Modelling Stochastic Volatility with Leverage and Jumps: A Simulated Maximum Likelihood Approach via Particle Filtering.

Sheheryar Malik and Michael K Pitt

No 897

WARWICK ECONOMIC RESEARCH PAPERS

DEPARTMENT OF ECONOMICS

THE UNIVERSITY OF WARWICK

Modelling Stochastic Volatility with Leverage and Jumps: A Simulated Maximum Likelihood Approach via Particle Filtering.

Sheheryar Malik and Michael K Pitt

Department of Economics, University of Warwick, Coventry CV4 7AL

April 3, 2009

Abstract

In this paper we provide a unified methodology in order to conduct likelihood-based inference on the unknown parameters of a general class of discrete-time stochastic volatility models, characterized by both a leverage effect and jumps in returns. Given the non-linear/non-Gaussian state-space form, approximating the likelihood for the parameters is conducted with output generated by the particle filter. Methods are employed to ensure that the approximating likelihood is continuous as a function of the unknown parameters thus enabling the use of Newton-Raphson type maximization algorithms. Our approach is robust and efficient relative to alternative Markov Chain Monte Carlo schemes employed in such contexts. In addition it provides a feasible basis for undertaking the non-trivial task of model comparison. The technique is applied to daily returns data for various stock price indices. We find strong evidence in favour of a leverage effect in all cases. Jumps are an important component in two out of the four series we consider.

Some key words: Particle filter, Simulation, SIR, State space, Leverage effect, Jumps.

1 Introduction

The aim of this paper is to conduct likelihood-based inference on a general class of stochastic volatility models using a smooth particle filter. Stochastic volatility (SV) models have gained considerable interest in theoretical options pricing and financial econometrics literature; in the latter as an alternative to the well documented ARCH/GARCH-type models. The SV framework allows variance to evolve according to some latent stochastic process.

In studying the relationship between volatility and asset price/return, a so-called "leverage effect" refers to the increase in future expected volatility following bad news. The reasoning underlying is that, bad news tends to decrease price thus leading to an increase in debt-to-equity ratio (i.e. financial leverage). The firms are hence riskier and this translates into an increase in expected future volatility as captured by a negative relationship between volatility and price/return. In the finance literature empirical evidence supportive of a leverage effect has been provided by Black (1976) and Christie (1982). The state-space form of SV model that is studied in the bulk of the literature assumes that the measurement and state equation disturbances are uncorrelated, thus ruling out leverage.

Another characteristic of financial data are "jumps" in the returns process. Jumps can basically be described as rare events; large, infrequent movement is returns which are an important feature of financial markets (see Merton (1976)). These have been documented to be important in characterizing the non-Gaussian tail-behaviour of conditional distributions of returns.

The case of SV with leverage has recently been considered by Christofferesen, Jacobs and Minouni (2007). They analyse various specifications of the stochastic volatility model with leverage, e.g. the affine SQR model of Heston (1993) and also various non-affine models. They demonstrate the generality and robustness of the smooth particle filter for purposes of parameter estimation (See Pitt (2003)). We add to the literature by providing a very general methodology for carrying out maximum likelihood estimation of the parameters of an SV model which incorporates both leverage and jumps, within a particle filtering framework.

The plan of this paper is as follow. In Section 2 we describe the standard SV model, the SV with leverage model and the SV with leverage with jumps model along with a brief review of the literature. In Section 3 we first describe how parameter estimation can be carried out using particle filters generally, and then specifically in the context of the SV with leverage and jumps model. We also describe the relevant diagnostic tests. Section 4 provides results for simulation experiments testing estimator performance. Section 5 provides empirical examples using daily returns data for S&P500, FTSE 100, Dow Jones and Nasdaq. Section 6 concludes.

2 Volatility Models

2·1 *Stochastic Volatility*

The standard stochastic volatility (SV) model with uncorrelated measurement and state equation disturbances is given by,

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t, \quad t = 1, ..., T \end{aligned} \tag{2·1}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \backsim N(0, \Sigma) \text{ and } \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Here y_t is the observed return, $\{h_t\}$ are the unobserved log-volatilities, μ is the drift in the state equation, σ_η^2 is the volatility of log-volatility and ϕ is the persistence parameter. Within the econometrics literature, this model is seen as a generalization Black-Scholes option pricing formula that allows for volatility clustering in returns. There have been different methodologies proposed in the context of parameter estimation for such models. Harvey, Ruiz and Shephard (1994), advocates a Quasi Maximum Likelihood procedures, whereas Jacquier, Polson and Rossi (1994) propose an MCMC method in order to construct a Markov chain that can be used to draw directly from the posterior distributions of the model parameters and unobserved volatilities (see also Shephard and Pitt (1997)). Shephard and Pitt (1997) and Durbin and Koopman (1997) consider importance sampling in order to obtain the likelihood.

2·2 *Stochastic Volatility with Leverage*

We can take the standard SV model just described and adapt it in order to incorporate a leverage effect. Given that,

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t, \quad t = 1, ..., T \end{aligned} \tag{2·2}$$

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \backsim N(0, \Sigma),$$

we now allow for the disturbances to be correlated (see e.g. Heston (1993)) [1] which implies

[1] Primary contributions in modelling leverage within an ARCH/GARCH framework have been made by Nelson(1991), Glosten, Jagannathan and Runkle (1994) and Engle and Ng (1993). Asymmetric models put forth in this regard, such as TARCH and EGARCH make conditional variance a function of the sign in addition to the size of returns.

that the covariance matrix has the form,

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \tag{2·3}$$

Furthermore, noting that the disturbances are conditionally Gaussian, we can write $\eta_t = \rho\ \epsilon_t + \sqrt{(1-\rho^2)}\xi_t$, where $\xi_t \backsim N(0,1)$.The state equation can be reformulated as,

$$h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \rho \epsilon_t + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t. \tag{2·4}$$

By substituting, $\epsilon_t = y_t \exp(-h_t/2)$ into (2.4), the model adopts the following Gaussian nonlinear state-space form where the parameter ρ measures the leverage effect.

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \rho\ y_t \exp(-h_t/2) + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t. \end{aligned} \tag{2·5}$$

Alternatively we could have written $\epsilon_t = \rho\ \eta_t + \sqrt{(1-\rho^2)}\zeta_t$, where ζ_t is again an independent standard Gaussian. In which case, the SV with leverage model is given by, $y_t|\eta_t \sim N(\rho \exp(h_t/2)\eta_t\ ; (1-\rho^2)\exp(h_t))$ where $h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t$.

Amongst the earliest contributions in modelling leverage in the stochastic volatility literature was made by Harvey and Shephard (1996). The authors extend the Quasi Maximum Likelihood (QML) technique used in parameter estimation in standard SV models (see Harvey et al. (1994)) to handle correlation between disturbances. Recognizing that information on correlation is lost as result of squaring the observations in the process of linearizing the model; the technique developed by Harvey and Shephard (1996) allows the information to be recovered by carrying out inference conditional on the signs of observations, i.e. by relating these to filtered volatilities. When applied to daily CRSP (Centre for Research in Security Prices) and SP30 (Standard and Poors), the authors find evidence of a leverage effect. A problem with the QML approach is that log ϵ_t^2 is poorly approximated by the normal distribution yielding a quasi-likelihood estimator with poor finite sample properties.

In order to correct for this Kim, Shephard and Chib (1998) develop an alternative approach for analysis of SV models employing MCMC techniques to provide a likelihood-based framework. The Kim et al. approach revolves around approximating $\log \epsilon_t^2$ by a mixture of seven normal densities which in turn facilitates the state-space representation associated with the Kalman Filter. Omori, Chib, Shephard and Nakajima (2007) extend this approach to handle leverage in SV models. They apply this approach to fit a model to daily returns of TOPIX and find evidence of leverage.

Jacquier, Polson and Rossi (2004) build upon the MCMC approach put forth in JPR (1994) to conduct inference in an extended SV model, i.e. to allow for both leverage effect but also fat-tails in the measurement equation disturbances, where evidence supportive of the latter has been uncovered by Gallant et al. (1998) and Gweke (1992), amongst others[2]. Application of their model to weekly CRSP, daily S&P 500 data as well as a few daily exchange rate series' yields evidence supportive of the extensions.

Meyer and Yu (2000) also employ a Gibbs sampling approach to perform posterior computations on an asymmetric SV model and find evidence of a leverage effect in daily Pound/Dollar exchange rate series. Yu (2005) documents the two main specifications for modelling leverage in the literature, and notes an important difference between the two which becomes apparent

[2]Jacquier et al assume $\epsilon_t = \sqrt{\lambda_t} z_t$ where z_t is a standard normal variate and λ_t is distributed as i.i.d. inverse gamma, whereby the marginal distribution is student-t.

