Sequential Monte Carlo Methods for Stochastic Volatility Models with Jumps

Davide Raggi[1*] **and Silvano Bordignon**[2]

[1] *University of Bologna*
Department of Economics
P.zza Scaravilli 2, 40126 Bologna Italy
e-mail: draggi@economia.unibo.it

[2] *University of Padova*
Department of Statistics
Via C. Battisti 241-243, 30123, Padova, Italy
e-mail: bordigno@stat.unipd.it

Abstract: In this paper we propose a sequential Monte Carlo algorithm to estimate a stochastic volatility model with leverage effect, non constant conditional mean and jumps. Our idea relies on the auxiliary particle filter algorithm together with the Markov Chain Monte Carlo (MCMC) methodology. Our method allows to sequentially evaluate the parameters and the latent processes involved in the dynamic of interest. An empirical application on simulated data and on the Standard & Poor's 500 index is presented to study the performance of the algorithm implemented.

Keywords: Stochastic volatility with jumps, leverage, return's predictability, Bayesian estimation, auxiliary particle filters, MCMC.
JEL classifications: C11, C15, G11.

1. Introduction

In this paper we propose a methodology to analyze the sequential parameter learning problem for a stochastic volatility model with jumps and a predictable component. We aim at updating the estimates of the parameters of interest together with the states on a daily basis according to the flow of new information that sequentially arrives in the markets. Sequential procedures seem suitable when we are interested to real time application, in which case we need to update regularly our estimates at each time. Economic agents need to produce estimates and forecasts in real time. In financial application this means that we want to adapt our estimates every time a new observation is available. These are the main reasons we think sequential methods are appealing from a practical and a theoretical point of view.

One of the most compelling advantage of sequential Monte Carlo methods is the reduced computational burden required if compared with other Monte Carlo procedures such as MCMC. The main limitation of MCMC in real time

*Financial support from the MIUR under grant PRIN 2005 Prot. N. 2005132539 and Prot. N. 2002135473 are gratefully acknowledged.

applications is that every time we observe new data, we have to restart the inferential procedure from the scratch causing a waste of time.

Our procedure builds on the particle filtering algorithm of Liu and West (Liu & West 2001) in which we integrate an MCMC step to prevent the algorithm to degenerate after a number of iterations. The use of MCMC together with particle filters has been proposed in Gilks & Berzuini (2001) and Berzuini & Gilks (2001) and has been proved to be an effective combination between the computational advantages of sequential algorithms and the statistical efficiency of the MCMC methods. The introduction of the MCMC step is really useful when dealing with long time series, since it sensibly reduces the degeneration troubles connected with sequential Monte Carlo methods.

We apply our methodology in a stochastic volatility context. Time varying conditional variance modelling represents an important topic for financial applications. A large literature has grown up on describing financial time series using stochastic volatility models (see Taylor 1994 and Ghysels et al. 1996 for a review). Furthermore, the introduction of a jump component have been proved to give an improved fit of the data, both in relation to their ability to describe the return's behaviour (Eraker et al. 2003), as well as for the pricing of financial derivatives (see Bakshi et al. 1997, Pan 2002 and Eraker 2004 amongst other).

Several variants of ARCH and SV models have been proposed so far to account for the empirical regularities of financial time series. Amongst these regularities three are tackled in this paper within a stochastic volatility 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 dynamic to describe extreme and rare events such as crashes on the market.

The leverage effect has been deeply investigated in the GARCH setting in Nelson (1991) whereas in a stochastic volatility framework this problem has been tackled in Yu (2005). This characteristic describe the relationship between returns and conditional variances. It is in fact reasonable to think that bad news in the markets, i.e. when the price decreases, leads to a boost on the variance, that is a measure for the financial risks. Looking from the opposite perspective, episodes of high volatility induce expectations of lower future returns, hence the negative correlation between these shocks. Furthermore the leverage effect seems also important to explain some characteristics of the data on financial derivatives.

Second, in financial applications there is a substantial evidence of some predictability on the returns. This finding has been noticed since the early works of Merton (1971), that gave a theoretical justification for this evidence. In applications related to optimal portfolio choices, it is important to take into account this predictable component. In fact, the economic theory shows that an investor gains from market predictability and volatility timing, but the impact of these benefits are difficult to quantify. This is why it is interesting to explicitly model the conditional expected value of the returns together with the dynamic of the volatilities.

Third, in the recent literature, there is a substantial evidence in favor of jumps on returns and volatilities. In fact, a diffusive behaviour of these two processes seems to be inadequate to describe the underlying dynamics (Eraker et al. 2003, Raggi 2005). This mis-specification has also some consequences when pricing european options (see for example Pan 2002 and Eraker 2004). Furthermore, 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).

The remainder of the paper is organized as follows. The basic model is described in Section 2. Our inferential solution for that class of models is outlined in Section 3. Finally some empirical results based on simulated data and on the analysis of the Standard & Poor's 500 index are illustrated in section 4.

2. The Model

A stochastic volatility model for the observable return process is usually specified as

$$y_{t+1} = \mu_t + \exp\{v_t/2\}\epsilon_{t+1} + \kappa_{t+1}J_{t+1} \tag{2.1}$$

$$v_{t+1} = \mu + \phi v_t + \sigma_\eta \eta_{t+1} \tag{2.2}$$

$$\mu_{t+1} = \alpha + \beta\mu_t + \sigma_\mu \zeta_{t+1}. \tag{2.3}$$

Returns are defined as the first difference on the logarithm of the observed prices, i.e. $y_{t+1} = \log p_{t+1} - \log p_t$. In this framework we assume that the error term ϵ_{t+1} is a standardized Gaussian white noise sequence. The conditional mean μ_{t+1} and the logarithm of the conditional variance or volatility v_{t+1} are described by two non observable processes.

