduction of the MCMC step in the Liu and West (2001) setup is particularly useful when dealing with long time series since it sensibly reduces the degeneration difficulties connected with sequential Monte Carlo methods. We also found that the sequential algorithm is more precise at detecting jump times and sizes than the standard MCMC approach.

We focus our analysis on stochastic volatility models (SV), a central topic in financial applications. A review of univariate models is shown by Ghysels et al. (1996), and for the multivariate context see Asai et al. (2006). Several variants of Autoregressive Conditional Heteroskedastic (ARCH) and SV models have been proposed to account for the empirical regularities of financial time series. In particular, in this article we deal with three such regularities within an SV framework. First, we consider the leverage effect between returns and conditional variances; second, we model the conditional mean, that is, the predictable component of the returns; finally, we take into account a jump's dynamics to describe extreme and rare events such as market crashes.

The introduction of a jump process in a volatility model has been proved to give an improved fit to data, both in relation to the model's ability to describe the return's behavior (Eraker et al., 2003), as well as for the pricing of financial derivatives (see Bakshi et al., 1997; Pan, 2002; and Eraker, 2004, amongst others). In the recent literature, there is in fact evidence in favor of jumps in returns and volatilities, since a diffusive behavior of these two processes seems to be inadequate to describe the underlying dynamics. Furthermore, if we consider the asset allocation problem in which the risky asset follows a jump diffusion process, there is some evidence that an extreme and rare event influences the conditional mean and the volatility, thus implying a modification on the optimal

portfolio weights (Liu et al., 2003). Leverage describes the relationship between returns and conditional variances. It is in fact reasonable to think that bad news in the markets (e.g., the price decreases), leads to an increase in the variance, which is a measure of the financial risks. On the other hand, episodes of high volatility induce expectations of lower future returns, hence, the negative correlation between these shocks. In the stochastic volatility literature leverage has been modelled by introducing a non-null correlation between returns and volatilities as suggested in Harvey and Shephard (1996), Yu (2005), and Omori et al. (2007), amongst others. This issue has also been extensively investigated in Asai and McAleer (2005), by introducing a dynamic asymmetric leverage model, in which volatility and returns are directly related to properly describe the impact of sign and magnitude of past returns on volatility. We, finally, deal with predictability of returns, an issue that has been considered since the early works of Merton (1971), that gave a theoretical justification for this feature. For optimal portfolio choices, it is important to take into account the conditional means and volatilities of returns, since economic theory shows that an investor gains from market predictability and volatility timing, even if the impact of these benefits is difficult to quantify.

The remainder of the article is organized as follows. The SV models considered are described in Section 2. Our inferential solution is outlined in Section 3. Empirical results based on simulated and real data are illustrated in Sections 4 and 5. Finally, Section 6 concludes.

2. STOCHASTIC VOLATILITY MODELS

The stochastic volatility model for the observable returns y_{t+1} is specified as

$$y_{t+1} = \mu_t + \exp\{v_t/2\}\epsilon_{t+1} + \kappa_{t+1}J_{t+1} \tag{1}$$

$$v_{t+1} = \alpha_v + \phi v_t + \sigma_\eta \eta_{t+1} \tag{2}$$

$$\mu_{t+1} = \alpha_\mu + \beta\mu_t + \sigma_\mu \zeta_{t+1}. \tag{3}$$

The nonobservable stationary processes μ_{t+1} and v_{t+1} are, respectively, the conditional mean and the log-volatility. We assume that the error vector $(\epsilon_{t+1}, \eta_{t+1}, \zeta_{t+1})$ is a standardized white Gaussian noise where the leverage effect has been modeled simply through $\mathrm{Cov}(\epsilon_{t+1}, \eta_{t+1}) = \rho$. In order to properly describe extreme events such as market crashes, a useful extension is to introduce a jump component both in the returns and in the volatilities. Duffie et al. (2000) for instance propose a model based on a stochastic differential equation with jumps driven by a marked point process. In the discrete time model, these discontinuities are governed by a sequence of independent Bernoulli random variables J_{t+1} with fixed

intensity[1] λ. A Gaussian random variable κ_{t+1} with mean μ_y and variance σ_y^2 describes the size or mark associated to each jump.

We directly model the conditional mean via an autoregressive process μ_{t+1}. Chernov et al. (2003) suggest that some serial dependence on μ_{t+1} can be motivated by the effect of non-synchronous trading and unexpected stochastic dividends. This dependence is assumed to be mean reverting. Similar dynamics for the conditional mean have been studied recently in Johannes et al. (2002). We also assume that v_0 and μ_0 are distributed according to their stationary distributions. Finally, the noise ζ_{t+1} is uncorrelated with ϵ_{t+1} and η_{t+1} even if there are no theoretical reasons to impose this constraint.

The prior distribution for the parameters vector $\boldsymbol{\theta}$ is consistent with Kim et al. (1998) and Eraker et al. (2003). We thus hypothesize the following prior distributions: $\alpha_v \sim \mathrm{N}(0;10)$, $\phi \sim \mathrm{Beta}(25;2)$, $\sigma_\eta^2 \sim \mathrm{IG}(2.5, 0.05)$, $\rho \sim \mathrm{U}(-1,1)$, $\alpha_\mu \sim \mathrm{N}(0;4)$, $\beta \sim \mathrm{Beta}(25;2)$, $\sigma_\mu^2 \sim \mathrm{IG}(2.5;0.05)$, $\lambda \sim \mathrm{Beta}(2;100)$, $\mu_y \sim \mathrm{N}(0;20)$, $\sigma_y^2 \sim \mathrm{IG}(2.5;0.05)$, where, in particular, *IG* denotes the inverse Gamma distribution.

In empirical applications, a first order autoregressive volatility specification can be restrictive and might not be able to describe long range dependencies. Different and more flexible long memory dynamics could be introduced within our framework. For instance, to help in describing a long memory behavior of the volatility in the ARCH literature, Engle and Lee (1999) use the combination of a highly persistent process together with a short memory dynamics. A similar parameterization for SV models to describe long memory has been studied in Chernov et al. (2003), Liesenfeld and Richard (2003), Durham (2006), Omori et al. (2007), and in Nakajima and Omori (2009). Volatility can thus be defined by a superposition of stationary autoregressive processes, that is, $v_t = v_{1,t+1} + v_{2,t+1}$ in which

$$v_{1,t+1} = \mu + \phi_1 v_{1,t} + \sigma_1 \eta_{1,t+1} \tag{4}$$

$$v_{2,t+1} = \phi_2 v_{2,t} + \sigma_2 \eta_{2,t+1}. \tag{5}$$

The first factor $v_{1,t}$ describes the long range dependencies whereas the second factor, $v_{2,t}$, is the short memory component that accommodates for extreme observations. We thus impose $-1 < \phi_2 < \phi_1 < 1$. As noted in Chernov et al. (2003) the low persistent volatility process $v_{2,t}$, as well as the jump factor J_t, describes the tail behavior of the returns. For this reason in this second specification we do not consider the jump dynamics that would be redundant. Moreover, the leverage effect is modeled through the

[1]This assumption has been relaxed in Eraker (2004) where the intensity is modeled by $\lambda_t = \lambda_0 + \lambda_1 v_t$.

non-null correlations $\text{Cov}(\eta_{1,t}, \epsilon_t) = \rho_1$ and $\text{Cov}(\eta_{2,t}, \epsilon_t) = \rho_2$, leading to the following distribution for the errors

$$\begin{pmatrix} \epsilon_t \\ \eta_{1,t} \\ \eta_{2,t} \end{pmatrix} \sim N(0, \Sigma), \quad \Sigma = 0 \begin{bmatrix} 1 & \sigma_1\rho_1 & \sigma_2\rho_2 \\ \sigma_1\rho_1 & \sigma_1^2 & 0 \\ \sigma_2\rho_2 & 0 & \sigma_2^2 \end{bmatrix}.$$

To make Σ positive definite we claim that $\rho_1^2 + \rho_2^2 < 1$. Our prior's choice is based on Omori et al. (2007) and states that $\mu \sim \text{N}(0, 1)$, $(\phi_1 + 1)/2 \sim \text{Beta}(20, 1.5)$, $(\phi_2 + 1)/2 \sim \text{Beta}(10, 10)$, $\sigma_1 \sim \text{IG}(2.5, 0.025)$, $\sigma_1 \sim \text{IG}(2.5, 0.05)$, and $\rho_i \sim \text{U}(-1, 1)$, $i = 1, 2$.

A different solution to introduce long run dependencies, not considered here but feasible in our inferential setup, has been proposed in Chan and Petris (2000), by describing v_t through a fractional integrated process. Following Chan and Palma (1998), an Autoregressive Fractional Integrated Moving Average (ARFIMA) process can be approximated through a Moving Average dynamics truncated at a finite lag M, leading to

$$v_t = \sum_{i=1}^{M} \psi_j \eta_{t-j}, \tag{6}$$

in which ψ_j is a given function of the long memory parameter. Then, following Chan and Petris (2000) it is possible to provide a state-space representation[2] of the stochastic volatility model that can be handled through our particle filtering method.

3. SEQUENTIAL PARAMETER LEARNING

Since their introduction, stochastic volatility models have provided an interesting benchmark for several estimation techniques, since their likelihood function is not available in closed form and then inference has to be based on approximations or numerical evaluations. Some of these techniques have been applied to estimate parameters of continuous time models with and without jumps,[3] others to inference on their discrete counterparts. Here we will focus on discrete time specifications. In the recent literature, Monte Carlo algorithms have provided a flexible yet powerful tool for inference on complex models possibly with non

[2]See also Raggi and Bordignon (2011) for a state space representation of a long memory ARFIMA model.

[3]Some of the most popular methods include Simulated Method of Moments (Chernov et al., 2003; Gallant and Tauchen, 1996), Approximate Maximum Likelihood (Aït-Sahalia, 2002), Generalized Method of Moments (Pan, 2002), Simulated Maximum Likelihood (Durham and Gallant, 2002), Markov Chain Monte Carlo (Elerian et al., 2001), Particle Filters (Golightly and Wilkinson, 2006; Johannes et al., 2009.)

observable factors. Several strategies based on importance sampling have been proposed for univariate models (Danielsson, 1994; Durbin and Koopman, 1997; Durham, 2006; Pitt, 2002; Shephard and Pitt, 1997). In particular Jungbacker and Koopman (2006) and Liesenfeld and Richard (2006) propose different sampling strategies which are important to estimate various versions of an SV model both in the univariate and in the multivariate case. The efficient importance sampling of Liesenfeld and Richard (2006) have been used to implement classical and Bayesian inference. A simulated maximum likelihood method for stochastic volatility models, jumps, and leverage has also been proposed in Malik and Pitt (2009). In a Bayesian framework, MCMC has been extensively used since the seminal papers of Jacquier et al. (1994) and Kim et al. (1998). Extensions to models with leverage have been considered, for instance, in Jacquier et al. (2004), Yu (2005), and Omori et al. (2007), whereas multivariate dynamics have been estimated in Chib et al. (2005), Bos and Shephard (2006), Chan et al. (2006), and Yu and Meyer (2006). Finally, applications to models with jumps have been developed through MCMC in Chib et al. (2002), Eraker et al. (2003), Raggi (2005), and in Nakajima and Omori (2009).