The fat-tailed extension is also explored in Harvey, Ruiz and Shephard (1994) and Kim et al (1998).

when the two specifications are written in a Gaussian nonlinear state-space form. Whereas, Kim et al. (1998) and Omori et al. (2007) work with the Euler-Maruyuma approximation for the continuous time asymmetric SV model. Yu notes that the timing of the variables makes it difficult to interpret the leverage effect in the Jacquier et al.(2004) specification given we can not obtain the relationship between $E(h_{t+1}|y_t)$ and y_t in analytical form. For further discussion, we refer the reader to Yu (2005, pg 6). He concludes that from an empirical stand point having tested both specifications on daily S&P 500 and CRSP data, that the specification of the basic model as used in Kim, Shephard and Chib (1998) is preferred.

2·3 *Stochastic Volatility with Leverage and Jumps*

The SV model with leverage which allows for jumps in the returns process can be written as,

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) + J_t \varpi_t \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t, \quad t = 1, ..., T \end{aligned} \tag{2·6}$$

where,

$$\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} \backsim N(0, \Sigma) \text{ and } \Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

$J_t = j$ is the time-t jump arrival where $j = 0, 1$ is a Bernoulli counter with intensity p. $\varpi_t \backsim N(0, \sigma_J^2)$ dictates the jump size. The leverage effect is incorporated as before noting $f(\eta_t|\epsilon_t) = N(\rho\epsilon_t; 1-\rho^2)$. This model can be considered a discrete-time counterpart to a general, continuous-time jump-diffusion model (see Duffie, Pan and Singleton (2000) and Johannas, Polson and Stroud (2009)). In brief, assume log of stock price $y(t)$ and the underlying state variable, i.e. the volatility $X(t)$ jointly solve:

$$\begin{aligned} \mathrm{d}y(t) &= a^y(X(t))\mathrm{d}t + \sigma^y(X(t))\mathrm{d}B(t) + \mathrm{d}\left(\sum_{n=1}^{N_t^y} Z_n^y\right), \\ \mathrm{d}X(t) &= g^x(X(t))\mathrm{d}t + \sigma^x(X(t))\mathrm{d}W(t) + \mathrm{d}\left(\sum_{n=1}^{N_t^x} Z_n^x\right). \end{aligned}$$

Here $B(t)$ and $W(t)$ are correlated Brownian motions, N_t^y and N_t^X are homogenous (or non-homogenous) Poisson processes with Z_n^y and Z_n^x being the jump sizes for stock returns and volatility respectively. The functions $a^y(.), \sigma^y(.), g^x(.)$ and $\sigma^x(.)$ are general functions subject to certain constraints.

There have been a several recent contributions in estimating SV models with jumps, albeit mostly within a Bayesian framework. Amongst the earliest are Bates (1996) and Bakshi, Cao and Chen (1997), which deal with models involving jumps in returns and parameter estimation carried out via a non-linear generalized least squares/Kalman filtration methodology. This is extended in Bates(2000) which employs a the same estimation methodology for two-factor SV models with jumps in returns. Eraker, Johannes and Polson (2003) provide an MCMC strategy for conducting inference on stochastic volatility models incorporating jumps in returns and also in the volatility process (initially introduced by Duffie et al.(2000)). They conduct empirical analysis on S&P500 and Nasdaq 100 index returns and find strong evidence of jumps in volatility. Jumps have been documented to be important in characterizing the non-Gaussian tail-behaviour of conditional returns distributions. In order to characterize this feature of returns, the approach of estimating SV models with student-t errors have been employed by, for example, Chib, Nedari and Shephard (2006) and Sandmann and Koopman (1998). For the same purposes, an alternative approach employed by Durham (2008) is to use a mixture of Gaussians for the measurement equation disturbance. His paper uses a simulated maximum likelihood approach to conduct inference.

3 Particle Filter Estimation

This paper is concerned with evaluation of state-space models via particle filter. We model the time series $\{y_t, t = 1, ..., T\}$ using a state space framework with the state $\{h_t\}$ assumed to be Markovian. The problem of state estimation within a filtering context can be formulated as the evaluation of the density $f(h_t|Y_{t,})$, $t = 1, ..., T$ where $Y_t = (y_1, ..., y_t)$ is contemporaneously available information. In linear Gaussian state space models the density is Gaussian at every iteration of the filter and the Kalman filter relations propagate and update the mean and covariance of the distribution. In nonlinear and/or non-Gaussian state space models we cannot obtain a closed form expression for the required conditional density and particle filters are employed in order to recursively generate (an approximation to) the state density.

There is has been considerable work done on the development of simulation based methods to perform filtering nonlinear Gaussian state space models. Leading contributions to the literature are by Gordon, Salmond and Smith (1993), Kitagawa (1996), Isard and Blake (1996), Muller (1991) and Pitt and Shephard (1999). A review is provided by Doucet et al. (2000). Most of the literature revolves around on-line filtering of the states with very little work done in the parameter estimation within this framework; see Liu and West (2000), Pitt (2003) and Polson, Stroud and Muller (2008).

We begin by providing a description of a particle filter, as put forth in the seminal paper by Gordon et al. (1993) and then describe how this framework can be adapted for parameter estimation.

3·1 *Preliminaries*

We assume a known measurement density $f(y_t|h_t)$ and the ability to simulate from the transition density $f(h_{t+1}|h_t)$. Particle filters involve using simulation to carry out on-line filtering, i.e. to learn about the state given contemporaneously available information. Suppose we have a set of random samples, 'particles', $h_t^1, ..., h_t^M$ with associated discrete probability masses $\lambda_t^1, ..., \lambda_t^M$, drawn from the density $f(h_t|Y_t)$. The principle of Bayesian updating implies that the density of the state conditional on all available information can be constructed by combining a prior with a likelihood; recursive implementation of which forms the basis for particle filtering. The particle filter is hence an algorithm to propagate and update these particles in order to obtain a sample which is approximately distributed as $f(h_{t+1}|Y_{t+1})$; the true filtering density,

$$f(h_{t+1}|Y_{t+1}) \propto f(y_{t+1}|h_{t+1}) \int f(h_{t+1}|h_t) dF(h_t|Y_t). \qquad (3{\cdot}1)$$

In order to sample from this density we use the Sampling Importance Resampling algorithm of Gordon, Salmond and Smith (1993) (hence forth referred to as SIR). The basic SIR algorithm is outlined below. We start at $t = 0$ with samples from $h_0^i \sim f(h_0)$, $i = 1, ..., M$, which is generally the stationary distribution, if it exists.

Algorithm : SIR *for t=0,..,T-1*:
We have samples $h_t^i \sim f(h_t|Y_t)$ *for* $i = 1, ..., M$.

1. *For* $i = 1 : M$, *sample* $\widetilde{h}_{t+1}^i \sim f(h_{t+1}|h_t^i)$.

2. *For* $i = 1 : M$ calculate normalized weights,

$$\lambda_{t+1}^{i} = \frac{\omega_{t+1}^{i}}{\sum_{k=1}^{M} \omega_{t+1}^{k}}, \text{ where } \omega_{t+1}^{i} = f(y_{t+1}|\widetilde{h}_{t+1}^{i}) \ .$$

3. *For* $i = 1 : M$, *sample (from the mixture)* $h_{t+1}^{i} \sim \sum_{k=1}^{M} \lambda_{t+1}^{k} \delta(h_{t+1} - \widetilde{h}_{t+1}^{k})$.

This will yield an approximation of the desired posterior density, $f(h_{t+1}|Y_{t+1})$ as t varies. Here $\delta(.)$ is a dirac-delta function. Sampling in **Step 3** is a multinomial sampling scheme (sometimes referred to as the weighted bootstrap) and is computationally $O(M)$. It relies on the following result of Smith and Gelfand (1993).

Theorem 3.1 *Suppose that our required density is proportional to* $L(x)G(x)$*, for example, and that we have samples* $x^i \backsim G(x), i = 1, ..., M$*. If* $L(x)$ *is a known function then, the theorem states that the discrete distribution over* x^i *with probability mass* $L(x^i)/\Sigma L(x^i)$ *on* x^i *tends in distribution to the required density as* $M \to \infty$.