The autoregressive specification of the conditional variance is an approximation of the Euler discretization of the continuous time dynamic proposed in Hull & White (1987) and in Heston (1993). We assume that the initial state v_0, i.e. the volatility at time 0, is distributed according to

$$N\left(\frac{\mu}{1-\phi}; \frac{\sigma_\eta^2}{1-\phi^2}\right),$$

that is the invariant law of the autoregressive model, identified by the first two marginal moments of the log-volatility process. The parameter ϕ is the persistence of the volatility process that allows also for the volatility clustering feature. In empirical applications this parameter is close to 1 even though it is assumed that $|\phi| < 1$. This condition imply that the conditional variance process is stationary, condition that is inherited by the returns. μ is the drift component of the process and σ_η can be interpreted as the volatility of the volatility factor. Furthermore the error term η_{t+1} is a Gaussian white noise. In order to describe the leverage effect, we assume that $\text{Cov}(\epsilon_{t+1}, \eta_{t+1})$ are correlated with correlation ρ. This parameter is in general negative, due to the relation between returns and risks described above, even though in some application, such as the exchange rates, its estimate is close to zero.

In the recent econometric literature this basic model has been extended in order to capture more characteristics of the data. In order to properly describe extreme events such as crashes in the markets a useful extension is to introduce a jump component in the return and in the volatility equations. Duffie et al. (2000) introduce a model based on a stochastic differential equation with discontinuities described by a marked point process that affects both the return and the volatility processes. In the discrete time counterpart of this model the discontinuities are generated by a sequence of independent Bernoulli random variables with fixed intensity[1] λ.

To each jump is associated a size or mark, described by a Gaussian random variable with mean μ_y and variance σ_y^2.

We also directly model the drift of the returns via further autoregressive process μ_{t+1}. Following Chernov et al. (2003) the serial correlation on the returns induced by μ_{t+1} can be explained by the non-synchronous trading and unexpected stochastic dividend effects. The specification for expected returns μ_{t+1} captures the common view that expected returns have a mean-reverting component. The parameters of interest for this equation have basically the same interpretation given before for the volatility's process parameters. As for the conditional variance, the conditional mean at time 0 is distributed as

$$\mu_0 \sim N\left(\frac{\alpha}{1-\beta}; \frac{\sigma_\mu^2}{1-\beta^2}\right)$$

where the quantities $\frac{\alpha}{1-\beta}$ and $\frac{\sigma_\mu^2}{1-\beta^2}$ are respectively the marginal mean and variance of the returns. A similar dynamic for the conditional mean has been studied recently in Johannes et al. (2002*b*).

In this paper we assume that the noise process ζ_{t+1} is uncorrelated with ϵ_{t+1} and η_{t+1} even if there are not theoretical motivations to impose this constraint.

We need also to define the prior distribution for the parameter vector θ. Our choice is consistent with Kim et al. (1998) and Eraker et al. (2003). We thus hypothesize the following prior distributions

- $\mu \sim \mathrm{N}(0; 10)$;
- $\phi \sim \mathrm{Beta}(25; 2)$;
- $\sigma_\eta^2 \sim \mathrm{IG}(2.5; 0.05)$;
- $(2\rho - 1) \sim \mathrm{U}_{(0,1)}(0.5; 1)$;
- $\alpha \sim \mathrm{N}(0; 4)$;
- $\beta \sim \mathrm{Beta}(25; 2)$;
- $\sigma_\zeta^2 \sim \mathrm{IG}(2.5; 0.05)$
- $\lambda \sim \mathrm{Beta}(2; 100)$
- $\mu_y \sim \mathrm{N}(0; 20)$

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

- $\sigma_y^2 \sim \mathrm{IG}(2.5; 0.05)$.

With IG we indicate the inverse of a Gamma distribution.

3. Sequential Parameter and States Learning

Since their introduction, stochastic volatility models have been an interesting benchmark for many estimation techniques. Some of these rely on the Efficient Method of Moments of Gallant & Tauchen (1996), others on the Implied-State Generalized Method of Moments (IS-GMM) of Pan (2002). Estimation through Maximum Likelihood has been carried out in Aït-Sahalia (2002), by approximating analytically the transition density through Hermite polynomials. Recently, many methods based on simulation techniques have been implemented in order to approximate the likelihood. Simulated maximum likelihood methods have been proposed in Brandt & Santa-Clara (2002), Durham & Gallant (2002) and Koopman & Hol Uspensky (2002) among others. Filtering techniques to evaluate the likelihood have been implemented in Johannes et al. (2002*a*) and in Pitt (2002).

In the recent literature, Monte Carlo algorithms have provided a flexible yet powerful tool for inference on complex models possibly with non observable components. MCMC methods have been introduced in Jacquier et al. (1994) and in Kim et al. (1998). An application to models with jumps has been implemented in Chib et al. (2002) and in Eraker et al. (2003). Furthermore, MCMC methods for inference on continuous time models have been implemented in Eraker (2001) and in Elerian et al. (2001). MCMC methods provide efficient and accurate estimates when applied to off-line applications but seems to be inadequate when dealing with real time applications where we need to update regularly our estimates at each time.

Particle filter algorithms, introduced in Gordon et al. (1993), have been successfully used in a variety of fields such as engineering, econometrics and biology for instance. They provide a sub-optimal but feasible solution to the Bayesian filtering problem. A detailed review on adaptive sequential algorithms is given in Liu & Chen (1998) and in Doucet et al. (2001), whereas an useful tutorial is Arulampalam et al. (2002).

Goal of the methods is to estimate the posterior distribution of the parameters $p(\theta|y_{1:t})$ and the filtering distribution of the latent states $p(v_{t-1}, \mu_{t-1}, J_t, \kappa_t|y_{1:t})$. We first describe the mechanics of these algorithms when the parameters are known. We then extend our solution to the parameter learning problem. Suppose for instance to deal with a general state-space model

$$y_t = h_m(x_t, \epsilon_t) \tag{3.1}$$

$$x_t = h_s(x_{t-1}, \eta_t) \tag{3.2}$$

where (3.1) and (3.2) are respectively the measurement and the state equations. Here x_t is the so called state sequence, y_t is the observed variable, (ϵ_t, η_t) is a white noise process and $h_s(\cdot)$ and $h_m(\cdot)$ are possibly nonlinear functions. Our goal is to estimate the distribution $p(x_{t+1}|y_{1:t+1})$ given the distribution $p(x_t|y_{1:t})$ in which $y_{1:t} = (y_1, \ldots, y_t)$ is the past history of the observable process.