Many of the mentioned techniques provide efficient and accurate estimates when used for off-line applications, but seem to be inadequate when dealing with real time problems where we need to regularly update the estimates at each time step. In the recent literature, filtering techniques for sequential parameter learning have been implemented in Liu and West (2001), Storvik (2002), Johannes et al. (2006), and in Polson et al. (2008).

Our idea to implement sequential parameter learning relies on sequential Monte Carlo methods. We first describe the background tools for particle filtering with known and unknown parameters. We then focus on our solution for the parameter learning problem.

3.1. Some Background on Particle Filters

Particle filter algorithms, introduced in Gordon et al. (1993), have been successfully used in a variety of fields such as engineering, econometrics, and biology. They provide a suboptimal but feasible solution to the Bayesian filtering problem. A detailed review on Monte Carlo sequential algorithms is given in Doucet et al. (2001) and Arulampalam et al. (2002). Consider the general state-space model defined by the density $p(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1})$ that describes the evolution of the latent states $\boldsymbol{x}_t$ and by $p(y_t \mid \boldsymbol{x}_t)$ that specifies the observable y_t. Our goal is to estimate the distribution $p(\boldsymbol{x}_{t+1} \mid y_{1:t+1})$ given $p(\boldsymbol{x}_t \mid y_{1:t})$ in which $y_{1:t} = (y_1, \ldots, y_t)$ is the past history of the observable process up to time t. We also require the

knowledge of the initial distribution $p(\boldsymbol{x}_0)$, of the transition distribution $p(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}_t), t \geq 0$, and of the measurement distribution $p(y_{t+1} \mid \boldsymbol{x}_{t+1}), t \geq 1$.

The key idea is to approximate the filtering density $p(\boldsymbol{x}_{t+1} \mid y_{1:t+1})$ by a discrete cloud of points called particles $\boldsymbol{x}^j_{t+1}$, $j = 1, \ldots, N$, and a set of weights ω^j_{t+1} as follows

$$\hat{p}(\boldsymbol{x}_{t+1} \mid y_{1:t+1}) = \sum_{j=1}^{N} \omega^j_{t+1} \delta(\boldsymbol{x}_{t+1} - \boldsymbol{x}^j_{t+1}), \tag{7}$$

where $\delta(\cdot)$ is the Dirac delta measure. The cloud of points at time $t+1$ are chosen using the importance sampling principle, in which the importance density is $q(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}^i_t, y_t)$ and the weights are

$$\omega^i_{t+1} \propto \omega^i_t \frac{p(y_{t+1} \mid \boldsymbol{x}_{t+1}) p(\boldsymbol{x}^i_{t+1} \mid \boldsymbol{x}^i_t)}{q(\boldsymbol{x}^i_{t+1} \mid \boldsymbol{x}^i_t, y_t)} \quad i = 1, \ldots, N. \tag{8}$$

With this setup, it can be proved that the variance of the weights increases systematically over t, leading to a poor approximation of the filtering distribution. For this reason a resampling step is often added in order to avoid numerical degeneracies by getting rid of the particles with low probability. This is the Sampling Importance Resampling (SIR) algorithm. An important variant of the SIR filter is the Auxiliary Particle (AP) filter, suggested by Pitt and Shephard (1999) in which the proposal depends on the whole stream of particles, i.e., $q(\boldsymbol{x}_{t+1} \mid \boldsymbol{x}^i_{1:t}, y_t)$ through a suitable auxiliary variable.

When parameters are unknown, inference is a challenging question, in fact, standard particle methods such as the SIR filter and the AP filter can evidence some degeneracy problems, mainly in the presence of outliers and high dimensionality of the latent states. Recently, a number of article have tackled the problem of estimating the fixed parameters in a sequential context. For instance Storvik (2002) proposes a filter in which the parameters are sequentially updated by simulating from their conditional distribution $p(\boldsymbol{\theta} \mid y_{1:t+1})$ through MCMC. A different approach, named the practical filter by Polson et al. (2008), is based on the idea that $p(\boldsymbol{x}_{t+1}, \boldsymbol{\theta} \mid y_{1:t+1})$ can be expressed as a mixture of lag-filtering distributions. The estimate is then based on a rolling-window MCMC algorithm.

A different solution has been proposed in Gordon et al. (1993) and in Kitagawa (1998) by artificially defining an autoregressive dynamics for the parameters, say $\boldsymbol{\theta}_{t+1}$, and by incorporating them in an augmented state vector $(\boldsymbol{x}_{t+1}, \boldsymbol{\theta}_{t+1})$. The main point against this approach is that it leads to time varying and not to fixed parameter estimates. More precisely, Liu and West (2001) point out that these artificial dynamics can lead to posterior variances that are larger, relative to the true posteriors for the actual fixed

parameters. To correct for this overdispersion, West (1993) and Liu and West (2001) propose to approximate the posterior distribution $p(\boldsymbol{\theta} \mid y_{1:t+1})$ by a smooth kernel density, leading to

$$p(\boldsymbol{\theta} \mid y_{1:t+1}) \approx \sum_{i=1}^{N} \omega_t^i N(\boldsymbol{m}_{t+1}^i; h^2 \boldsymbol{\Sigma}_{t+1}). \tag{9}$$

The quantity $\boldsymbol{m}_{t+1}^i = a\boldsymbol{\theta}_{t+1}^i + (1-a)\bar{\boldsymbol{\theta}}_{t+1}$ is the kernel location for the ith component of the mixture whereas the matrix $\boldsymbol{\Sigma}_{t+1}$ and the vector $\bar{\boldsymbol{\theta}}_{t+1}$ are respectively estimates of the variance-covariance matrix and of the mean of the posterior distribution at time $t+1$. Furthermore, $\boldsymbol{\theta}_{t+1}^i$, $i = 1, \ldots, N$ is a sample from $p(\boldsymbol{\theta} \mid y_{1:t+1})$. The constants h and a, which measure the extent of the shrinkage and the degree of overdispersion of the mixture, are given by $h^2 = 1 - ((2\delta - 1)/2\delta)^2$ and $a = \sqrt{1 - h^2}$, whereas the discount factor δ ranges between 0.95-0.99. It can be proved that the variance of the mixture approximation in (9) is $\boldsymbol{\Sigma}_{t+1}$ and the mean is obviously $\bar{\boldsymbol{\theta}}_{t+1}$. According to this setup, at time $t+1$, a reasonable proposal for the posterior is then

$$\boldsymbol{\theta}_{t+1} \mid \boldsymbol{\theta}_t \sim \mathrm{N}(a\boldsymbol{\theta}_t + (1-a)\bar{\boldsymbol{\theta}}_t, h^2 \boldsymbol{\Sigma}_t). \tag{10}$$

This methodology has been successfully used in Liu and West (2001) in a dynamic factor stochastic volatility context and in Carvalho and Lopes (2007) in a switching regime stochastic volatility framework.

3.2. Particle Filters with MCMC Correction

For the stochastic volatility model with jumps defined in Eq. (1)–(3) we found that the Liu and West (2001) setup described above performs poorly. The major drawback with this algorithm is that the estimated posterior variance-covariance matrix $\boldsymbol{\Sigma}_{t+1}$ collapses to zero after a few hundred iterations. This problem is probably due to the sample impoverishment phenomenon caused by the resampling procedure, that in this context is likely due to the high dimension of the extended states vector and to the discontinuous nature of the jump process. In fact, particles with high probability are selected many times, causing a loss of diversity in the cloud of points. This effect is severe when the noise of the latent process is small.[4] To reduce degeneracies, following the suggestion of Gilks and Berzuini (2001), we propose a generalization of the Liu and West (2001) method by adding an MCMC move to increase sample variability of the particles. This should also help to reduce the correlation

[4]Note that this feature can be emphasized by the choice of $h^2 < 1$.

between particles after resampling. In practice, calling $\tilde{\boldsymbol{x}}_{t+1} = (\boldsymbol{x}_{t+1}, \boldsymbol{\theta})$, the particles $\tilde{\boldsymbol{x}}^i_{t+1}$ approximating $p(\boldsymbol{\theta}, \boldsymbol{x}_{t+1} \mid y_{1:t+1})$, can be moved to a different location $\tilde{\boldsymbol{x}}'^i_{t+1}$ according to a Markov transition kernel $T(\tilde{\boldsymbol{x}}_{t+1}, \tilde{\boldsymbol{x}}'_{t+1})$, that is invariant with respect to the same filtering distribution. For this reason, a burn-in period for the MCMC step is not necessary.

More formally, given the posterior distribution $p(\tilde{\boldsymbol{x}}_{t+1} \mid y_{1:t+1})$, the importance weights $\omega_{t+1}(\tilde{\boldsymbol{x}}_{t+1})$ and the proposal $q(\tilde{\boldsymbol{x}}_{t+1} \mid \tilde{\boldsymbol{x}}_{t+1}, y_t)$, it is easy to check that

$$\begin{aligned} p(\tilde{\boldsymbol{x}}_{t+1} \mid y_{1:t+1}) &= \int \omega_{t+1}(\tilde{\boldsymbol{x}}_{t+1}) q(\tilde{\boldsymbol{x}}_{t+1} \mid \tilde{\boldsymbol{x}}_t, y_t) T(\tilde{\boldsymbol{x}}_{t+1}, \tilde{\boldsymbol{x}}'_{t+1}) d\tilde{\boldsymbol{x}}_{t+1} \\ &= p(\tilde{\boldsymbol{x}}'_{t+1} \mid y_{1:t+1}). \end{aligned} \tag{11}$$

In other words, we move all the particles $(\boldsymbol{x}^i_{t+1}, \boldsymbol{\theta}^i)$, that approximate the posterior, through $T(\cdot, \cdot)$ thus obtaining a further approximation of the filtering distribution based on the weighted sample $(\boldsymbol{\theta}'^i, \boldsymbol{x}'^i_{t+1}, \omega^i_{t+1})$.