The algorithm is iterated through the data to in order to produce empirical filtering densities,

$$\widehat{f}(h_{t+1}|Y_{t+1}) \propto f(y_{t+1}|h_{t+1}) \sum_{i=1}^{M} \lambda_t^i f(h_{t+1}|h_t^i), \tag{3·2}$$

for each time step. It is worth noting that we need to know $f(y_{t+1}|h_{t+1})$ only up to a proportionality. Furthermore, we can estimate all moments, for example $E[h_{t+1}|Y_{t+1}]$ by either $\frac{1}{M}\sum_{i=1}^{M} h_{t+1}^{i}$ using **Step 3** or Rao-Blackwellisation, $\sum_{i=1}^{M} \widetilde{h}_{t+1}^{i}.\lambda_{t+1}^{i}$ using **Step 1 and Step 2**. Next we look at how this SIR particle filter framework can be exploiting and modified in order to carry out likelihood evaluation for parameter estimation.

3·2 *Likelihood Evaluation*

We now assume the model is indexed, possibly in both state and measurement equations, by a vector of fixed parameters θ. In order to carry out parameter estimation we need to estimate the likelihood function, which in log terms is given by;

$$\log L(\theta) = \log f(y_{1,....,}y_T|\theta)$$

$$= \sum_{t=1}^{T} \log f(y_{t+1}|\theta; Y_t),$$

via the prediction decomposition (e.g. see Harvey(1993)). In order to estimate this function, we exploit the relationship,

$$f(y_{t+1}|\theta; Y_t) = \int f(y_{t+1}|h_{t+1}; \theta) f(h_{t+1}|Y_t; \theta) dh_{t+1}. \tag{3·3}$$

The particle filter delivers samples from $f(h_t|Y_t; \theta)$, and we can sample from the transition density $f(h_{t+1}|h_t; \theta)$ in order to estimate the integral. The resampling step is crucial and there are various methods for implementing **Step 3** of the SIR algorithm e.g. stratification. We replace this step with a smooth resampling procedure. The reason for this is as follows.

As noted in Pitt (2003), if particles $h_t^i, i = 1, ..., M$ drawn from the filtering density $f(h_t|Y_t; \theta)$ are slightly altered then the proposal samples, $h_{t+1}^i, i = 1, ..., M$ will also alter only slightly, as in the case of a highly persistent transition function, for example. But on the other hand, the discrete probabilities associated with these proposals will change as well, the implication of which is that the even if we generate the same uniforms at each time step, the resampled particles will not be close. Hence, the conventional weighted bootstrap methods are not smooth, in the sense of yielding an estimator of the likelihood which is not continuous as a function of the parameters θ. This has important implications for using gradient based maximization and computation of standard errors using conventional techniques (see also Liu and West (2000) and Polson, Stroud and Muller (2008)).

More specifically, it may be seen that in **Step 3** of the SIR algorithm we are sampling from the following empirical distribution function,

$$\widehat{F}(h_{t+1}) = \sum_{k=1}^{M} \lambda_{t+1}^{k} I(h_{t+1} - \widetilde{h}_{t+1}^{k}),$$

where $I(\bullet)$ is an indicator function. Sampling from this step function is what leads to the discontinuities as we change the parameters even if we keep the random number seed fixed. However, we may replace this empirical distribution function by,

$$\widetilde{F}(h_{t+1}) = \sum_{k=1}^{M} \lambda_{t+1}^{k} G\left(\frac{h_{t+1} - \widetilde{h}_{t+1}^{k}}{h_{t+1}^{k+1} - \widetilde{h}_{t+1}^{k}}\right),$$

where the $\widetilde{h}_{t+1}^{k}$ are sorted in ascending order and some adjustments, found in Appendix A, are imposed for the smallest and largest points. We have chosen the distribution function $G(x) = x$ corresponding to a Uniform distribution although other choices are possible. Importantly as $M \to \infty$, $\widetilde{F}(h_{t+1}) \to \widehat{F}(h_{t+1}) \to F(h_{t+1}|Y_t)$. It is straightforward and quick to invert this function. The computational overhead is in principle $O(M \times \log M)$ due to sorting though in practice we found this to be largely irrelevent. The method of smooth resampling is described in further detail in Appendix A. In the following section we shall describe the general method for estimation for the stochastic volatility model with both jumps and leverage. The simpler models, standard SV and SV with leverage, may of course be estimated in the same way imposing the necessary restrictions.

3·3 *Implementation of Stochastic Volatility with Leverage and Jumps Model*

Given the replacement of resampling step (**Step 3**) of the basic SIR algorithm with a smooth resampling scheme, implementing the particle filter for parameter estimation in the context of the vanilla SV model (see Section 2.1) is straightforward. In the SV with leverage model equation (2.5), $f(h_{t+1}|h_t; y_t)$ is highly non-linear. This make it difficult to obtain a good approximation via procedures such as the Extended Kalman Filter or by linearizing the state-space form by taking log-square transformations (See Harvey and Shephard (1996)). There are non-trivial implementational complications arising due to this non-linearity if we were to estimate such a model using MCMC or importance sampling, for example. Our method circumvents these issues since **Step 1** of the algorithm is still implemented straightforwardly.

Let us now consider the SV with leverage and jumps model. **Steps 2** and **3** remain the same but **Step 1** is now slightly more complicated and nests two additional steps (**1a** and **1b**) which will be described below. Since the returns process can jump with a certain probability, this necessitates simulating ϵ_t from the mixture density $f(\epsilon_t|h_t, y_t)$, which is then fed into the state equation, $h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \rho \epsilon_t + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t$. Here,

$$f(\epsilon_t|h_t, y_t) = \sum_{j=0}^{1} f(\epsilon_t|J_t = j; h_t, y_t) \Pr(J_t = j|h_t, y_t), \tag{3·4}$$

where $\Pr(J_t = 1|h_t, y_t)$ is the conditional probability of a jump. We establish that the functional form of the mixture is given by,

$$f(\epsilon_t|h_t, y_t) = \delta(\frac{y_t}{\exp(h_t/2)}) \Pr(J_t = 0|h_t, y_t) + N(\upsilon_{\epsilon_1}, \sigma^2_{\epsilon_1}). \Pr(J_t = 1|h_t, y_t).$$

The derivation of this mixture in addition to computation of the moments υ_{ϵ_1}, $\sigma^2_{\epsilon_1}$, probability $\Pr(J_t = 1|h_t, y_t)$ and the associated distribution function is detailed in the Appendix B. We need to implement a sub-algorithm for **Step 1** in the case of SV with leverage and jumps.

Sub-algorithm used within SIR, for t=0,..,T-1:
We have samples $h^i_t \sim f(h_t|Y_t)$ for $i = 1, ..., M$.

Step 1. $\begin{cases} \textbf{(1a)} & \textit{For } i = 1 : M, \textit{ sample } \epsilon^i_t \backsim f(\epsilon^i_t|h^i_t, y_t). \\ \textbf{(1b)} & \textit{For } i = 1 : M, \textit{ sample } \widetilde{h}^i_{t+1} \sim f(h_{t+1}|h^i_t; y_t; \epsilon^i_t). \end{cases}$

It is evident from the description of the components of $f(\epsilon_t|h_t, y_t)$ that this density with be characterized by mass at a unique point, $y_t \exp(-h_t/2)$, but continuous elsewhere, and governed by the moments of $N(\upsilon_{\epsilon_1}, \sigma^2_{\epsilon_1})$.

In the context of the particle filter we require samples $\epsilon^i_t \backsim f(\epsilon^i_t|h^i_t, y_t)$, $i = 1, ..., M$. The method for obtaining these sample is provided in the Appendix B. Once these samples are obtained they are fed through the state equation in order to implement **Step 1b.** The states are initialized using the unconditional density $f(h_1) \backsim N(\mu, \frac{\sigma^2_\eta}{1-\phi^2})$. The non-normalized weights for **Step 2** in the SIR algorithm are of the form,

$$\begin{aligned} f(y_{t+1}|\widetilde{h}^i_{t+1}, \sigma^2_J) &= (1-p)\left\{\frac{1}{\sqrt{2\pi\exp(\widetilde{h}^i_{t+1})}}\exp\left(-\frac{1}{2}\frac{y^2_{t+1}}{\exp(\widetilde{h}^i_{t+1})}\right)\right\} \\ &\quad + p\left\{\frac{1}{\sqrt{2\pi(\exp(\widetilde{h}^i_{t+1}) + \sigma^2_J)}}\exp\left(-\frac{1}{2}\frac{y^2_{t+1}}{\exp(\widetilde{h}^i_{t+1}) + \sigma^2_J}\right)\right\}, \end{aligned} \quad (3{\cdot}5)$$

for $i = 1, ..., M$. Once we are able to resample in a smooth manner as described in Section 3.2 and Appendix A, the log-likelihood function associated with the particle filtering scheme becomes straight forward to construct[3]. We record at each time step the Monte Carlo estimator of the empirical prediction density[4], i.e.

$$\widehat{l}_{t+1} = \log \widehat{f}(y_{t+1}|\theta; Y_t) = \log\left\{\frac{1}{M}\sum_{i=1}^{M} f(y_{t+1}|\widetilde{h}^i_{t+1}, \sigma^2_J)\right\}, \quad (3{\cdot}6)$$

[3]See Pitt (2002) for a detailed discussion of other possible schemes.