To implement the filter, it is required the knowledge of the initial distribution $p(x_0)$ of the transition distribution $p(x_{t+1}|x_t),\ t \geq 0$ and of the measurement distribution $p(y_t|x_t),\ t \geq 1$. The key idea is to approximate the filtering density $p(x_{t+1}|y_{1:t+1})$ by a discrete cloud of points, $\{x^j_{t+1} : j = 1, \ldots N\}$, called particles as follows

$$\hat{p}(x_{t+1}|y_{1:t+1}) = \sum_{j=1}^{N} \omega^j_{t+1} \delta(x_{t+1} - x^j_{t+1}), \tag{3.3}$$

where ω^j_{t+1} are suitable weights and $\delta(\cdot)$ is an indicator function. Obtaining a sample from (3.3) is easy by recurring to the importance sampling method in which the proposal $q(x_{t+1}|y_{1:t+1})$ can be set to $p(x_{t+1}|x_t)$ with weights $\omega^j_{t+1} \propto \omega^j_t\ p(y_{t+1}|x^j_{t+1})$. At time $t+1$ the new particles are generated from a proposal distribution $q(x_{t+1}|x^i_t, y_t)$ and then reweighted according to the weights ω_{t+1} given by

$$\omega^i_{t+1} \propto \omega^i_t \frac{p(y_{t+1}|x_{t+1})p(x^i_{t+1}|x^i_t)}{q(x^i_{t+1}|x^i_t, y_t)} \qquad i = 1, \ldots N \tag{3.4}$$

It can be proved that the variance of the weights increases systematically over t with the consequence that we associate unit weight to one particle and zero to the others. For this reason a resampling step is added to this simple scheme in order to avoid numerical degeneracies by getting rid of the points with low probability.

An important variant of the basic filter is the auxiliary particle filter suggested in Pitt & Shephard (1999) in which the proposal depends on the past stream of particles through an auxiliary variable $J = 1, \ldots, N$ that index the past trajectories (more details on this method are provided in Liu & Chen 1998 and in Godsill & Clapp 2001). In the original version of the algorithm, the probability ω_{t+1} is corrected by an adjustment multiplier that should diversificate the particles with high probability. In general this factor is taken to be proportional to $p(y_{t+1}|\mu^j_{t+1})$ in which μ^j_{t+1} is a likely value of $p(x_{t+1}|x^j_t)$ such as the mean or the mode for example. In many applications this extension generate particles that are most likely close to the true distribution.

In general Monte Carlo filtering techniques provides a viable and efficient solution to the states filtering problem when the parameters are known. However, inference for parameters of the model is a challenging question. In the sequential Monte Carlo literature, a common practice is to artificially define a dynamic for the parameters, let say θ_t and then include these *dynamics* in an augmented state vector (x_t, θ_t) (See Gordon et al. 1993 and Kitagawa 1998 for example). The main point against this approach is that this procedure leads to time varying and not to fixed parameter estimates.

Recently a number of paper have tackled the problem of estimating fixed parameters in a sequential context. For example Storvik (2002) propose a filter in which the parameters are sequentially updated by simulating their conditional distribution $p(\theta|y_{1:t})$. A different approach named practical filter by Johannes et al. (2006) is based on the idea that the $p(x_t, \theta|y_{1:t})$ can be expressed as a mixture of lag-filtering distributions. The estimate is then based on a rolling-window MCMC algorithm.

In the context of stochastic volatility models, however, these methods seems to provide unstable results for some parameters[2].

A different approach for the filtering problem of a dynamic state space model based on the auxiliary particle filter has been proposed in Liu & West (2001). Their approach generalizes in a dynamic context the kernel smoothing approximation of the posterior $p(\theta|y_{1:t})$ proposed in West (1993), thus leading to

$$p(\theta|y_{1:t}) \simeq \sum_{i=1}^{N} \omega_t^i N(\theta|m_t^j; h^2\Sigma_t). \tag{3.5}$$

The matrix Σ_t is an estimate of the posterior variance covariance matrix, $m_t^i = a\theta^i + (1-a)\bar{\theta}_t$ is the kernel location for the i-th component of the mixture, where $\bar{\theta}_t$ is an estimate of the posterior's mean. Finally the constants h and a, that measure the extent of the shrinkage and the degree of overdispersion of the mixture are given by $h^2 = 1-((2\delta-1)/2\delta)^2$, $a = \sqrt{1-h^2}$, whereas the discount factor δ ranges between 0.95-0.99.

In a dynamic context, this approximation is equivalent to impose a dynamic on the parameters given by

$$\theta_{t+1}|\theta_t \sim \mathrm{N}\left(a\theta_t + (1-a)\bar{\theta}_t, h^2\Sigma_t\right). \tag{3.6}$$

The loss of information induced by this *artificial* evolution is automatically corrected by the shrinkage factors that constraint $\mathrm{Var}(\theta_{t+1}|y_{1:t})$ to be equal to $\mathrm{Var}(\theta_t|y_{1:t})$. Once the model has been completed the filtering is performed through the auxiliary particle filter of Pitt & Shephard (1999).

This methodology has been successfully used in Liu & West (2001) in a dynamic factor stochastic volatility context and in Carvalho & Lopes (2006) for a switching regime stochastic volatility framework. For a stochastic volatility model with jumps we found that the basic setup described above perform poorly in practice, providing unstable estimates of the posterior mean of θ over time. A second problem we noticed is that the estimated posterior variance-covariance matrix Σ_t collapses to zero after few hundreds iterations.