We now provide the details of the algorithm proposed considering the version we implement for the model described in Eq. (1)–(3). We write the vector of the states as $\boldsymbol{x}_{t+1} = (v_t, \mu_t, J_{t+1}, \kappa_{t+1})$ and we estimate the posterior distribution $p(v_t, \mu_t, J_{t+1}, \kappa_{t+1}, \boldsymbol{\theta} \mid y_{1:t+1})$.

In order to perform the MCMC step we need to keep track of the whole trajectory of each particle. A useful way to store all of this information is through a set of sufficient statistics S_t (see Fearnhead, 2002, for an extensive treatment on this point). In order to implement the sequential Monte Carlo procedure, it is thus necessary to update recursively the sufficient statistics set. We update S_t through a recursive updating rule $f(\cdot)$, such that $S_{t+1} = f(S_t, \boldsymbol{x}_{t+1})$. For our model with jumps, the sufficient statistics up to time t are

$$\begin{aligned} S_t = \Bigg(& v_0, \quad \sum_{i=1}^{t} v_i, \quad \sum_{i=1}^{t} v_{i-1}, \quad \sum_{i=1}^{t} v_{i-1}^2, \quad \sum_{i=1}^{t} v_i^2, \quad \sum_{i=1}^{t} v_i v_{i-1}, \\ & \sum_{i=1}^{t} a_i b_i, \quad \sum_{i=1}^{t} a_i b_i v_{i-1}, \quad \sum_{i=1}^{t} a_i b_i v_i, \quad \sum_{i=1}^{t} a_i^2 b_i^2, \quad \mu_0, \quad \sum_{i=1}^{t} \mu_i, \\ & \sum_{i=1}^{t} \mu_{i-1}, \quad \sum_{i=1}^{t} \mu_i^2, \quad \sum_{i=1}^{t} \mu_{i-1}^2, \quad \sum_{i=1}^{t} J_i, \quad \sum_{i=1}^{t} \kappa_i, \quad \sum_{i=1}^{t} \kappa_i^2 \quad \Bigg). \end{aligned}$$

where $a_i = y_i - \mu_{i-1} - \kappa_i J_i$ and $b_i = \exp\{-v_{i-1}/2\}$. It can be noticed that the sufficient statistics may depend on v_t and μ_t that belong to $\boldsymbol{x}_{t+1}$. In this case we estimate these quantities by simulating them from their dynamics. The amount of computer memory required is, thus, sensibly reduced.

The resulting algorithm is summarized as follows

`Parameter learning algorithm`

1. Set $t = 0$. Simulate N particles from the prior $p(\boldsymbol{\theta})$, from $p(v_0)$ and from $p(\mu_0)$ with equal weights, $J_0 = 0$ and $\kappa_0 = 0$;
2. For $t = 1$ to T:

   i) *Auxiliary particle step:*

      a) Given $\boldsymbol{x}_t^j = (v_{t-1}^j, \mu_{t-1}^j, J_t^j, \kappa_t^j, \boldsymbol{\theta}_t^j)$ and ω_t^j, $j = 1, \ldots, N$, set $\overline{J}_{t+1}^j = 0$ and compute

$$\bar{v}_t^j = E[v_t \mid v_{t-1}^j, \boldsymbol{\theta}_t^j]$$
$$\bar{\mu}_t^j = E[\mu_t \mid \mu_{t-1}^j, \boldsymbol{\theta}_t^j]$$
$$\boldsymbol{m}_t^j = a\boldsymbol{\theta}_t^j + (1-a)\bar{\boldsymbol{\theta}}_t;$$

      b) Compute $g_{t+1}^j \propto \omega_t^j p(y_{t+1} \mid \bar{v}_t^j, \bar{\mu}_t^j, \overline{J}_{t+1}^j, \boldsymbol{m}_t^j)$, such that $\sum_{j=1}^N g_{t+1}^j = 1$, $j = 1, \ldots, N$;
      c) Select with replacement from $(\boldsymbol{x}_t^j, \boldsymbol{m}_t^j)$ with probability $\{g_{t+1}^j\}$;

   ii) *Updating the parameters:*

      a) Update $\boldsymbol{\theta}_{t+1}$ from $N(\boldsymbol{m}_t^j, h^2\boldsymbol{\Sigma}_t)$, $j = 1, \ldots, N$;

   iii) *Updating the states:*

      a) Update v_t from $N(\alpha_v^j + \phi^j v_{t-1}^j, \sigma_\eta^{2\,j})$, $j = 1, \ldots, N$;
      b) Update μ_t from $N(\alpha_\mu^j + \beta^j \mu_{t-1}^j, \sigma_\mu^{2\,j})$, $j = 1, \ldots, N$;
      c) Update J_{t+1} from $Bi(1, \lambda^j)$, $j = 1, \ldots, N$;
      d) Update κ_{t+1} from $N(\mu_y^j, \sigma_y^{2\,j})$, $j = 1, \ldots, N$;

   iv) *Updating the sufficient statistics:*

      a) Update the sufficient statistics $S_{t+1}^j = f(S_t^j, \boldsymbol{x}_{t+1}^j)$, $j = 1, \ldots, N$;

   v) *Computing the importance weights:*

      a) Compute $\omega_{t+1}^j \propto \frac{p(y_{t+1} \mid v_t^j, \mu_t^j \boldsymbol{\theta}_{t+1}^j)}{p(y_{t+1} \mid \bar{\mu}_t^j, \bar{v}_t^j, \boldsymbol{m}_t^j)}$, $\sum_{j=1}^N \omega_{t+1}^j = 1$;
      b) Record $\boldsymbol{x}_{t+1}^j$, ω_{t+1}^j and S_{t+1}^j, $j = 1, \ldots, N$;

   vi) *MCMC step (optional):*

      a) Move all of the the particles through a systematic scan Gibbs sampler, i.e., $T(\boldsymbol{x}_{t+1}^j, \boldsymbol{x}_{t+1}^{\prime\,j})$ using $S_{t+1}^j, j = 1, \ldots, N$;

b) Update the sufficient statistics according to the MCMC move, that is, $S^j_{t+1} = f(S^j_t, \boldsymbol{x}'^{\,j}_{t+1}),\ j = 1, \ldots, N$.

3. End.

We perform the MCMC step through a Gibbs sampler. It is also convenient to use some transformation of the parameters $\boldsymbol{\theta}$ in order to extend their support to the real line. In fact, the posterior is approximated by a mixture of Normals, and then a convenient reparameterization of the model is in terms of parameters lying on the real line. This is important in order to perform *step 3* of the algorithm. We then consider the transformed parameter $\phi^* = \log\phi - \log(1-\phi)$ and $\beta^* = \log\beta - \log(1-\beta)$. We also define $\rho^* = \log(1+\rho) - \log(1-\rho)$. For the same reason, we consider the logarithm of σ_η, σ_μ, σ_ζ and of the intensity λ.[5]

4. SIMULATION STUDY

In this section we provide some illustrative examples to show the performance of the algorithm. All the calculations are based on software written using the Ox©5.00 language of Doornik (2001). We simulate a time series of length $T = 2000$ from the model described by Eqs. (1)–(3). The true parameters, consistent with empirical findings on similar stochastic volatility models with jumps, are the following one:

a) Volatility process: $\alpha_v = 0.06$, $\phi = 0.95$, $\sigma_\eta = 0.15$, $\rho = -0.5$;
b) Conditional mean: $\alpha_\mu = 0.001$, $\beta = 0.90$, $\sigma_\mu = 0.1$;
c) Jump Process: $\lambda = 0.01$, $\mu_y = -4$, $\sigma_y = 2$.

We approximate the posterior distributions of interest through a cloud of 25,000 particles. We use the MCMC correction for $\boldsymbol{\theta}$ every 50 iterations of the algorithm, whereas J_{t+1} and κ_{t+1} are updated systematically. This choice provides a reasonable compromise between statistical precision and computational burden. We noticed that the introduction of an MCMC correction makes the inferential procedure more robust. First, it prevents the algorithm from degenerating after a few hundred iterations, keeping $\boldsymbol{\Sigma}$ different from the null matrix and thus providing a non degenerate posterior distribution for $\boldsymbol{\theta}$. Second, our empirical evidence suggested that also the estimated latent vectors may degenerate without any correction,

[5]For the superposition model the following transformations have been used in order to achieve stationary and positive definitness of Σ: $\theta_1 = \mu$, $\theta_2 = \log(1+\phi_1) - \log(1-\phi_1)$, $\theta_3 = \log(1+\phi_1) - \log(\phi_1 - \phi_2)$, $\theta_4 = \log\sigma_1^2$, $\theta_5 = \log\sigma_2^2$, $\theta_6 = \log(1+\rho_1) - \log(1-\rho_1)$, $\theta_7 = \log\left(\sqrt{1-\rho_1^2} + \rho_2\right) - \log\left(\sqrt{1-\rho_1^2} - \rho_2\right)$.