[4]**Bias correction:** It should be noted that at the present the log-likelihood will not be unbiased. In order to correct this we can use the usual Taylor expansion method. Abstarcting from likelihoods we have the large sample result that our estimated likelihood, $\overline{L}$, is unbiased for the true likelihood L, with $E[\overline{L}] = L$ and $Var[\overline{L}] = \frac{\sigma^2}{M}$. Therefore we have,

$$E[\log \overline{L}] \simeq \log L \; - \frac{1}{2}\frac{\sigma^2}{M\; L^2},$$

an approximation which is very good for large M. Hence we can bias correct by substituting L as $\overline{L}$, setting

$$\widehat{\log\; L} = \log \overline{L} + \frac{1}{2}\frac{\widehat{\sigma}^2}{M\; \overline{L}^2}.$$

After running through time we calculate,

$$\log \widehat{L}(\theta) = \sum_{t=1}^{T} \widehat{l}_t. \tag{3·7}$$

Essentially, our approach utilizes simulation to approximate the true likelihood. It has been demonstrated in Del-Moral (2004) that the particle filter provides unbiased estimate of the true likelihood function $L(\theta)$, such that $\widehat{L}(\theta) \overset{a.s.}{\to} L(\theta)$ as $M \to \infty$ and $E[\widehat{L}(\theta)] = L(\theta)$. The resulting simulated maximum likelihood estimator has asymptotic properties as discussed in Gourieroux and Monfort (1996, Ch. 3). The estimator is consistent if T and $M \to \infty$. In addition, when T and $M \to \infty$ and $\sqrt{T}/M \to 0$ the simulated maximum likelihood estimator is asymptotically equivalent to the maximum likelihood estimator[5].

As long as the transition and measurement densities are continuous in h_{t+1} and θ, we can sufficiently ensure $\log \widehat{L}(\theta)$ will be continuous in θ. The important point to note here is that within the implementation framework set out for the general SV with leverage and jumps model by setting parameters, σ_J^2 and p to zero we recover the SV with leverage specification. Furthermore, setting $\rho = \sigma_J^2 = p = 0$ we recover the standard SV specification.

Our implementation of the particle filter in the context of the SV with leverage and jumps also allows us to estimate the probability of a jump, i.e. $\Pr(J_t = 1|Y_{t-1}) = \int \Pr(J_t = 1|y_t; h_t) f(h_t|Y_{t-1}) dh_t$ straightforwardly as,

$$\widehat{\Pr}(J_t = 1|Y_{t-1}) = \frac{1}{M} \sum_{i=1}^{M} \Pr(J_t = 1|y_t, h_t^i). \tag{3·8}$$

3·4 *Diagnostics*

Standard approaches involved in specification analysis of time-series models is to investigate the properties of residuals in terms of their dynamic structure and unconditional distributions. This is infeasible given the latent dimension of the model under consideration. Alternatively therefore, in order to test the hypothesis that the prior and model are true, we require the distribution function,

$$\mathrm{u}_t = F(y_t|Y_{t-1}) = \int F(y_t|h_t) f(h_t|Y_{t-1}) dh_t. \tag{3·9}$$

In the specific case of SV with leverage and jumps, the distribution function can be estimated by,

$$\widehat{\mathrm{u}}_t = (1-p) \left\{ \frac{1}{M} \sum_{i=1}^{M} \Phi\left(\frac{y_t}{\exp(h_t^i/2)} \right) \right\} + p \left\{ \frac{1}{M} \sum_{i=1}^{M} \Phi\left(\frac{y_t}{\sqrt{\exp(h_t^i) + \sigma_J^2}} \right) \right\}, \tag{3·10}$$

where $\Phi(.)$ denotes the standard normal distribution function .If the prior and model were true, then the estimated distribution functions, $\widehat{\mathrm{u}}_t \backsim UID(0,1)$, for $t = 1, ..., T$, as $M \to \infty$ (See Rosenblatt (1952)).

[5]These are Propositions 3.1 and 3.2 in Chapter 3 of Gourieroux and Monfort (1996). Simulation-based estimators have been implemented and developed in other contexts such as discrete response models by Pakes and Pollard (1989), Lee (1992) and Sauer and Keane (2009).

3·5 *Model Comparison*

Given the framework described we can conduct model comparison be computing marginal likelihoods of competing models. We have the marginal likelihood as,

$$f(y|W_k) = \int f(y|\theta_k; W_k)f(\theta; W_k)d\theta_k,$$

where $f(y|\theta_k; W_k)$ is our likelihood approximation via the particle filter for model W_k ($k = 1, ..., K$) given the model specific maximum likelihood estimate of the parameter vector θ_k resulting from the optimization of the likelihood function. We may express this as,

$$f(y|W_k) = \int \frac{f(y|\theta_k; W_k)f(\theta_k; W_k)}{g(\theta_k|y, W_k)} g(\theta_k|y, W_k)d\theta_k,$$

where $g(\theta_k|y, W_k)$ is a multivariate Gaussian or t-distribution centered at maximum likelihood estimate (or the mode of $f(y|\theta_k; W_k)f(\theta_k; W_k)$) with the variance given by the inverse of the observed information matrix. This importance sampling scheme leads to an approximation,

$$f(y|W_k) = \sum\nolimits_{j=1}^{S} \frac{f(y|\theta_k^j; W_k)f(\theta_k^j; W_k)}{g(\theta_k^j|y, W_k)},$$

where $\theta_k^j \sim g(\theta_k|y, W_k)$. In practice this may only take a small number of draws as the posterior may be close to being log-quadratic (asymptotically under the usual assumptions this will be the case). Once the the appropriate prior density $f(\theta_k; W_k)$ is selected this model comparison scheme based on the ratios of marginal likelihoods between competing models can be implemented. Given the fact that we integrate out the parameter vector when computing the marginal likelihoods, we do not fall victim to the nuisance parameter problem encountered in similar contexts using likelihood ratio tests.

4 Simulation Experiments

4·1 *Stochastic Volatility with Leverage*

After running the smooth particle filter we maximize the estimated log-likelihood function with respect to $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$. We now investigate the performance of our maximum likelihood estimator for the SV with leverage case. First we simulate two time series of length 1000 and 2000 with parameter values $\theta = (\mu, \phi, \sigma_\eta^2, \rho) = (0.5, 0.975, 0.02, -0.8)$ and run the smooth particle filter 50 times using different random number seeds for the smooth particle filter for each run. These values for parameters are typical of those used in the literature in similar contexts. The resulting estimated log-likelihoods for each run are then maximized estimates with respect to θ. This is carried out for $M = 300$ and 600. The average of 50 maximum likelihood estimates ($\overline{ML_s}$) and 50 variance estimates ($\overline{Var}$) along with the variance for the sample of maximum likelihood estimates ($Var(ML_s)$), are reported for each case considered. The variance covariance matrix is estimated using the variance of the scores, i.e. the outer product of gradients (OPG) estimator. We chose to use this estimator as opposed to taking the negative of the inverse of the Hessian matrix at the mode in order to maintain the robustness of our procedure. Results are given in Table 1.