This latter problem is probably due to the sample impoverishment phenomenon caused by the resampling step. In fact, particles with high probability are selected many times causing a loss of diversity in the cloud. This problem is severe when the noise of the latent process is small[3]. A possible remedy is

[2]For this reason, (Johannes et al. 2006) need to fix the parameter σ_η to be constant for both the methods

[3]Note that this feature can be emphasized by the choice of the shrinking factor $h^2 < 1$ defined above.

to choose an efficient resampling scheme that keeps low the Monte Carlo variance. We found the *residual sampling* proposed in Liu & Chen (1995) an useful alternative.

`Residual Sampling`

- Retain $k_j = \lfloor N\omega_t^j \rfloor$ copies of x_{t+1};
- Sample the remaining $N - \sum_{i=1}^{N} k_i$ with probability proportional to $N\omega_t^j - \lfloor N\omega_t^j \rfloor$;
- Reset the weights to $1/N$.

where $\lfloor z \rfloor$ is the greater integer less than z.

Another approach to increase the sample variability is to recur to MCMC moves. This wariness should also help to reduce the correlation between particles. This idea has been recently developed in Gilks & Berzuini (2001) and Berzuini & Gilks (2001). It is worth noting that the outcome of the auxiliary particle filter is a weighted sample approximating the posterior distribution $p(\theta, x_{t+1}|y_{1:t+1})$. All these particles can be *rejuvenated* or moved according to a Markov transition with the same posterior as invariant distribution. For this reason it is not necessary a burn-in time for the MCMC step.
We apply this idea to the parameter learning methodology proposed in Liu & West (2001). More precisely we update in this way the parameters and the jump process.

We now provide the details of the algorithm considering the version we implement for the model described in (2.2)-(2.3). Using the notation introduced in Johannes et al. (2002*a*), the vector of the states is $x_{t+1} = (v_t, \mu_t, J_{t+1}, \kappa_{t+1})$. Furthermore, to simplify the notation, θ_t is the posterior conditional on the past up to time t, i.e. $\theta_t \sim p(\theta|y_{1:t})$. We estimate the posterior distribution $p(v_t, \mu_t, J_{t+1}, \kappa_{t+1}\theta|y_{1:t+1})$.

The resulting algorithm is summarized as follows

`Parameter learning algorithm`

0. Simulate N particles from the prior $p(\theta_0)$, from $p(v_0)$ and from $p(\mu_0)$, $J_t = 0$ and $\kappa_t = 0$ with equal weights;
   For $t = 1$ to T:
1. Given $x_t^j = (v_{t-1}^j, \mu_{t-1}, J_t, \kappa_t, \theta_t^j)$ and $\omega_t^j,\ j = 1, \ldots, N$, compute
$$\begin{aligned} \bar{\nu}_t^j &= E[v_t|v_{t-1}^j, \theta_t^j] \\ \bar{\mu}_t^j &= E[\mu_t|\mu_{t-1}^j, \theta_t^j] \\ m_t^j &= a\theta_t^j + (1-a)\bar{\theta}_t \\ \bar{J}_{t+1}^j &= 0 \end{aligned}$$
2. Draw an integer τ from $\tau \in \{1, \ldots, N\}$ using the residual sampling with probabilities
$$g_{t+1}^j \propto \omega_t^j p(y_{t+1}|\bar{\nu}_t^j, \bar{\mu}_t^j, \bar{J}_{t+1}, m_t^j)$$
3. Update θ_{t+1} from $\theta_{t+1}^\tau \sim N(m_t^\tau, h^2\Sigma_t)$

4. Update v_t from $p(v_t|v_{t-1}^{\tau}, \theta_{t+1}^{\tau})$
5. Update μ_t from $p(\mu_t|\mu_{t-1}^{\tau}, \theta_{t+1}^{\tau})$
6. Update J_{t+1} from $p(J_{t+1}|\theta_{t+1}^{\tau})$
7. Update κ_{t+1} from $p(\kappa_{t+1}|\theta_{t+1}^{\tau})$
8. Update the sufficient statistics according to the draws in step 3 to 7.
9. Compute $\omega_{t+1}^{\tau} \propto \frac{p(y_{t+1}|v_t^{\tau}, \mu_t^{\tau} \theta_{t+1}^{\tau})}{p(y_{t+1}|\bar{\mu}_t^{\tau}, \bar{\nu}_t^{\tau}, m_t^{\tau})}$
10. Repeat step (2)-(9) N times. Record $x_{t+1}^j = (v_t^j, \mu_t^j, J_{t+1}^j, \kappa_{t+1}^j, \theta_{t+1}^j)$.
11. (Optional) Move the former particles according to MCMC with invariant distribution the posterior.and update the sufficient statistics according to the former MCMC move.

We perform the MCMC step for the parameters every 50 iteration of the algorithm, whereas we update J_{t+1} and κ_{t+1} systematically. This choice provides a reasonable compromise between statistical precision and computational burden.
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 these information is through a set of sufficient statistics S_t (Fearnhead 2002). The amount of computer memory requirements is sensibly reduced. It is also convenient to use some transformation of the parameters θ in order to extend their support to the real line. In fact the posterior is approxiamted 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 the *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 the three variances and of the intensity λ.

4. Empirical Results

In this section we test our algorithm using simulated data and Standard's & Poor 500 index observed on a daily basis during the last 18 years from January 1985 to July 2003.

4.1. Simulated Data

We first run our methodology on simulated data. We consider the complete model with leverage and conditional mean and variances.