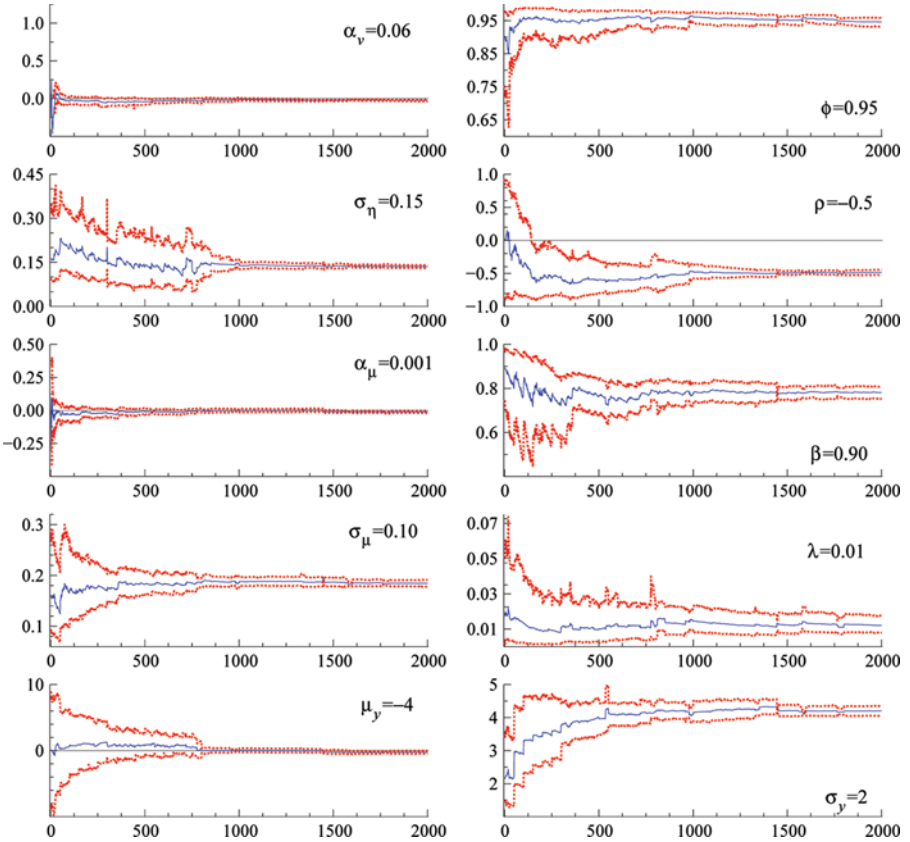

FIGURE 1 Estimated parameters together with the 2.5 and the 97.5% posterior quantiles.

providing a poor approximation of the true factors. Figure 1 reports the sequential learning process for the parameters, i.e., the evolution of the posterior mean together with the 2.5 and the 97.5% posterior quantiles.

Our algorithm provides accurate estimates for the parameters of the log-volatility process and, in fact, the posterior means of ϕ, σ_η and ρ quickly converge to their true values. In particular, the algorithm provides very precise estimates of the leverage ρ and of the persistence ϕ. It is also interesting to note the accuracy obtained for the volatility of the volatilities parameter σ_η. The top panel of Fig. 2 shows that the estimated log-volatility closely follows the true process.

More difficulties arise with the conditional mean parameters. Even though Fig. 2 suggests that the true trajectory of μ_t is well approximated, we find that the persistence parameter β is slightly underestimated, while the estimate of σ_μ is slightly larger than its true value. However, the difference between true and posterior means is not dramatic and we think this gap

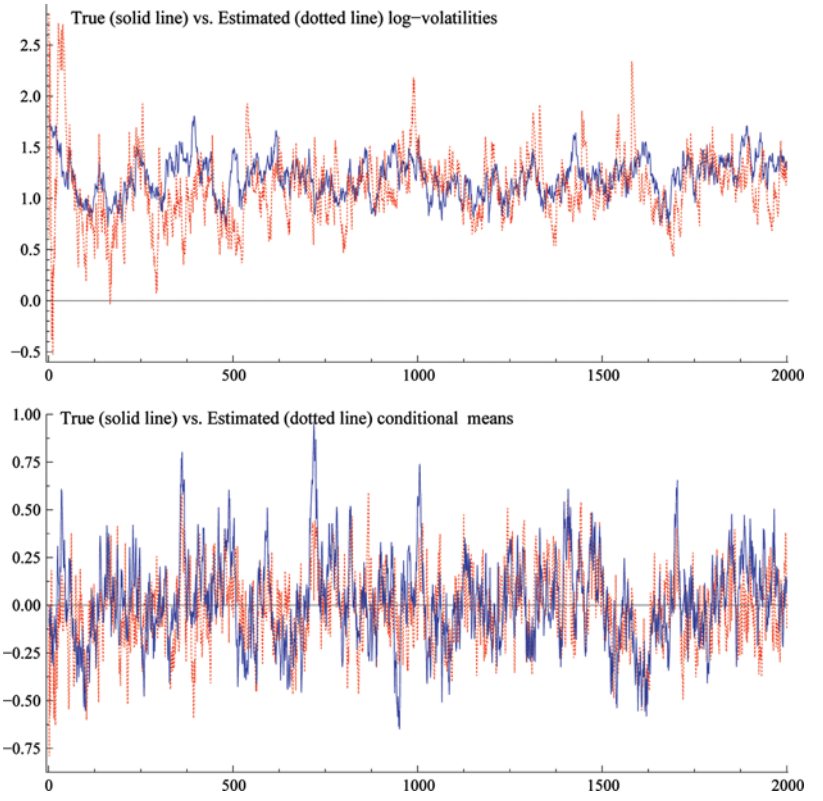

FIGURE 2 True vs. estimated log-volatilities (upper panel) and conditional means (lower panel).

can be reduced by introducing a non null correlation between y_{t+1} and μ_{t+1}, thus making the observed data more informative for these parameters.

It is interesting to note that the algorithm detects the jumps accurately. This feature is displayed in Fig. 3. In a few other cases we have noted an occasional inability of the algorithm to distinguish between outliers and actual jumps. This is especially evident when an extreme return is observed at the beginning of the series and when the jump size is small. However, Fig. 3 suggests that the algorithm is very accurate in detecting expected size and timing.[6] In some occasions difficulties arise when estimating the parameters related to the jump process, in which case some care has to be taken in the empirical analysis. The reason for these occasional pitfalls is most likely due to the rare nature of the jumps. Apart from λ, it is difficult to identify the parameters describing κ_t, i.e., μ_y and σ_y. Difficulties related

[6]Some further Monte Carlo results, not reported here, show that the algorithm perform well even when the jump's size is heterogeneous, meaning that we considered jumps that range between −7.5% and +5%.

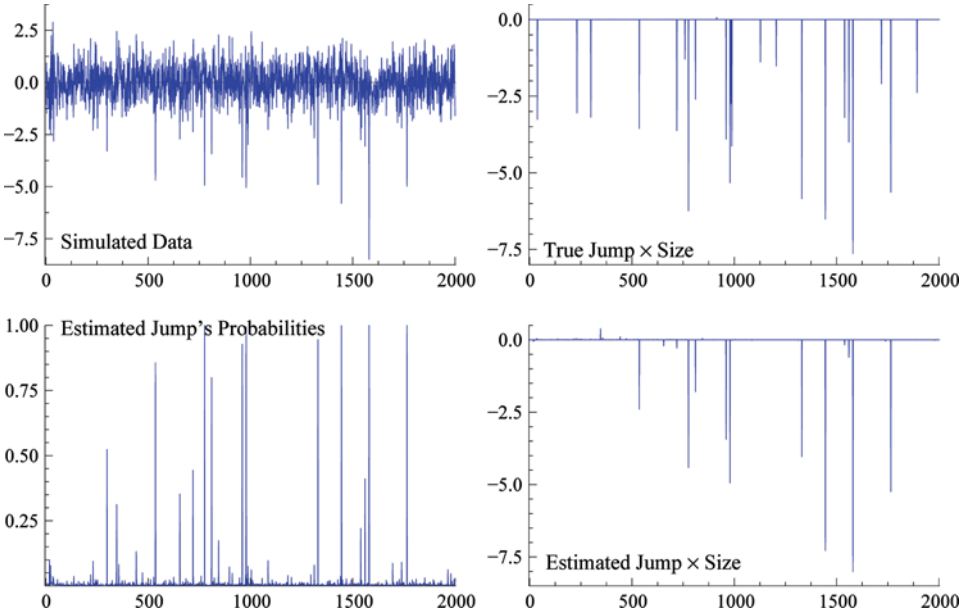

FIGURE 3 The simulated data are in the top left panel; in the bottom left panel, the estimated probabilities of jumps; on the right, the true and estimated impact of a jump event.

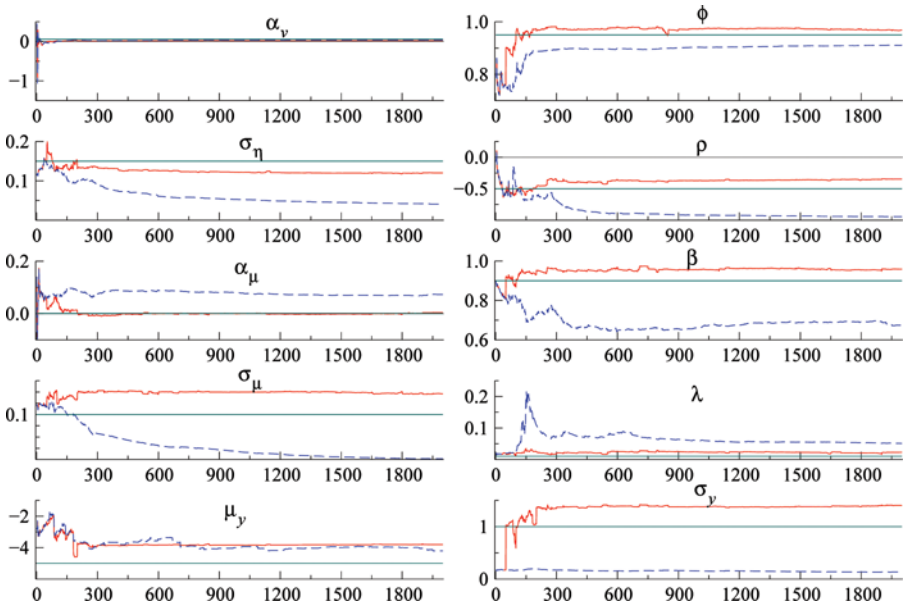

FIGURE 4 Particle filter with and without MCMC correction. Comparison between the sequential estimates with MCMC correction (solid lines) and the sequential estimates with no correction (dashed lines). The true parameter values are indicated with the horizontal lines.