It is informative to consider the ratio of the variance of the maximum likelihood estimates in Table 1 to the variance of each parameter with respect to the data. These are, for $M = 300, T = 1000$: (0.0281, 0.0124, 0.0095, 0.0497); $M = 600, T = 1000$:(0.0078, 0.0046, 0.0062, 0.0192) and $M = 300, T = 2000$: (0.0171, 0.0142, 0.0094, 0.0223) and $M = 600, T = 2000$:(0.00757,

M=300, T=1000				M=300, T=2000			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$		$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5447	0.6491	0.01832	μ	0.4087	0.3848	0.0066
ϕ	0.9770	0.0033	0.00004	ϕ	0.9766	0.0022	0.00003
σ_η^2	0.0143	0.0015	0.00002	σ_η^2	0.0153	0.0010	0.000009
ρ	-0.7938	0.1867	0.00931	ρ	-0.8166	0.1106	0.00247
M=600, T=1000				M=600, T=2000			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$		$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5461	0.6792	0.00534	μ	0.4095	0.4181	0.00322
ϕ	0.9767	0.0034	0.000016	ϕ	0.9765	0.0023	0.000012
σ_η^2	0.0144	0.0016	0.0000098	σ_η^2	0.0154	0.0011	0.000004
ρ	-0.7946	0.1868	0.00392	ρ	-0.8175	0.1178	0.00150

Table 1: *Fixed dataset. Performance of the smooth particle filter for the stochastic volatility with leverage model for two cases, T=1000 and 2000; considering M=300, 600 for each case.True parameters,* $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$.

M=200			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5107	1.3207	2.3506
ϕ	0.9726	0.0081	0.0073
σ_η^2	0.0206	0.0043	0.0045
ρ	-0.7859	0.80467	0.6067
M=500			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5154	1.3499	2.3482
ϕ	0.9728	0.0087	0.0057
σ_η^2	0.0204	0.0042	0.0044
ρ	-0.7895	0.8008	0.5722
M=3000			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5164	1.3047	2.353
ϕ	0.9728	0.0097	0.0053
σ_η^2	0.0205	0.0043	0.0040
ρ	-0.7911	0.7877	0.5627

Table 2: *50 different datasets. Analysis of the maximum likelihood estimator for stochastic volatility with leverage model for cases, M=200, 500 and 3000. T=1000 in all cases.* True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$

0.00489, 0.00393, 0.01186). There is a substantial reduction in these ratios as M increase which is illustrated by kernel density estimates in Figure 1.

Next, we generated 50 different time series each of length $T = 1000$, with fixed values of parameters $\theta = (\mu, \phi, \sigma_\eta^2, \rho) = (0.5, 0.975, 0.02, -0.8)$. Keeping the random number seed fixed we run the smooth particle filter in turn for each of the time series and maximize the estimated log-likelihood with respect to θ for each run. The average of 50 maximum likelihood estimates ($\overline{ML_s}$)and 50 variance estimates ($\overline{Var}$) along with mean squared errors ($Var(ML_s)$) are reported in Table 2 for each of three cases considered. The histograms in Figure 2 indicate that the distribution of the parameters is not too far from normality. In all cases we find that biases are not significantly different from zero[6] and the true values of the parameters lie well within their 95% confidence limits. The procedure does not throw up any extreme outliers and we have no problem with convergence to the mode. It is worth noting that the variance estimates ($\overline{Var}$) alter only slightly when taking different values of M. The reason for this is that as $|\rho| \longrightarrow 1$, the state equation (2.5) tends to a deterministic GARCH-type diffusion which anchors the variability around the point estimate thus decreasing sensitivity to the value of M[7].

4·2 *Stochastic Volatility with Leverage and Jumps*

Now we investigate parameter estimation in the case of SV with leverage and jumps model. We run the smooth particle filter and maximize the estimated log-likelihood with respect to the parameter vector $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$.We again begin by simulating two time series of length 1000 and 2000, setting parameters $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p) = (0.5, 0.975, 0.02, -0.8, 10, 0.10)$. These values for parameters are in line with those that have been adopted in similar contexts in the literature.

M=300, T=1000				M=300, T=2000			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$		$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5595	3.0020	0.06023	μ	0.4770	1.2653	0.03098
ϕ	0.9648	0.0103	0.00021	ϕ	0.9680	0.00522	0.00013
σ_η^2	0.0458	0.0186	0.00020	σ_η^2	0.0338	0.00661	0.000123
ρ	-0.7072	1.0326	0.01629	ρ	-0.7419	0.7275	0.01352
σ_J^2	10.176	813.98	6.9054	σ_J^2	7.7568	207.71	1.1959
p	0.0769	0.0754	0.00120	p	0.11263	0.0659	0.00079
M=600, T=1000				M=600, T=2000			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$		$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5650	2.9623	0.03853	μ	0.4830	1.2760	0.01097
ϕ	0.9648	0.0103	0.00013	ϕ	0.9681	0.0052	0.00005
σ_η^2	0.0461	0.0192	0.00012	σ_η^2	0.0338	0.0067	0.00008
ρ	-0.7026	1.0333	0.00665	ρ	-0.7394	0.7425	0.00622
σ_J^2	10.174	823.13	2.5625	σ_J^2	7.7929	216.21	0.87021
p	0.0764	0.0771	0.00045	p	0.1115	0.0667	0.00047

Table 3: *Fixed dataset. Performance of the smooth particle filter for the stochastic volatility model with leverage and jumps for two cases, T=1000 and 2000; considering M=300, 600 for each case.*

[6] $E(\widehat{\theta}) - \theta = Bias \backsim N(0, \frac{MSE}{50})$ where the mean squared error (MSE) is $E[(\widehat{\theta} - \theta)^2]$.

[7] As a conseuence of this, since now we only require small M, incorporating leverage also indirectly also reduces computation time of our procedure.

M=200			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.49151	2.0908	1.7937
ϕ	0.97101	0.013972	0.018073
σ_η^2	0.022110	0.0086659	0.0071614
ρ	-0.84684	1.3943	1.1835
σ_J^2	9.8470	954.42	621.81
p	0.10458	0.13002	0.069915
M=500			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.50006	2.2045	1.5714
ϕ	0.97186	0.015317	0.010667
σ_η^2	0.022389	0.0097163	0.0064737
ρ	-0.83714	1.4793	1.1215
σ_J^2	9.8013	1018.7	637.60
p	0.10358	0.13667	0.063125
M=900			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.49720	2.1724	1.6280
ϕ	0.97203	0.014559	0.0099983
σ_η^2	0.022474	0.0090217	0.0075645
ρ	-0.84500	1.5008	1.1664
σ_J^2	9.8524	1007.0	648.20
p	0.10367	0.13505	0.065325

Table 4: *50 different datasets. Analysis of the maximum likelihood estimator for stochastic volatility with leverage and jumps model for cases, M=200, 500 and 900. T=1000 in all cases.*

The smooth particle filter is run 50 times using a different random number seed but keeping the dataset fixed. The estimated log-likelihood is maximized with respect to θ for each run. In Table 3, the average of the resulting 50 maximum likelihood estimates ($\overline{ML_s}$)and 50 variance estimates ($\overline{Var}$), along with the variance for the sample of maximum likelihood estimates ($Var(ML_s)$),are reported for different cases considered. The variance covariance matrix is again estimated using the OPG estimator.

We examine the ratio of the variance of the maximum likelihood estimates to the variance of each parameter with respect to the data. These are, for $M = 300, T = 1000$: (0.0201, 0.0209, 0.0108, 0.01578, 0.0085, 0.0159); $M = 600, T = 1000$:(0.0131, 0.0132, 0.0062, 0.0064, 0.0032, 0.0059); $M = 300, T = 2000$: (0.0245, 0.0251, 0.0186, 0.0186, 0.0058, 0.0121) and $M = 600, T = 2000$: (0.0086, 0.0095, 0.0121, 0.0084, 0.0040, 0.0070). These ratios suggest that the variance of the simulated estimates is small in comparison to the variance induced by the data. The reduction in these ratios as M increases is illustrated by kernel density estimates in Figure 3.

Next, we generate 50 different time series each of length $T = 1000$, setting values of parameters $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p) = (0.5, 0.975, 0.02, -0.8, 10, 0.10)$. Keeping the random number seed fixed we run the smooth particle filter in turn for each of the time series and maximize the estimated log-likelihood with respect to θ for each run. The average of 50 maximum likelihood estimates ($\overline{ML_s}$) and 50 variance estimates ($\overline{Var}$) along with mean squared errors ($Var(ML_s)$) are reported in Table 4, for each of three cases considered. Variance estimates are computed using the OPG estimator for the variance covariance matrix.

Small Jump - High Intensity			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.21240	3.4545	2.7098
ϕ	0.97290	0.0066247	0.0072527
σ_η^2	0.029170	0.013178	0.014784
ρ	-0.85636	0.70314	0.66880
σ_J^2	0.63322	95.169	60.103
p	0.23544	4.3614	6.8037

Table 5: *50 different datasets. Analysis of the maximum likelihood estimator for stochastic volatility with leverage and jumps model. We set parameter values;* $\mu = 0.25$, $\phi = 0.975$, $\sigma_\eta^2 = 0.025$, $\rho = -0.8$, $\sigma_J^2 = 0.5$ and $p = 0.10$. *M=300 and T=1000* .