To analyze the algorithm performance, we simulated 2000 observations from the model in eq. (2.2)-(2.3) in which the parameters are the following

$$\text{Volatility process: } \mu = 0.06, \quad \phi = 0.95, \quad \sigma_\eta = 0.15, \quad \rho = -0.5;$$

$$\text{Conditional mean: } \alpha = 0.001, \quad \beta = 0.90, \quad \sigma_\mu = 0.1;$$

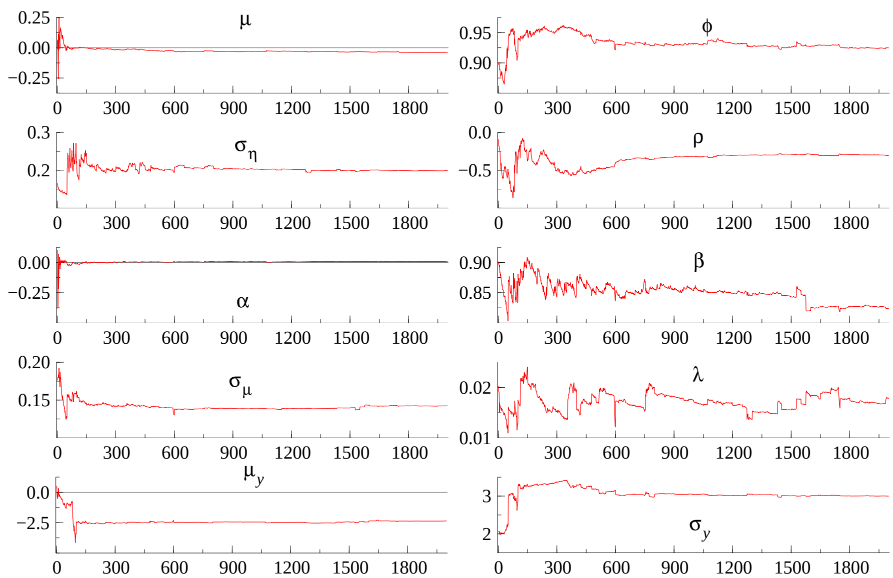

Fig 1. *Estimated Parameters*

$$\text{Jump Process: } \lambda = 0.01, \quad \mu_y = -4, \quad \sigma_y = 2.$$

These parameter value are consistent with some empirical findings on similar jump diffusion models (see for example Raggi 2005). For this experiment, we use 25000 particles to approximate the distributions of interest. The results are reported on Figure 1.

In general we notice that the algorithm provides stable estimates for the parameters and the estimates are consistent with the true parameters.

In some occasions difficulties arise for the parameters related to the jump process, in which case some care has to be taken in the empirical analysis. The reason of these occasional pitfalls probably depends on the nature of the jumps, that are rare events. Thus it is difficult to identify the parameters that describe the jumps, i.e. μ_y and σ_y, through the data. However, even if jumps are rare events, the algorithm detect properly the jump events and their sizes. This evidence is stressed in Figure 2.

A second problem related to the parameters's jump process might arise with simulated data. In fact in some case we have noticed an occasional inability of the algorithm to distinguish efficiently between outliers and actual jumps. In practice, with this data set, we slightly over-estimate the parameter λ that is the probability to detect jumps *and* outliers. However, Figure 2 shows that almost all the jumps are detected properly. This effect on λ is negligible with real data applications, in which the estimates of the parameters are really close

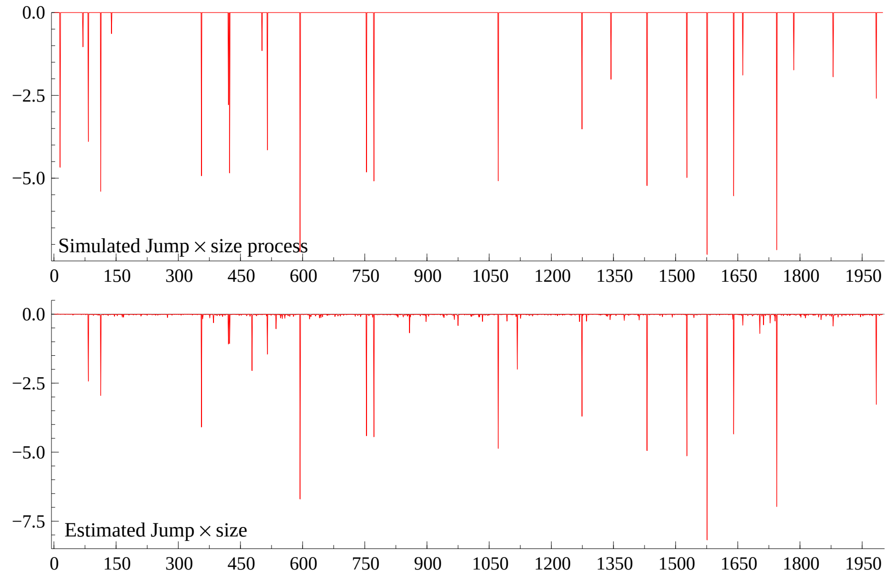

FIG 2. *Estimated Parameters*

to the one obtained with similar models through MCMC (see Raggi 2005 for example). Troubles related to the lack of identification of jump models has also been noticed in Chib et al. (2002) and Eraker et al. (2003).

We also obtain a reasonable estimate for the parameters μ_y and σ_y.

The log-volatility process is estimated as reported in figure 3.

The algorithm provides quite precise estimates for the leverage ρ and also a good estimate of the persistence ϕ. The only parameter that is slightly over-estimated is σ_η, but its magnitude is not far from the true value.

The conditional mean is reported in Figure 4

4.2. S&P 500 Index

In this section we report some empirical results based on the analysis of the S&P 500 index observed from January 1985 to July 2003. The data set has been downloaded from `Datastream`. All the calculations made in this paper are based on software written using the Ox©3.2 language of Doornik (2001). By letting Y_t be the observed log-prices, the returns are defined as $100 \times (Y_t - Y_{t-1})$.

We estimate our model approximating the distribution of interest through a cloud of 50,000 particles. The results are summarized in figures from 5 to 8.