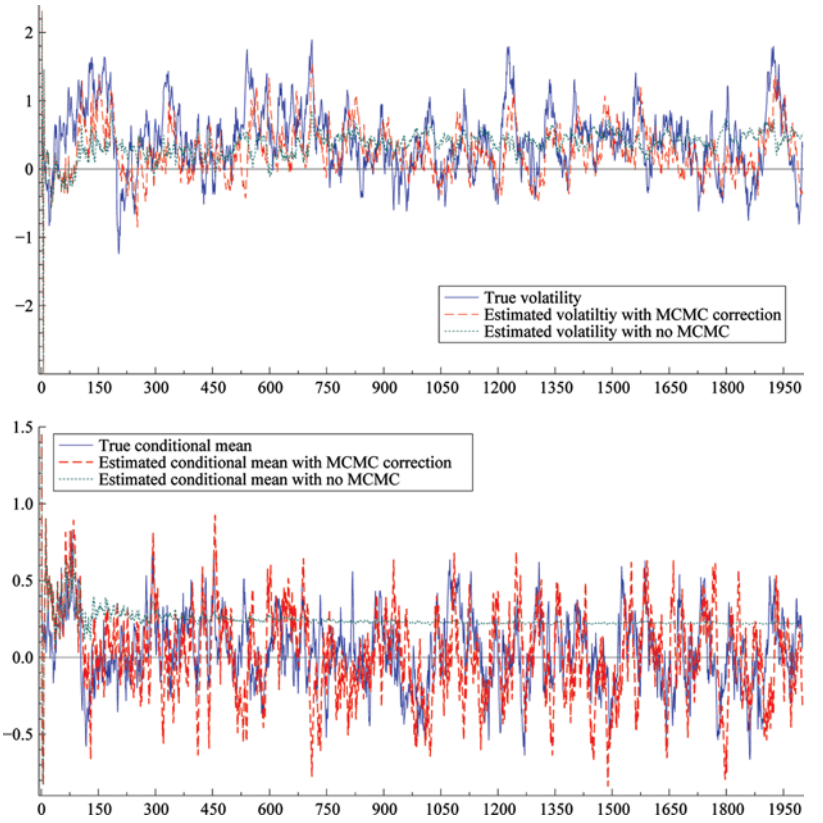

FIGURE 5 Particle filter with and without MCMC correction. Comparison between the estimated latent processes with MCMC correction (dashed lines) and the sequential estimates with no correction (dotted lines). The simulated processes are indicated with solid lines.

to the lack of identification of jump models are, however, a common problem in this field and have also been noticed in Chib et al. (2002), Eraker et al. (2003), Raggi (2005), and in Nakajima and Omori (2009).

In a second set of experiments we compared the performance of our procedure against the particle filtering algorithm that does not use the MCMC correction. Figure 4 plots the sequential estimates versus the true parameters. In particular, the estimates obtained introducing the MCMC correction are always closer to the true parameters with respect to the plain method. We also report in Fig. 5 the estimated latent states. We observed that the estimates obtained through the plain algorithm sensibly differ with respect to the *true* processes and seem to converge to the marginal means of v_t and μ_t. On the other hand, the approximation provided by our methodology closely follows the true trajectories of v_t and μ_t.

This analysis, corroborated by many other experiments not reported here for conciseness reasons, suggests that the MCMC correction is useful,

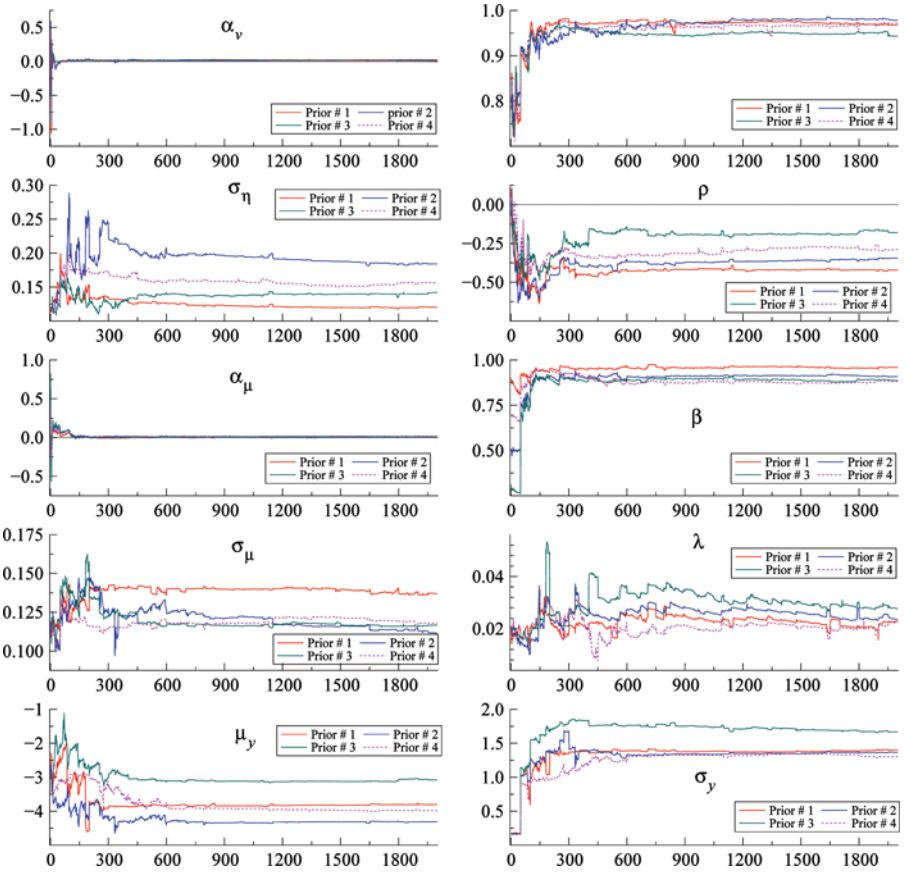

FIGURE 6 Sensitivity analysis of the algorithm for different priors for β.

and provides an improvement with respect to the plain algorithm of Liu and West (2001). We found also cases in which the MCMC correction was not strictly necessary, even though, in our experience, it has never been worst than the benchmark method.

Furthermore, to assess the robustness properties of the algorithm for inference, we also considered different initial distributions for β, λ, μ_y, and σ_y by running two more experiments. In the first case, we changed the initial distribution for β, choosing priors with expected values ranging between 0.28 and 0.92 as follows:

Prior 1: $\beta \sim \text{Beta}(25, 2) \Rightarrow \text{E}[\beta] = 0.9259$, $\sigma_\beta = 0.0495$;
Prior 2: $\beta \sim \text{Beta}(20, 20) \Rightarrow \text{E}[\beta] = 0.5$, $\sigma_\beta = 0.0781$;
Prior 3: $\beta \sim \text{Beta}(2, 5) \Rightarrow \text{E}[\beta] = 0.2857$, $\sigma_\beta = 0.1597$;
Prior 4: $\beta \sim \text{Beta}(25, 10) \Rightarrow \text{E}[\beta] = 0.7142$, $\sigma_\beta = 0.0753$;

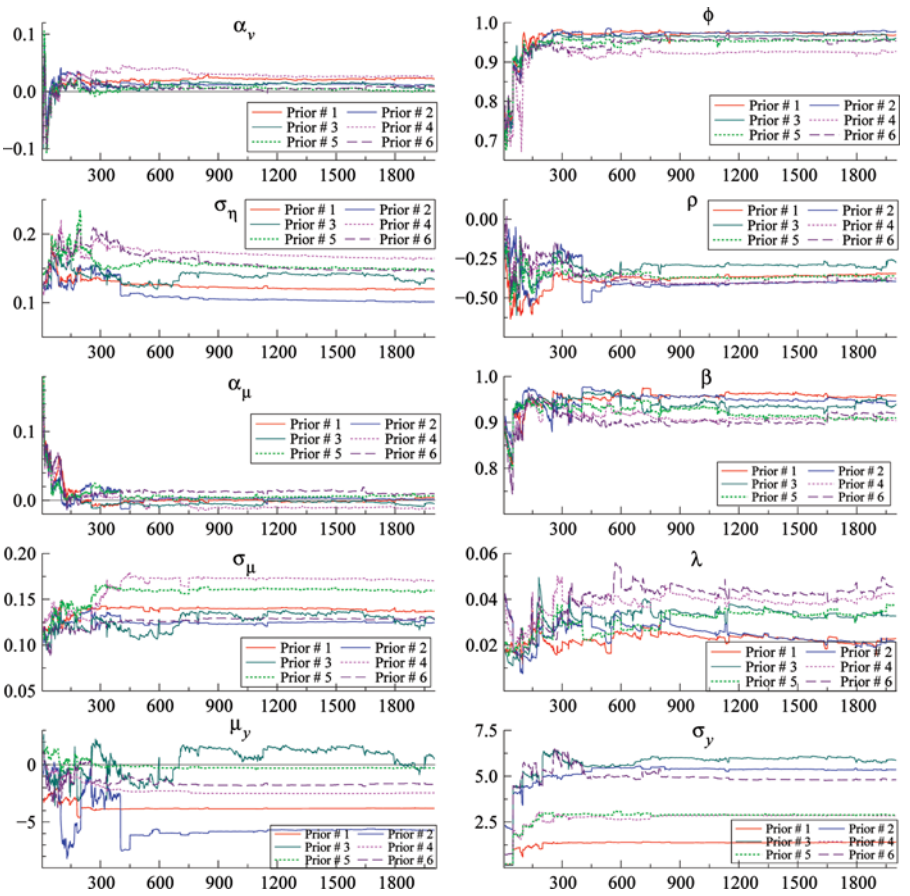

FIGURE 7 Sensitivity analysis of the algorithm for different priors for λ, μ_y and σ_y.

As shown in Fig. 6, we observed precise and stable estimates for the posterior mean of this parameter. We also noticed some small variations on the posterior means of σ_μ and σ_η. In particular, we noticed that the posterior means of σ_η ranged between 0.12 and 0.18, whereas the posterior means of σ_μ were between 0.11 and 0.135. This behavior is likely explained by the difficulty of identifying these two parameters, since they are very sensitive to outliers, as stressed in Johannes et al. (2006).

We also performed a sensitivity analysis for λ, μ_y and σ_y. We considered different initial distributions coherent with a setup characterized by infrequent but large jumps (see Eraker et al., 2003, and Nakajima and Omori, 2009, on this point). In particular, we considered the following six cases:

Prior 1: $\lambda \sim \text{Beta}(2, 100)$, $\mu_y \sim \text{N}(-3, 4)$, and $\sigma_y \sim \text{IG}(2.5, 0.05)$;
Prior 2: $\lambda \sim \text{Beta}(2, 40)$, $\mu_y \sim \text{N}(0, 100)$, and $\sigma_y \sim \text{IG}(5, 20)$;

Prior 3: $\lambda \sim \text{Beta}(2, 100)$, $\mu_y \sim \text{N}(0, 100)$, and $\sigma_y \sim \text{IG}(2.5, 0.025)$;
Prior 4: $\lambda \sim \text{Beta}(2, 40)$, $\mu_y \sim \text{N}(0, 20)$, and $\sigma_y \sim \text{IG}(2.5, 0.05)$;
Prior 5: $\lambda \sim \text{Beta}(2, 100)$, $\mu_y \sim \text{N}(0, 20)$, and $\sigma_y \sim \text{IG}(2.5, 0.05)$;
Prior 6: $\lambda \sim \text{Beta}(5, 100)$, $\mu_y \sim \text{N}(0, 100)$, and $\sigma_y \sim \text{IG}(10, 5)$;

Empirical results, reported in Fig. 7, evidenced that all of the other parameters were insensitive to the prior's choice, even though some changes occurred for μ_y and σ_y. On the other hand, the estimated latent processes remained substantially unchanged, and all of the jumps have been precisely detected, with high probability and correct size. This finding evidenced some lack of identification for these two parameters that is related with the rare nature of the process they describe.