Large Jump - Low Intensity			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.25359	1.9024	1.3926
ϕ	0.97293	0.0063159	0.0074348
σ_η^2	0.026733	0.0066633	0.0070814
ρ	-0.82253	0.55547	0.42255
σ_J^2	9.6201	2162.1	3884.2
p	0.013252	0.075626	0.020192

Table 6: *50 different datasets. Analysis of the maximum likelihood estimator for stochastic volatility with leverage and jumps model. We set parameter values;* $\mu = 0.25$, $\phi = 0.975$, $\sigma_\eta^2 = 0.025$, $\rho = -0.8$, $\sigma_J^2 = 10$ and $p = 0.01$. *M=300 and T=1000* .

The corresponding histograms in Figure 4 suggest convergence towards the mode and that we are not far from normality. In testing for bias we find very encouraging results. We find that all parameters, except the leverage parameter ρ which is estimated with slight bias, are either within, or on the boundary of their 95% confidence limits. It should be pointed out that unbiasedness is an asymptotic property associated with the likelihood and there is no reason for us to not expect some degree of bias given a time series of moderate length such as what we are considering for purposes of our experiments. The results are stable across different values of M. We note that the settings for this experiment were one of a large jump variance σ_J^2 with very high intensity, p. One would expect the additional noise induced by these setting to render the estimation of the stochastic volatility components less accurate. Our findings suggest that in spite of having large jumps with high intensity, our procedure delivers highly reliable estimates for all the parameters.

We proceed to investigate how the error in estimation is affected by varying the intensity and jump size. The results in Table 5 suggest that having small jumps occurring with high intensity induces a slight amount of bias is estimating of σ_η^2, ρ and p. In contrast, if large jumps occur at a very low frequency, i.e. setting $p = 0.01$, the accuracy of our estimates is greatly enhanced (see Table 6). All parameters fall well within their 95% confidence limits with only moderate bias in the estimate of leverage (see Figures 5 and 6). Using simulated data generated with large jump-low intensity calibration for θ, we provide the diagnostic check (see Section 3.4) for the SV with leverage and jumps model in addition to a plot of the data, filtered standard deviation and filtered jump probabilities in Figure 7 [8]. The diagnostic test appears to be indicate the prior and model are correct.

5 Empirical Examples

We now employ our methodology to estimate the three models; (i) stochastic volatility (SV), (ii) stochastic volatility with leverage (SVL) and (iii) stochastic volatility with leverage and jumps (SVLJ) using daily returns data for four different price indices, namely S&P 500, FTSE 100, Dow Jones and Nasdaq. For each of the series, the parameter estimates along with standard errors[9], log-likelihood values and Akaike information criterion (AIC) for the three specifications are reported in Tables 7,8,9 and10. The results reported indicate that the gain in likelihood points moving from the SVL to SVLJ specification is small compared to the gain in points by incorporating only leverage in the SV specification.

	ML Estimate	Standard Error
SV: Log-lik value $= -3044.1$, AIC $= 6094.2$		
μ	0.1717	0.1872
ϕ	0.9832	0.0056
σ_η^2	0.0218	0.0048
SVL: Log-lik value $= -2996.4$, AIC $= 6000.8$		
μ	0.2432	0.0983
ϕ	0.9739	0.0040
σ_η^2	0.0307	0.0044
ρ	-0.7944	0.0426
SVLJ: Log-lik value $= -2993.7$, AIC $= 5999.4$		
μ	0.2498	0.1010
ϕ	0.9766	0.0041
σ_η^2	0.0266	0.0048
ρ	-0.8303	0.0444
σ_J^2	5.2607	2.0453
p	0.0079	0.0026

Table 7: *Parameter estimates for S&P 500 daily returns data for period, 16/05/1995 - 24/04/2003. M=500.*

For the time span of data considered we find that leverage is extremely important component in modelling stochastic volatility whereas including jumps in addition to leverage yield a statistically significant gain in the case of Dow Jones and Nasdaq. We illustrate the actual returns data, along with the quantiles of filtered standard deviation and filtered jump probabilities for S&P 500 and Dow Jones in Figure 8 and 9. Results of the diagnostic check on the SVLJ specification for all four series are provided in Figure 10.

[8] Note that the plots in each of these figures illustrate output generated by a single run of the smooth particle filter.

[9] We use the outer product of gradients estimator for the variance covariance matrix.

	ML Estimate	Standard Error
SV: Log-lik value = -3004.4, AIC = 6014.8		
μ	0.0751	0.2093
ϕ	0.9859	0.0052
σ_η^2	0.0176	0.0046
SVL: Log-lik value = -2972.8, AIC = 5965		
μ	0.1135	0.1257
ϕ	0.9842	0.0037
σ_η^2	0.0201	0.0040
ρ	-0.7825	0.0509
SVLJ: Log-lik value = -2972.2, AIC = 5956.4		
μ	0.0638	0.1262
ϕ	0.9836	0.0038
σ_η^2	0.0212	0.0042
ρ	-0.8029	0.0584
σ_J^2	1.4652	1.0376
p	0.0132	0.0229

Table 8: *Parameter estimates for FTSE 100 daily returns data for period, 01/07/1996 - 01/03/2004. M=500.*

	ML Estimate	Standard Error
SV: Log-lik value = -2623.5, AIC = 5253		
μ	-0.2379	0.1717
ϕ	0.9830	0.0061
σ_η^2	0.0183	0.0043
SVL: Log-lik value = -2586.7, AIC = 5181.4		
μ	-0.1745	0.0963
ϕ	0.9805	0.0035
σ_η^2	0.0213	0.0037
ρ	-0.8282	0.0410
SVLJ: Log-lik value = -2579.2, AIC = 5170.4		
μ	-0.1557	0.0988
ϕ	0.9825	0.0034
σ_η^2	0.0189	0.0036
ρ	-0.8640	0.0451
σ_J^2	18.706	12.43
p	0.0018	0.0014

Table 9: *Parameter estimates for Dow Jones Composite daily returns data for period, 01/05/2000 - 31/12/2007. M=500.*

	ML Estimate	Standard Error
SV: Log-lik value = −3457.6, AIC = 6921.2		
μ	0.7193	0.5488
ϕ	0.9973	0.0016
σ_η^2	0.0054	0.0156
SVL: Log-lik value = −3429.3, AIC = 6866.6		
μ	0.4877	0.1834
ϕ	0.9942	0.0014
σ_η^2	0.0077	0.0016
ρ	-0.8291	0.0543
SVLJ: Log-lik value = −3423.3, AIC = 6858.6		
μ	0.2615	0.1564
ϕ	0.9930	0.2284
σ_η^2	0.0131	0.0016
ρ	-0.8411	0.0034
σ_J^2	0.4781	0.0503
p	0.5599	0.0848

Table 10: *Parameter estimates for Nasdaq Composite daily returns data for period, 01/05/2000 - 31/12/2007. M=500.*

6 Conclusion

In this paper we have attempted to provide a unified methodology in order to conduct likelihood-based inference on the unknown parameters of discrete-time stochastic volatility models incorporating a leverage effect and jumps in the returns process. It is demonstrated how the likelihood can be approximated using output generated by the particle filter and how smooth resampling can be undertaken in order ensure that the likelihood estimator is continuous as a function of the unknown parameters. The latter enabling the use of gradient-based (Newton-Raphson type) maximization algorithms. A great advantage of our unified methodology is that it allows us to easily obtain the filtered path of the states, jump probabilities (i.e. in the case of SV with leverage and jumps) and output required to perform diagnostics. This is in contrast to competing methodologies which deliver these objects following often complicated modifications to their basic structures.

Implementation is easy and has the benefit of being both faster is terms of CPU computation time and more general than many alternatives in the literature. With regards to generality, we note that the standard SV and SV with leverage models are restricted forms of the SV with leverage and jumps model; it is highlighted how the proposed methodology can easily facilitate parameter estimation for all three types of models without any alteration in the basic structure of the algorithm and as a consequence also allow for model comparison. Our Monte Carlo experiments indicate that the method is both robust and efficient. Especially, when examining finite sample bias we find very encouraging results even when considering very high jump intensity. On our simulated data, the methodology was efficient and quick even for very long time series. In unreported results we tried $T = 20,000$.