Figure 5 provides the plot of the original time series together with the estimates of the latent processes. It is remarkable to note that to each spike in the returns corresponds an estimated jump. This phenomenon is evident for the

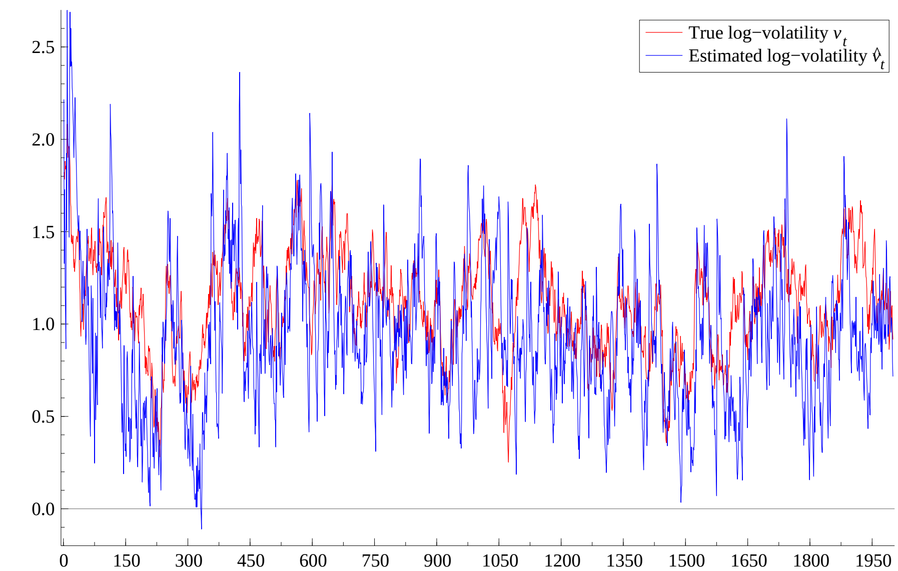

FIG 3. *Estimated Volatilities*

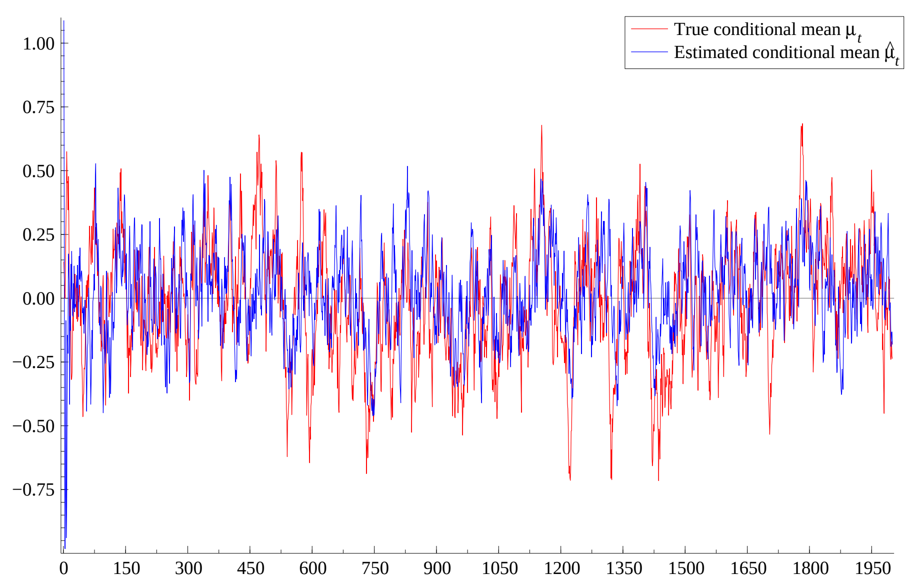

FIG 4. *Estimated Conditional Means*

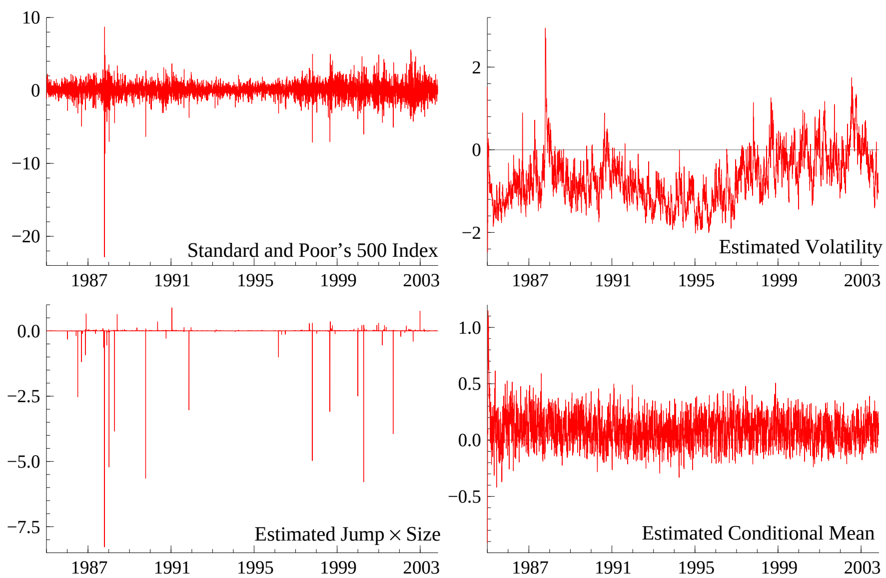

FIG 5. *Estimated latent Processes*

crash observed in October 1987 but is also clear for many others extreme days. It is also clear that the log-volatility burst when an extreme event happens. This finding, that is consistent to the current literature on jump diffusion models, is a reasonable feature since to an extreme event corresponds a huge increase on risks or variability on the returns.

In Figure 6 we provide the estimates for the parameters of the volatility process. The sequential estimates for ρ confirm a leverage effect between volatility and returns and the estimate ($\rho \approx -0.33$) is negative and sensibly different to zero. The log-volatility process is persistent. We found that ϕ tends to increase along t, but this behaviour can be explained by the increase in the volatilities observed in the last four years.

The conditional mean provide some evidence on the predictability of the returns. The intercept α is close to zero but the persistency parameter converge to 0.76, thus indicating that the process is not a white noise. This evidence stresses that the arrival of a jump has a persistent effect on the returns. It is important to notice this feature, since many models consider independent jumps that have a just a transient impact on returns.