5. DOW JONES INDEX

In this section we report some empirical results based on the Dow Jones index observed daily from January 1985 to mid August 2009. The data set has been downloaded from Yahoo Finance. As usual, the returns are defined as $y_{t+1} = 100 \times (\log p_{t+1} - \log p_t)$. The upper panels of Fig. 9 provide the plot of the observed time series. We estimate the model by approximating the distributions of interest through 25,000 particles. We initially estimated the model through the algorithm of Liu and West (2001) and we noticed that, after a few hundred iterations, $\boldsymbol{\Sigma}$ collapses to the null matrix inducing a posterior distribution of $\boldsymbol{\theta}$ approximated by just one particle with probability one and all of the others with null weights (Fig. 8).

We then compare the parameter learning algorithm with MCMC correction to the standard single-move Gibbs sampler proposed in Jacquier et al. (2004), in which the results are obtained by running the algorithm for 50,000 iterations with a burn-in of 25,000. The output of both the procedures is summarized in Table 1 and in Figs. 9 and 10. More precisely, Fig. 9 compares the particle filter estimates of the latent processes with the full MCMC posterior averages, whereas Fig. 10 shows the sequential estimates of the parameters. Finally, Table 1 provides the full sample comparison between the sequential and the full MCMC estimation.

We first comment the particle filter results and then we compare them with the MCMC output. The left panels of Fig. 9 report the sequential posterior means of the latent processes. For the log-volatility and the conditional mean we also give 95% confidence bands. It is remarkable to note that associated with each spike in the original data set there is an estimated high probability of jump. This is particularly evident for the crash observed during October 1987 in which we observe a daily move of about -25%. It is also clear from the bottom-left panel in Fig. 9, that the

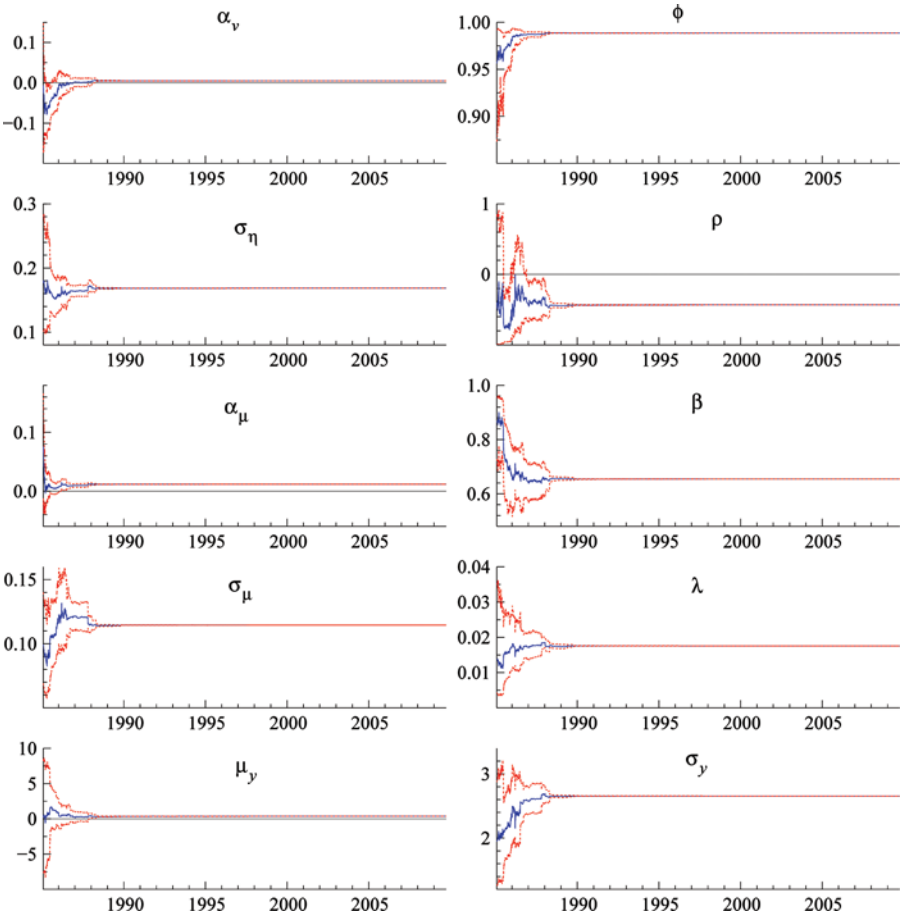

FIGURE 8 Dow Jones Index. Sequential estimate of the posterior means of the parameters of the model (solid line) and the 95% confidence bounds (dotted line).

particle filter detects with high probability almost all the jumps observed between 1997 to 2003. Together with the jumps, it is easy to note that the log-volatility bursts every time a jump is detected, which is a reasonable feature since an extreme and negative event leads to a sudden and large increase in the variability of the financial asset. In the last two years of the crisis, the algorithm detects just one jump. This is in some ways surprising, even though this behavior can be easily explained by a very high volatility that accomodates for extreme observations. The introduction of the jump factor is also useful in explaining the second moment of the interest rates process. We compute the ratio $\frac{\text{Var}[(J_t \kappa_t)]}{\text{Var}[y_t]}$ which expresses the percentage of the total variance due to jumps. In our analysis, we find that jumps explain 21% of the total variance.

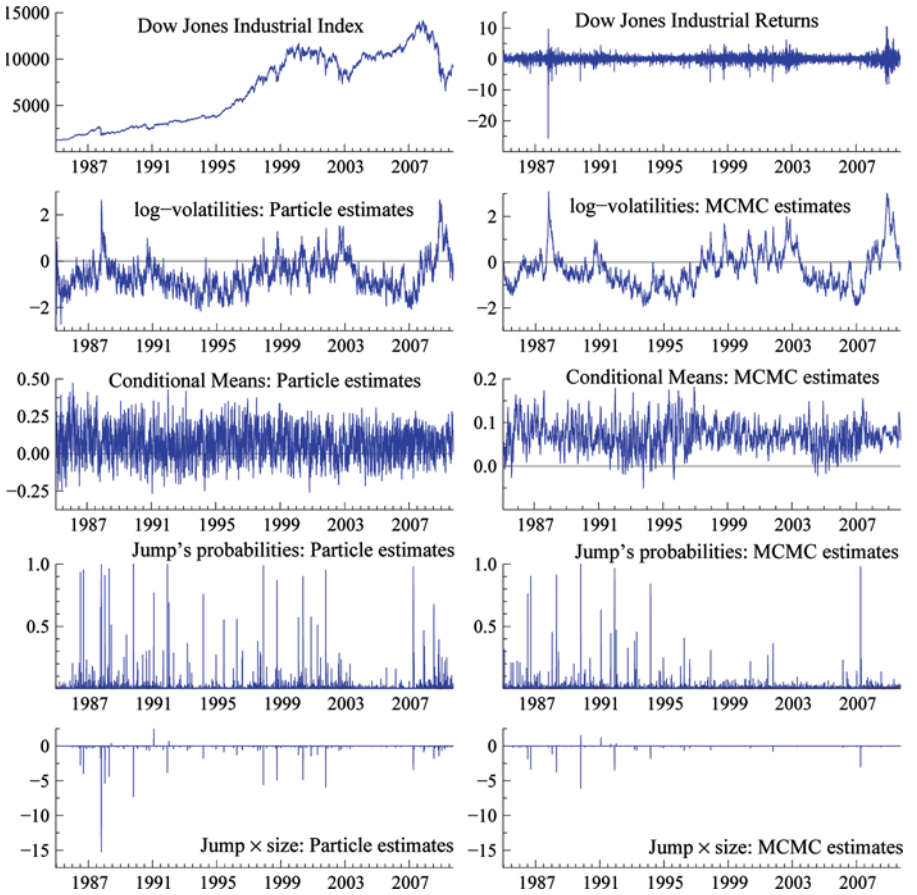

FIGURE 9 Dow Jones Index. On the upper-left panel, we report the observed prices, whereas on the left, we plot the returns. On the other panels, we show the sequential, and the MCMC posterior means for the latent processes. On the left panels, we report the sequential estimates, whereas on the right we report the MCMC estimates.

In Fig. 10 we show the sequential evolution of the parameters' posterior means. The estimate of ρ is approximately -0.32 and confirms a marked leverage effect, since it is negative and substantially different from zero. The log-volatility process is persistent since ϕ is greater than 0.97. We found that ϕ tends to increase slightly in time, but this behavior can be explained by the rising volatility observed during the last years. The parameter σ_η is approximately 0.15. These results are consistent with the current literature on stochastic volatility. The analysis of μ_t provides evidence about the predictability of the returns. The intercept α_μ is positive but close to zero, and the persistence parameter β converges to 0.87. This high estimate of β clearly implies a non-null autocorrelation of μ_t and suggests that the effect of a jump is persistent over time, thus influencing

TABLE 1 Dow Jones Index. Stochastic volatility with jumps model. Comparison between sequential and MCMC estimates. We report the posterior means and the 95% confidence bounds

	Particle		MCMC	
	Mean	95% Conf. Int.	Mean	95% Conf. Int.
$p(\alpha_v \mid y_{1:T})$	−0.01694	[−0.021, −0.012]	−0.00272	[−0.006,0.0008]
$p(\phi \mid y_{1:T})$	0.97199	[0.966,0.976]	0.98866	[0.984,0.992]
$p(\sigma_\eta \mid y_{1:T})$	0.15353	[0.150,0.156]	0.13815	[0.120,0.157]
$p(\rho \mid y_{1:T})$	−0.32867	[−0.351,−0.306]	−0.65645	[−0.756,−0.561]
$p(\alpha_\mu \mid y_{1:T})$	0.01091	[0.008,0.013]	0.01610	[0.009,0.023]
$p(\beta \mid y_{1:T})$	0.87814	[0.863,0.891]	0.76834	[0.689,0.835]
$p(\sigma_\mu \mid y_{1:T})$	0.08940	[0.087,0.091]	0.07247	[0.053,0.094]
$p(\lambda \mid y_{1:T})$	0.01097	[0.008,0.013]	0.00996	[0.003,0.020]
$p(\mu_y \mid y_{1:T})$	−1.38785	[−1.493,−1.286]	−0.74232	[−2.634,0.317]
$p(\sigma_y \mid y_{1:T})$	3.97957	[3.9054.053,]	2.16437	[1.602,2.870]

future returns. We think it is important to notice this feature, since in the current literature jumps are often taken to be independent with a transient impact on returns. This is one of the reasons why jumps are usually added to the volatility process. It is interesting to compute the half-life of the two autoregressive processes, defined as the number of periods required for the impulse response to a unit shock to a time series to dissipate by half. In practice, if the persistence parameter is ϕ, the half-life is defined as $\frac{\log 0.5}{\log \phi}$. The half-life for the log-volatility process is about 24.4, whereas for the conditional mean it is 5.33. These quantities imply that it takes about one month for the volatility and about one week for the conditional mean to absorb 50% of a shock.