As empirical examples we model four different daily returns series using our approach. Our results reveal that the inclusion of a leverage effect is extremely important when modelling stochastic volatility, as indicated by the substantial gain in the log-likelihood over the standard SV model. Additionally, we find that inclusion of jumps in returns, after having incorporated leverage leads to a relatively less dramatic gain in log-likelihood.

References

Bates, D. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark Options. Review of Financial Studies, 9, 69-107.

Bates, D. (2000). Post-'87 crash fears in S&P 500 futures options. Journal of Econometrics, 94, 181-238.

Bakshi, G., C. Cao and Z. Chen (1997). Empirical performance of alternative options pricing models. Journal of Finance, 52, 2003-2049.

Black, F. (1976). Studies of stock market volatility changes. Proceedings of the American Statistical Association, Business and Economic Statistics Section 177-181.

Carpenter, J. R., P. Clifford, and P. Fearnhead (1999). An improved particle filter for non-linear problems. IEE Proceedings on Radar, Sonar and Navigation 146, 2-7.

Chib, S., F. Nardari, and N. Shephard (2006). Analysis of high dimensional multivariate stochastic volatility models. Journal of Econometrics 134, 341-371.

Christie, A. A. (1982).The stochastic behaviour of common stock variances. Journal of Financial Economics 10, 407-432.

Christoffersen, P., Jocobs, K., and Mimouni, K. (2007). Models for S&P dynamics: Evidence from realized volatility, daily returns, and options prices.(McGill University, unpublished working paper).

Del-Moral, P. (2004). Feynman-Kac Formulae: Genealogical and Interacting Particle Systems with Applications. New York. Springer.

Durham, G. B. (2006). SV mixture models with application to S&P 500 index returns. Journal of Financial Economics (forthcoming).

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

Doucet, A., J.F.G. De Freitas and N. Gordon (2000). Sequential Monte Carlo Methods in Practice. Cambridge University Press, Cambridge.

Eraker, B., M. Johannes, and N. Polson (2003). Journal of Finance, 58(3), 1269-3000.

Engle, R., and Ng, V. (1993). Measuring and testing the impact of news in volatility. Journal of Finance, 43, 1749-1778

Gallant, A. R. and G. Tauchen (1998). Reprojection partially observed systems with applications to interest rate diffusions. Journal of the American Statistical Association, 93, 10-24.

Glosten, L.R., Jagannanthan, R. and D. Runkle (1993). Relationship between the expected value and the volatility of the excess return on stocks. Journal of Finance, 48, 1779-1802.

Gordon, N. J., D. J. Salmond, and A. F. Smith (1993). A novel approach to non-linear and non-Gaussian Bayesian state estimation. IEE-Proceedings F 140, 107-13.

Gourieroux, C. and Monfort, A. (1996). Simulation Based Econometric Methods. Oxford University Press, Oxford, pp 41-53.

Gweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculation of moments (with discussion). In: Bernardo, J.M., Berger, J.O., Dawid, A.P., Smith, A.F.M. (Eds.), Bayesian Statistics,Vol. 4. Oxford University Press, Oxford, pp. 169-193.

Harvey, A. C. and N. Shephard (1996). The estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business and Economic Statistics 14, 429-434.

Harvey, A. C. , E. Ruiz, and N. Shephard (1994). Multivariate stochastic variance models. Review of Economic Studies 61, 247-246.

Isard, M. and A. Blake (1996). Contour tracking by stochastic propagation of conditional density. Proceedings of the European Conference on Computer Vision, Cambridge 1, 343-356.

Jacquier, E., N.G. Polson and P.E. Rossi (1994). Bayesian analysis of stochastic volatility models. Journal of Business and Economic Statistics 12, 371-389.

Jacquier, E., N. G. Polson and P. E. Rossi (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics, 122(1), 185-212.

Johannas, M., Polson, N. G., and Stroud, J. R (2009). Optimal Filtering of Jump-Diffusions: Extracting Latent States from Asset Prices. Forthcoming Review of Financial Studies.

Sauer, R. and Keane M. P. (2009). A Computationally Practical Simulation Estimation Algorithm for Dynamic Panel Data Models with Unobserved Endogenous State Variables, International Economic Review, forthcoming.

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

Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. Journal of Computational and Graphical Statistics, 5, 1-25.

Liu, J. and R. Chen (1998). Sequential Monte Carlo methods for dynamic systems. Journal of American Statistical Association, 93, 1032-1044.

Liu, J. and West, M. (2001). Combined parameters and state estimation in simulation-based filtering. In: Sequential Monte Carlo Methods in Practice (by Doucet, A., J.F.G. De Freitas and N. Gordon) 97-233. Springer-Verlag, New York.

Merton, R. C. (1976). Option pricing when underlying stock returns and discontinuous. Journal of Financial Economics 3, 125-144.

Meyer, R. and J. Yu (2000). BUGS for Bayesian analysis of stochastic volatility model models. Econometrics Journal, 3, 198-215.

Nelson, D. (1991). Conditional heteroskedasticity in asset pricing: A new approach. Econometrica, 59, 347-370.

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

Pakes, A. and Pollard, D. (1989). Simulation and Asymptotics of Optimization Estimators. Econometrica, 57, No. 5, 1027-1057.

Lee, L. (1992). On Efficiency of Methods of Simulated Moments and Maximum Likelihood Estimation of Discrete Response Models. Econometric Theory, 8, 515-552.

Pitt, M. K. (2003). Smooth particle filters for likelihood evaluation and maximization. Unpublished working paper, University of Warwick.

Pitt, M. K. with Shephard, N.(1999). Filtering via simulation: auxiliary particle filter Journal of the American Statistical Association , 94, 590-9.

Polson, N. G., Stroud, J. S. and Muller, P. (2008). Particle Filtering with Sequential Parameter Learning, Journal of Royal Statistical Society, B, 70, 413-428.

Sandmann, G. and S.J. Koopman (1998). Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood, Journal of Econometrics, 87, No.2, 271-301.

Shephard, N. and M. K. Pitt (1997). Likelihood analysis of non-Gaussian measurement time series. Biometrika 84, 653-67.

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

7 Appendix A

Smooth Resampling

This procedure works by replacing the discrete cumulative distribution function (cdf) given by one that is smooth, thus providing particles from the filter which are smooth as a function of θ. Let us begin by assuming that we have a $1 \times M$ vector of elements h^i sorted in ascending order, with associated discrete probabilities, λ^i. The time subscript is suppressed for notational convenience. Abstracting from the notation used in the main text slightly, let the discrete cdf used in SIR be given by $\widehat{F}(h) = \sum_{i=1}^{M} \lambda^i I(h < h^i)$. This approximates the true cdf $F(h)$. In order to obtain a continuous interpolation for, $\widehat{F}(h)$ we proceed as follows.

We construct partitions of the sample space for h by defining region.i, $S_i = [h^i, h^{i+1}]$, $i = 1, ..., M-1$. Next we assign $\Pr(i)=\frac{1}{2}(\lambda^i + \lambda^{i+1})$, $\Pr(1)=\frac{1}{2}(2\lambda^1 + \lambda^2)$ and $\Pr(M-1)=\frac{1}{2}(\lambda^{M-1} + 2\lambda^M)$, such that these probabilities sum to unity. Within each region we have conditional densities given by,

$$
\begin{aligned}
g(h|i) &= \frac{1}{h^{i+1} - h^i},\ h \in S_i,\ i = 2, ..., M-2 \\
g(h|1) &= \begin{cases} \frac{\lambda^1}{2\lambda^1 + \lambda^2}, \text{ when } h = h^1 \\ \frac{\lambda^1 + \lambda^2}{2\lambda^1 + \lambda^2} \frac{1}{(h^2 - h^1)}, \text{ when } h \in S_1 \end{cases} \\
g(h|M-1) &= \begin{cases} \frac{\lambda^M}{\lambda^{M-1} + 2\lambda^M}, \text{ when } h = h^M \\ \frac{\lambda^1 + \lambda^2}{\lambda^1 + 2\lambda^2} \frac{1}{(h^M - h^{M-1})}, \text{ when } h \in S_{M-1} \end{cases}
\end{aligned}
$$

By following the above procedure we obtain a continuous interpolation for the discrete cdf; this 'continuous' cdf $\widetilde{F}(h)$ will pass through the mid-point of each step. As $M \to \infty$, $\widetilde{F}(h) \to \widehat{F}(h) \to F(h)$. We sample from the continuous density by selecting region i with $\Pr(i)$ and the sample from $g(h|i)$.We detail this resampling procedure below.