Finally the parameter estimates of the jump and sizes processes are plotted in Figure 8. λ evidences that the model detects about three extreme events per year. However, this estimate is about 6 time bigger during the 1987 crisis. During that period, in fact, there have been a number of smaller jumps closed to the main one dated 19th of October. For the expected size ($\mu_y = 0.33$) and the variance ($\sigma_y = 3.63$) we notice some parameter uncertainty. This is likely due

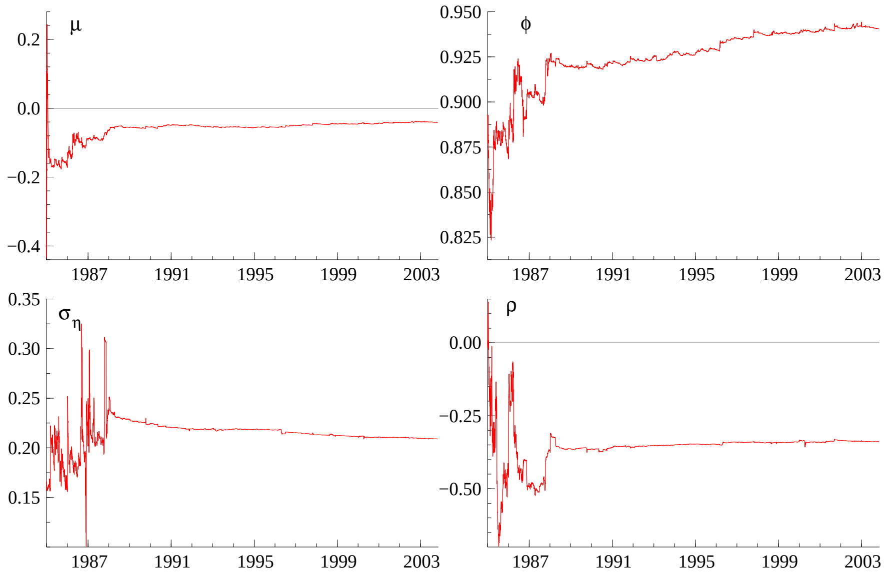

FIG 6. *Estimated Parameters of the volatility dynamic* $\mu, \phi, \sigma_\eta, \rho$

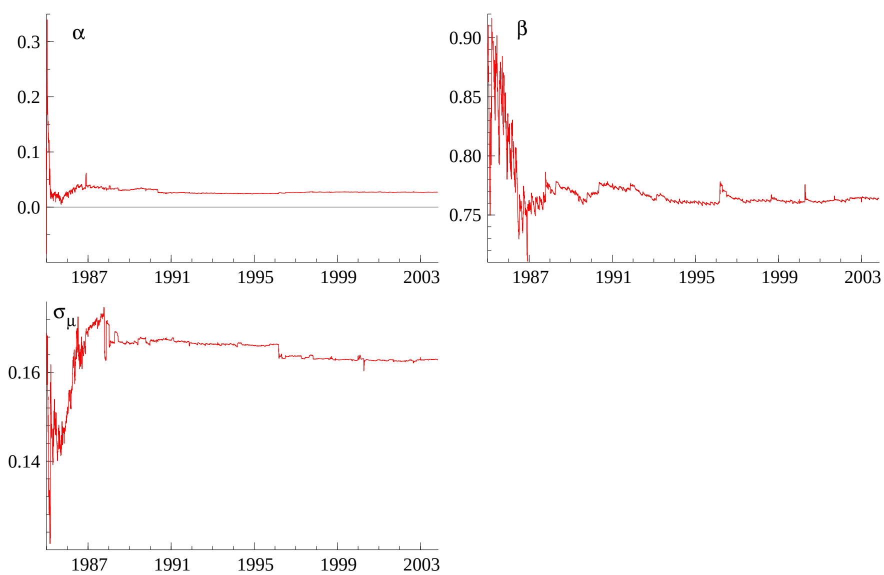

FIG 7. *Estimated Parameters of the conditional mean dynamic* $\alpha, \beta, \sigma_\zeta$

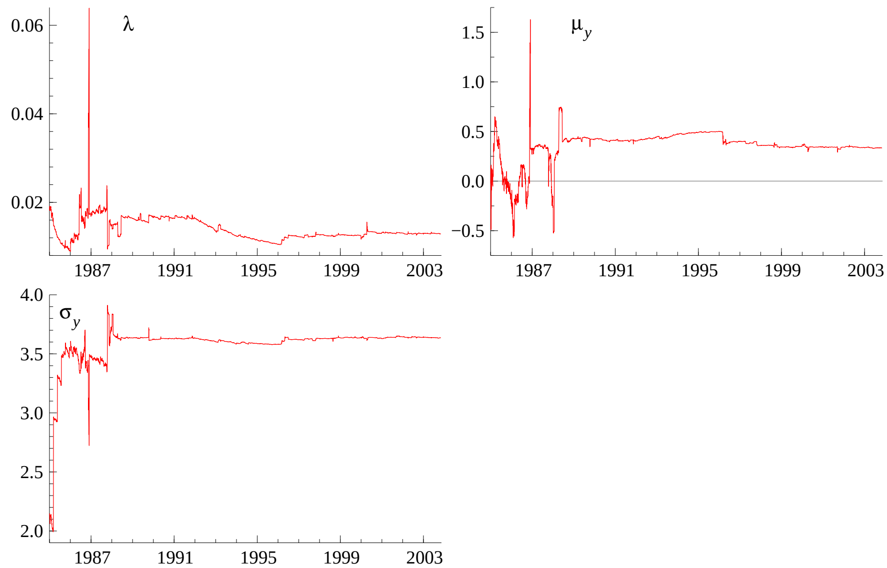

Fig 8. *Estimated Parameters of the jump dynamic* λ, μ_y, σ_y

to the rare nature of the jumps, that can cause some problem of identification of these parameters.