Finally, we consider the parameter estimates related to J_t and κ_t. The intensity λ suggests that the model detects about five extreme events per year. Concerning μ_y and σ_y, the expected size and the variability of κ_t, we obtain that $\mu_y \approx -1.3$ and $\sigma_y \approx 4$. This high value of σ_y implies that the impact of jumps on the returns is heterogeneous. More precisely, it seems that the model accurately describes the timing of the jumps, but their effect is quite variable. The estimates reported, in fact, indicate that κ_t likely ranges between −9.5 and +6.5 percent.

This analysis suggests that the model can be generalized to allow for a time dependent intensity λ_t. On closer inspection, Fig. 9 suggests that jumps arrive in clusters. For example, we estimate many jumps of large size between 1986 and 1992, none in the subsequent six years, several jumps again between 1998 and 2003 and then just one in the last two years. It is also easy to note that jumps of large size are more frequent in periods with high volatility, thus suggesting that the intensity λ and the jump's size κ_t

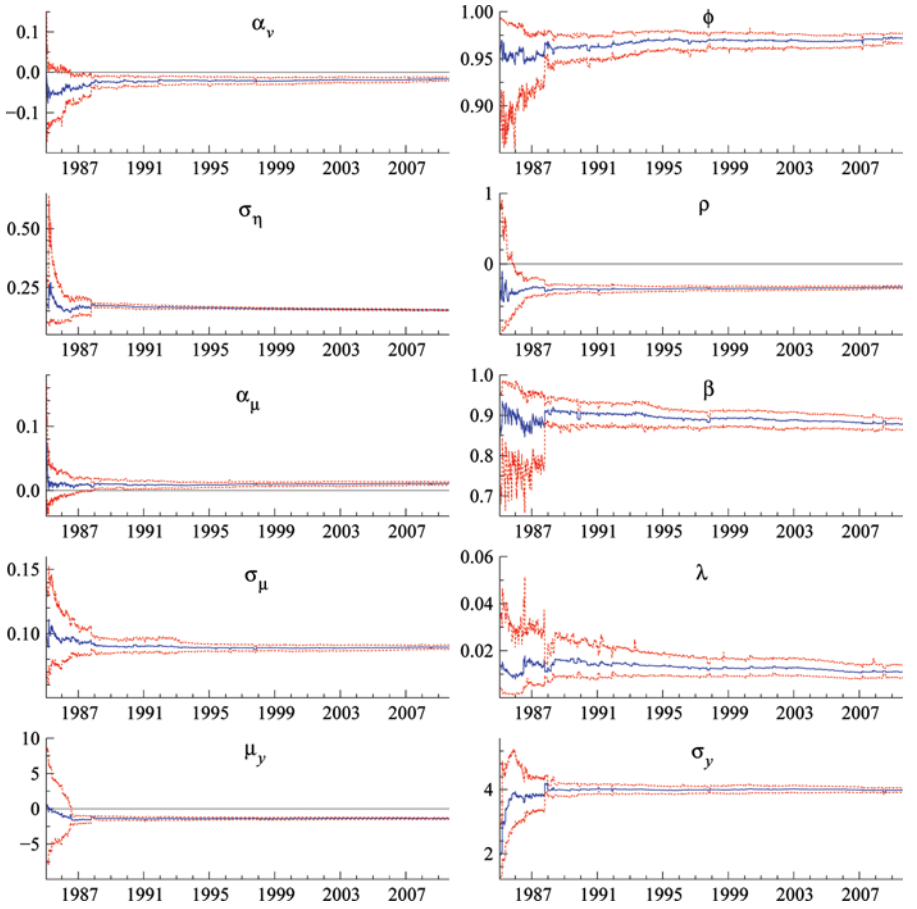

FIGURE 10 Dow Jones Index. Sequential estimate of the posterior means of the parameters of the model (solid line) and the 95% confidence bounds (dotted line).

may be time varying and dependent on the volatility. Surprisingly, in the last two years, we do not observe many jumps. This finding, however, can be explained by an extremely high volatility that accomodates for extreme events.

We now compare our results with those obtained through the standard MCMC procedure. In Table 1 we report the full sample MCMC posterior estimates and the posterior means estimated at the last observation of the particle filter, whereas plots of the MCMC estimated latent processes are on the right panels of Fig. 9. The second and third row panels evidence a substantial equivalence between the estimates of the log-volatilities and of the conditional means. Surprisingly, the MCMC procedure seems to be inadequate to detect the major jumps observed in the real data and associates to these events a low probability of jump. For example, the

crash of October 1987 is detected with low probability (about 10%) and negligible size, whereas the sequential procedure estimates the same jump with probability 1 and size of about −16%. A comparison with an independent benchmark suggests also that the sequential approach outperform MCMC on detecting jumps. In our analysis, the particle filtering procedure recognize more than 60% of the jumps detected by the Lee and Mykland (2008) test whereas MCMC detects just about the 30%. However, it is important to stress that this benchmark can be imprecise when daily data are taken into account.

Finally, Table 1 compares the posterior means and the confidence bounds of the parameters. There is a substantial equivalence of the two methods for almost all the parameters. The major differences concern σ_y, that halves with MCMC, and ρ. However, these differences can be explained by the inability of MCMC to detect the major jumps.

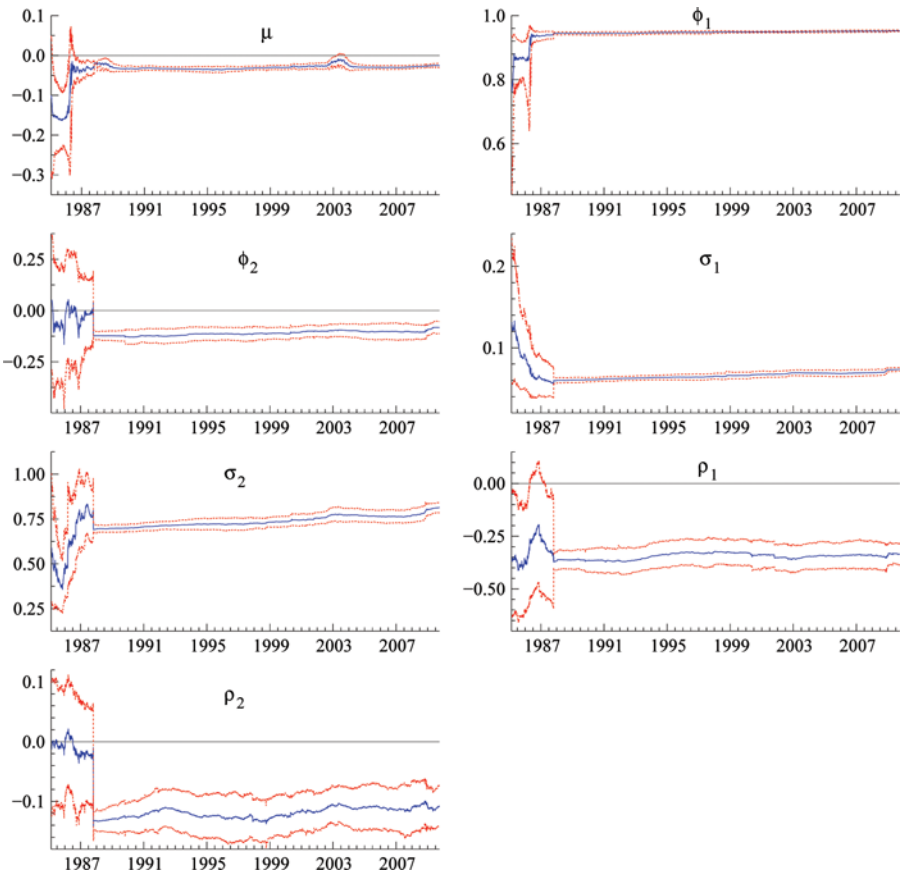

FIGURE 11 Dow Jones Index. Estimated parameters of the superposition model (solid line) and the 95% confidence bounds (dotted line).

We further report the results for the superposition model. In this case we do not consider the conditional mean. Figure 11 displays the sequential estimates of the parameters, whereas Fig. 12, shows the estimated volatility factors.

Our results are in line with Liesenfeld and Richard (2003) that consider a similar two factors specification without leverage effects. We observe that σ_2 is much larger than σ_1. Furthermore, the estimated value of the persistence parameter of the long-run component is about 0.95, while those of the short-run component is -0.08. The two correlation parameters are negative, in particular $\rho_1 = -0.33$ and $\rho_2 = -0.10$, thus indicating a leverage effect due to both of the factors. We observe from Fig. 12 that $v_{1,t}$ is underestimated with respect to the volatility model with jumps; however, $v_{2,t}$ compensates for this. Furthermore, the short-run volatility sensibly increases in correspondence to extreme observations, thus suggesting that this factor can be useful to describe the tail behavior of the return's distribution.