5. Conclusions and Further Developments

Monte Carlo sequential methods represents a valuable and a reliable methodology to estimate non linear and non gaussian state-space models. They also seem to be particularly useful to estimate stochastic volatility models and their extensions. In this paper we propose an algorithm based on the kernel smoothing approximation of the posterior suggested in Liu & West (2001) in which we incorporate a MCMC step in order to reduce sampling impoverishment problems related to sequential Monte Carlo strategies. Furthermore, we noticed in our empirical applications that this algorithm provides consistent and stable results also with longer time series that are typical in financial econometrics.

A first possible extension is to generalize the current model to allow for a time dependent intensity's dynamic. It is in fact evident from the estimates on the S&P500 series that extreme events arrive in clusters. For example there are many jumps between 1986 and 1991, really few and with negligible size between 1992 and 1996 and again many jumps in the last years. There is also a strong evidence that the number of jumps with high size is directly proportional to the volatility.

An interesting economic issue to explore is to quantify how an extreme event like a jump has an impact on the optimal portfolio weights. In an affine jump diffusive framework (see Duffie et al. 2000 for a theoretical treatment for these models), Liu et al. (2003) prove that these optimal weights can be computed through the solution of an ordinary differential equation. We think it would be interesting to estimate sequentially these quantities prior to a crash and after this event taking into account the parameters and states uncertainty related to the inferential procedure.

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.

Bakshi, G., Cao, C. & Chen, Z. (1997), 'Empirical Performance of Alternative Option Pricing Models', *Journal of Finance* **52**, 2003–2049.

Berzuini, C. & Gilks, W. (2001), RESAMPLE-MOVE Filtering with Cross Model Jump, *in* A. Doucet, N. de Freitas & N. Gordon, eds, 'Sequential Monte Carlo Methods in Practice', Springer, New York, pp. 117–138.

Brandt, M. & Santa-Clara, P. (2002), 'Simulated Likelihood Estimation of Diffusions with an Applicaiton to Exchange Rate Dynamics in Incomplete Markets', *Journal of Financial Economics* **63**, 161–210.

Carvalho, C. & Lopes, H. (2006), 'Simulation-Based Sequential Anlysis of Markov Switching Stochastic Volatility Models', *Computational Statistics & Data Analysis* **(forthcoming)**.

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.

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

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

Duffie, D., Pan, J. & Singleton, K. (2000), 'Trasform Analysis and Asset Pricing for Affine Jump-Diffusions', *Econometrica* **68**, 1343–1376.

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.

Eraker, B. (2001), 'MCMC Analysis of Diffusion Processes with Application to Finance', *Journal of Business & Economic Statistics* **19**, 177–191.

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* C. Rao & G. Maddala, eds, 'Statistical Methods in Finance', North Holland, Amsterdam, 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.

Godsill, S. & Clapp, T. (2001), Improvement Strategies for Monte Carlo Particle Filters, *in* A. Doucet, N. de Freitas & N. Gordon, eds, 'Sequential Monte Carlo Methods in Practice', Springer, New York, pp. 139–158.

Gordon, N., Salmond, D. & Smith, A. (1993), 'Novel Approach to Non-linear and Non-Gaussian Bayesian State Estimation', *IEE Proceedings-F* **140**, 107–

113.

Heston, S. (1993), 'A Closed-Form Solution of Options with Stochastic Volatility with Applications to Bond and Currency Options', *Review of Financial Studies* **6**, 327–343.

Hull, J. & White, A. (1987), 'The Pricing of Options on Assets with Stochastic Volatilities', *Journal of Finance* **42**, 281–300.

Jacquier, E., Polson, N. & Rossi, P. (1994), 'Bayesian Analysis of Stochastic Volatility Models (with discussion)', *Journal of Business & Economic Statistics* **12**, 371–417.

Johannes, M., Polson, N. & Stroud, J. (2002*a*), Nonlinear Filtering of Stochastic Differential Equations with Jumps, Working paper, Graduate School of Business, University of Chicago.

Johannes, M., Polson, N. & Stroud, J. (2002*b*), 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.

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.

Koopman, S. & Hol Uspensky, E. (2002), 'The Stochastic Volatility in Mean model: Empirical evidence from international stock markets', *Journal of Applied Econometrics* **17**, 667–689.

Liu, J. & Chen, R. (1995), 'Blind Deconvolution via Sequential Imputations', *Journal of the American Statistical Associacion* **90**, 567–576.

Liu, J. & Chen, R. (1998), 'Sequential Monte Carlo Methods for Dynamic Systems', *Journal of the American Statistical Associacion* **93**, 1032–1998.

Liu, J., Longstaff, F. & Pan, J. (2003), 'Dynamic Asset Allocation with Event

Risk', *Journal of Finance* **63**, 231–259.

Liu, J. & West, M. (2001), Combined Parameter and State Estimation in Simulation-Based Filtering, *in* A. Doucet, N. de Freitas & N. Gordon, eds, 'Sequential Monte Carlo Methods in Practice', Springer-Verlag, New York, pp. 197–223.

Merton, R. C. (1971), 'Optimum Consumption and Portfolio Rule in Continuous Time', *Journal of Economic Theory* **3**, 373–413.

Nelson, D. (1991), 'Conditional Heteroskedasticity in Asset Pricing: a New Approach', *Econometrica* **59**, 347–370.

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.

Raggi, D. (2005), 'Adaptive MCMC Methods for Inference on Affine Stochastic Volatility Models with Jumps', *Econometrics Journal* **8**, 235–250.

Storvik, G. (2002), 'Particle Filters for State-Space Models with the Presence of Unknown Static Parameters', *IEEE Transactions on Signal Processing* **50**, 281–289.

Taylor, S. (1994), 'Modelling Stochastic Volatility: a Review and Comparative Study', *Mathematical Finance* **4**, 183–204.

West, M. (1993), 'Approximating Posterior Distribution by Mixture', *Journal of the Royal Statistical Society, Ser. B* **55**, 409422.

Yu, J. (2005), 'On Leverage in a Stochastic Volatility Model', *Journal of Econometrics* **127**, 165–178.