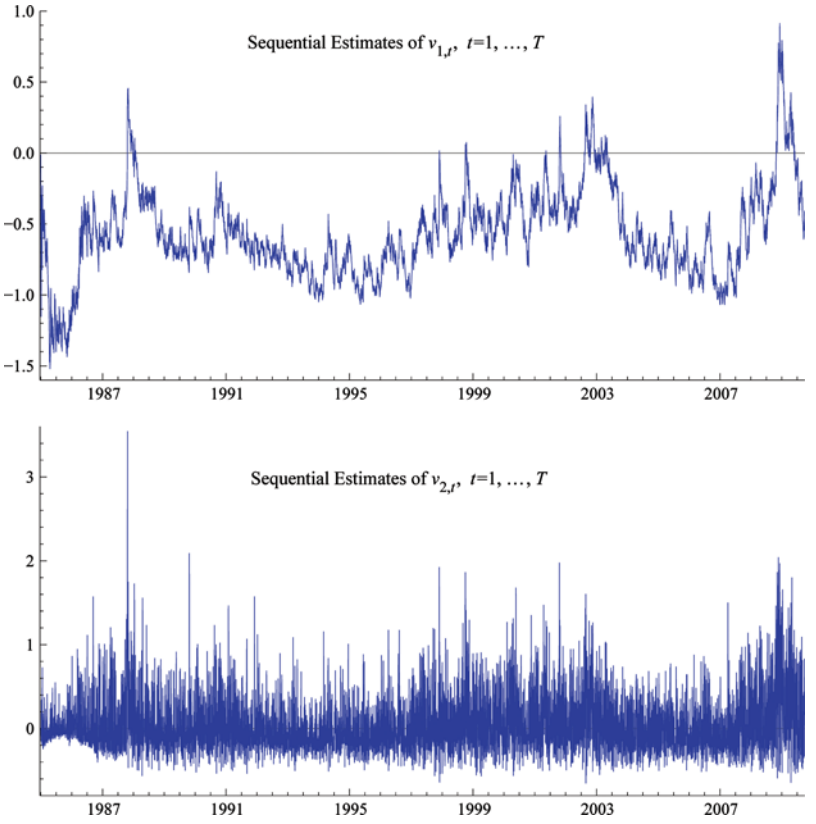

FIGURE 12 Dow Jones Index. Estimated volatilities $v_{i,t}, i = 1, 2$.

6. CONCLUSIONS

Sequential Monte Carlo methods represent a valuable and reliable methodology to estimate nonlinear and nongaussian state-space models. Their application also seems to be useful in the analysis of stochastic volatility models. In this article, we have proposed an algorithm based on the kernel smoothing approximation of the posterior suggested in Liu and West (2001), in which an MCMC step is incorporated in order to reduce sampling impoverishment problems related to sequential Monte Carlo strategies. Analysis based on simulated and real data demonstrate the effectiveness of our proposal, since the algorithm provides robust and stable results with longer time series that are typical in financial econometrics.

ACKNOWLEDGMENTS

Financial support from the MIUR under grant PRIN 2008 Prot. N. 2008MRFM2H_003 is gratefully acknowledged. We wish to thank two anonymous referees for helpful comments, Stuart Coles, Andrea Torsello, and the participants at the Workshop *New Directions in Monte Carlo Methods* held in Fleurance (France) on June 2007. All errors are our own.

REFERENCES

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed-form approximation approach. *Econometrica* 70:223–262.

Arulampalam, M. S., Maskell, S., Gordon, N., Clapp, T. (2002). A tutorial on particle filters for on-line nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing* 50:174–188.

Asai, M., McAleer, M. (2005). Dynamic asymmetric leverage in stochastic volatility models. *Econometric Reviews* 24:317–332.

Asai, M., McAleer, M., Yu, J. (2006). Multivariate stochastic volatility: A review. *Econometric Reviews* 25:145–175.

Bakshi, G., Cao, C., Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance* 52:2003–2049.

Bos, C. S., Shephard, N. (2006). Inference for adaptive time series models: stochastic volatility and conditionally gaussian state space form. *Econometric Reviews* 25:219–244.

Carvalho, C., Lopes, H. (2007). Simulation-based sequential anlysis of markov switching stochastic volatility models. *Computational Statistics & Data Analysis* 51:4526–4542.

Chan, D., Kohn, R., Kirby, C. (2006). Multivariate stochastic volatility models with correlated errors. *Econometric Reviews* 25:254–274.

Chan, N. H., Palma, W. (1998). State space modeling of long-memory processes. *Annals of Statistics* 26:719–740.

Chan, N. H., Petris, G. (2000). Long memory stochastic volatility: a bayesian approach. *Communications in Statistics: Theory and Methods* 29:1367–1378.

Chernov, M., Gallant, A., Ghysels, E., Tauchen, G. (2003). Alternative models for stock price dynamics. *Journal of Econometrics* 116:225–257.

Chib, S., Nardari, F., Shephard, N. (2002). Markov chain monte carlo methods for stochastic volatility models. *Journal of Econometrics* 108:281–316.

Chib, S., Nardari, F., Shephard, N. (2005). Analysis of high dimensional multivariate stochastic volatility models. *Journal of Econometrics* 134:341–371.

Danielsson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. *Journal of Econometrics* 64:375–400.

Doornik, J. (2001). *Ox: An Object-Oriented Matrix Programming Language*. London: Timberlake Consultants Press.

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

Duffie, D., Pan, J., Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica* 68:1343–1376.

Durbin, J., Koopman, S. (1997). Monte carlo maximum likelihood estimation of non-gaussian state space model. *Biometrika* 84:669–684.

Durham, G. (2006). Monte carlo methods for estimating, smoothing, and filtering one-and two-factor stochastic volatility models. *Journal of Econometrics* 133:273–305.

Durham, G., Gallant, R. (2002). Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes (with comments). *Journal of Business & Economic Statistics* 20:297–337.

Elerian, O., Chib, S., Shephard, N. (2001). Likelihood inference for discretely observed non-linear diffusions. *Econometrica* 69:959–993.

Engle, R. F., Lee, G. (1999). A long-run and short-run Component model of stock return volatility. In: Engle, R., White, H., eds. *Cointegration, Causality and Forecasting – A Festschrift in Honour of Clive W.J. Granger*. New York: Oxford University Press, pp. 475–497.

Eraker, B. (2004). Do equity prices and volatility jumps?. Reconciling evidence from spot and option prices. *Journal of Finance* 59:1367–1403.

Eraker, B., Johannes, M., Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance* 58:1269–1300.

Fearnhead, P. (2002). Markov chain monte carlo, sufficient statistics and particle filters. *Journal of Computational and Graphical Statistics* 11:848–862.

Gallant, A., Tauchen, G. (1996). Which moments to match? *Econometric Theory* 12:657–681.

Ghysels, E., Harvey, A., Renault, E. (1996). Stochastic volatility. In: Rao, C., Maddala, G., eds. *Statistical Methods in Finance*. Amsterdam: North Holland, pp. 119–191.

Gilks, W., Berzuini, C. (2001). Following a moving target - monte carlo inference for dynamic bayesian model. *Journal of the Royal Statistical Society, Ser. B* 63:127–146.

Golightly, A., Wilkinson, D. (2006). Bayesian sequential inference for nonlinear multivariate diffusions. *Statistics and Computing* 16:323–338.

Gordon, N., Salmond, D., Smith, A. (1993). Novel approach to non-linear and non-gaussian bayesian state estimation. *IEE Proceedings-F* 140:107–113.

Harvey, A. C., Shephard, N. (1996). Estimation of an asymmeric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics* 14:429–434.

Jacquier, E., Polson, N., Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business & Economic Statistics* 12:371–417.

Jacquier, E., Polson, N., Rossi, P. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics* 122:185–212.

Johannes, M., Polson, N., Stroud, J. (2002). *Sequential Optimal Portfolio Performance: Market and Volatility Timing*. Working paper, Columbia University.

Johannes, M., Polson, N., Stroud, J. (2006). *Sequential Parameter Estimation in Stochastic Volatility Models with Jumps*. Working paper, Graduate School of Business, University of Chicago.

Johannes, M., Polson, N., Stroud, J. (2009). Optimal filtering of jump diffusions: Extracting latent states from asset prices. *The Review of Financial Studies* 22:2759–2799.

Jungbacker, B., Koopman, S. J. (2006). Monte carlo likelihood estimation for three multivariate stochastic volatility models. *Econometric Reviews* 25:385–408.

Kim, S., Shephard, N., Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies* 65:361–393.

Kitagawa, G. (1998). Self organizing state space models. *Journal of the American Statistical Association* 93:1203–1215.

Lee, S., Mykland, P. (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *The Review of Financial Studies* 21:2535–2563.

Liesenfeld, R., Richard, J. (2003). Univariate and multivariate stochastic volatility models: estimation and diagnostics. *Journal of Empirical Finance* 10:505–531.

Liesenfeld, R., Richard, J. (2006). Classical and bayesian analysis of univariate and multivariate stochastic volatility models. *Econometrics Review* 25:335–360.

Liu, J., West, M. (2001). Combined parameter and state estimation in simulation-based filtering. In: Doucet, A., de Freitas, N., Gordon, N., eds. *Sequential Monte Carlo Methods in Practice*. New York: Springer-Verlag, pp. 197–223.

Liu, J., Longstaff, F., Pan, J. (2003). Dynamic asset allocation with event risk. *Journal of Finance* 63:231–259.

Malik, S., Pitt, M. (2009). *Modelling Stochastic Volatility with Leverage and Jumps: A Simulated Maximum Likelihood Approach via Particle Filtering*. Working paper, The University of Warwick, Department of Economic.

Merton, R. C. (1971). Optimum consumption and portfolio rule in continuous time. *Journal of Economic Theory* 3:373–413.

Nakajima, J., Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics & Data Analysis* 53:2335–2353.

Omori, Y., Chib, S., Shephard, N., Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics* 140:425–449.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics* 63:3–50.

Pitt, M. (2002). *Smooth Particle Filters for Likelihood Evaluation and Maximization*. Working paper, University of Warwick.

Pitt, M. K., Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association* 94:590–599.

Polson, N. G., Stroud, J. R., Müller, P. (2008). Practical filtering with sequential parameter learning. *Journal of the Royal Statistical Society B* 70:413–428.

Raggi, D. (2005). Adaptive MCMC methods for inference on affine stochastic volatility models with jumps. *Econometrics Journal* 8:235–250.

Raggi, D., Bordignon, S. (2011). Long memory and nonlinearities in realized volatility: A Markov Switching approach. *Computational Statistics and Data Analysis* (forthcoming).

Shephard, N., Pitt, M. (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika* 84:653–667.

Storvik, G. (2002). Particle filters for state-space models with the presence of unknown static parameters. *IEEE Transactions on Signal Processing* 50:281–289.

West, M. (1993). Approximating posterior distributions by mixture. *Journal of the Royal Statistical Society, Ser. B* 55:409–422.

Yu, J. (2005). On leverage in a stochastic volatility model. *Journal of Econometrics* 127:165–178.

Yu, J., Meyer, R. (2006). Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews* 25:361–384.