Once we obtain the continuous empirical cdf we the task is to implement smooth sampling, which will yield an ordered sample of particles, say, $h^{*1}, ..., h^{*M}$ .We use a stratified sampling scheme for purposes of this paper. Stratification reduces sample impoverishment and has been suggested by Kitagawa (1996), Carpenter et al. (1999) and Liu and Chen (1998). In an extreme case, after a certain amount of updates, the particle system may collapse to a single point resulting in a poor approximation to the required density[10]. In contrast to SIR which involves generating uniforms $u_1, ..., u_M \backsim UID(0,1)$, stratified sampling will require us to generate a single random variate $u \backsim UID(0,1)$ from which we can propagate sorted uniforms given by $u_j = (j-1)/M + u/M$, $j = 1, ..., M$.

If $\Pr(i) = \widetilde{\lambda}^i = \frac{1}{2}(\lambda^i + \lambda^{i+1})$, then the cumulative probability is given by $\overline{\lambda}^i = \sum_{s=1}^{i} \widetilde{\lambda}^s$, where $i = 1, ..., M-1$. Next we define the interval corresponding to region i as,

$$\left(\sum_{s=1}^{i-1} \widetilde{\lambda}^s , \sum_{s=1}^{i} \widetilde{\lambda}^s \right],$$

and the uniform(s) falling within the interval by,

$$u_j^* = \frac{u_j - (\sum_{s=1}^{i-1} \widetilde{\lambda}^s)}{\widetilde{\lambda}^i}.$$

We can now sample conditional upon that region, i.e. from $g(h|i)$ using the corresponding uniform(s) u_j^*. Since $g(h|i)$ is uniform, the sampled particles can be backed-out as

$$h^{*i} = (h^{i+1} - h^i) \times u_j^* + h^i. \tag{7·1}$$

Alrorithm: Smooth resampling

The algorithm given below samples the index corresponding to the region which are stored as, $r^1, r^2....., r^M$ and also the uniforms $u_1^*, ..., u_M^*$.

```
set s=0,j=1;
for (i=1 to M-1)
{
s=s+λ̃^i;

while (u_j ≤ s AND j ≤ M)
   {
      r^j = i;
      u*_j = ( u_j − (s − λ̃^i )) /  λ̃^i;
              j = j + 1;
   }
}
```

[10]In the less extreme case, a few particles may survive, but as noted by Carpenter et al (1999), the high degree of internal correlation yields summary statistics reflective of a substantially smaller sample. In order to compensate a very large number of particle will need to be generated.

8 Appendix B

Deriving the functional form of density $f(\epsilon_t|h_t, y_t)$

The conditional probability of a jump is given by,

$$\Pr(J_t = 1|h_t, y_t) = \frac{\Pr(y_t|J = 1)\Pr(J = 1)}{\Pr(y_t|J = 1)\Pr(J = 1) + \Pr(y_t|J = 0)\Pr(J = 0)},$$

$$= \frac{N(y_t|0; \exp(h_t) + \sigma_J^2)p}{N(y_t|0; \exp(h_t) + \sigma_J^2)p + N(y_t|0; \exp(h_t))(1-p)}. \tag{8·1}$$

Hence, $\Pr(J_t = 0|h_t, y_t) = 1 - \Pr(J_t = 1|h_t, y_t)$. Now since,

$$f(\epsilon_t|J = 1; h_t, y_t) \propto f(y_t|J = 1, h_t, \epsilon_t)f(\epsilon_t), \tag{8·2}$$

we can reformulate the conditional density $f(\epsilon_t|J = 1; h_t, y_t) \propto N(\ y_t|\ \epsilon_t \exp(h_t/2); \sigma_J^2) \times N(\epsilon_t|0; 1)$ in logarithmic form as,

$$\log\ f(\epsilon_t|J = 1; h_t, y_t) = const - \frac{1}{2}\frac{(y_t - \epsilon_t \exp(h_t/2))^2}{\sigma_J^2} - \frac{1}{2}\epsilon_t^2, \tag{8·3}$$

The resultant quadratic form facilitates completing the square to yield,

$$\log\ f(\epsilon_t|J_t = 1; h_t, y_t) = K - \frac{1}{2}\frac{(\epsilon_t - \upsilon_{\epsilon_1})^2}{\sigma_{\epsilon_1}^2}. \tag{8·4}$$

Computing moments υ_{ϵ_1} **and** $\sigma_{\epsilon_1}^2$

Taking the expression $\log\ f(\epsilon_t|J = 1; h_t, y_t) = const - \frac{1}{2}\frac{(y_t - \epsilon_t \exp(h_t/2))^2}{\sigma_J^2} - \frac{1}{2}\epsilon_t^2$,

First collect the squared terms corresponding to $-\frac{1}{2}\epsilon_t^2$;

$$\frac{1}{\sigma_{\epsilon_1}^2} = \frac{\exp(h_t)}{\sigma_J^2} + 1 = \frac{\exp(h_t)+\sigma_J^2}{\sigma_J^2},$$
$$\Longrightarrow \sigma_{\epsilon_1}^2 = \frac{\sigma_J^2}{\exp(h_t) + \sigma_J^2}.$$

Next those corresponding to ϵ_t;

$$\frac{\upsilon_{\epsilon_1}}{\sigma_{\epsilon_1}^2} = \frac{y_t \exp(h_t/2)}{\sigma_J^2},$$
$$\Longrightarrow \upsilon_{\epsilon_1} = \frac{y_t \exp(h_t/2)}{\exp(h_t) + \sigma_J^2}.$$

We hence establish that,

$$f(\epsilon_t|J_t = 1; h_t, y_t) = N(\upsilon_{\epsilon_1}, \sigma_{\epsilon_1}^2) \text{ where, } \upsilon_{\epsilon_1} = \frac{y_t \exp(h_t/2)}{\exp(h_t) + \sigma_J^2} \text{ and } \sigma_{\epsilon_1}^2 = \frac{\sigma_J^2}{\exp(h_t) + \sigma_J^2},$$

If the process does not jump, there is a dirac delta mass at the point,

$$f(\epsilon_t|J_t = 0; h_t, y_t) = \frac{y_t}{\exp(h_t/2)}. \tag{8·5}$$

Functional form of distribution funtion $F(\epsilon_t|h_t, y_t)$

$F(\epsilon_t|h_t, y_t)$ can thus be split into three regions with boundaries delineated as follows.

- $\Pr(J_t = 1|h_t, y_t). \int_{-\infty}^{\epsilon_t} f(\epsilon_t|J_t = 1, h_t, y_t)d\epsilon_t \;\; for\ \epsilon_t < y_t \exp(-h_t/2)$

- $\Pr(J_t = 1|h_t, y_t). \int_{-\infty}^{y_t \exp(-h_t/2)} f(\epsilon_t|J_t = 1, h_t, y_t)d\epsilon_t + (1 - \Pr(J_t = 1|h_t, y_t)) \;\; for\ \epsilon_t = y_t \exp(-h_t/2)$

- $\Pr(J_t = 1|h_t, y_t). \int_{-\infty}^{y_t \exp(-h_t/2)} f(\epsilon_t|J_t = 1, h_t, y_t)d\epsilon_t + (1 - \Pr(J_t = 1|h_t, y_t))$

  $+ \Pr(J_t = 1|h_t, y_t). \int_{\epsilon_t}^{+\infty} f(\epsilon_t|J_t = 1, h_t, y_t)d\epsilon_t \;\; for\ \epsilon_t > y_t \exp(-h_t/2)$

Sampling continuously from the mixture density $f(\epsilon_t|h_t, y_t)$

In the context of the particle filter, the generation of h_t^i, $i = 1, ..., M$ particles each time step, will give rise to densities, $f(\epsilon_t^i|h_t^i, y_t)$, $i = 1, ..., M$. The aim is thus to simulate $\epsilon_t^1, ..., \epsilon_t^M$, from corresponding densities $f(\epsilon_t^1|h_t^1, y_t), ..., f(\epsilon_t^M|h_t^M, y_t)$. We shall illustrate the procedure to simulate ϵ_t^1 from density $f(\epsilon_t^1|h_t^1, y_t) = \sum_{j=0}^{1} f(\epsilon_t^1|J_t = j; h_t^1, y_t) \Pr(J_t = j|h_t^1, y_t)$. Given that the density corresponding to particle h_t^1 is of the form,

$$f(\epsilon_t^1|h_t^1, y_t) = \delta(y_t \exp(-h_t^1/2)). \Pr(J_t = 0|h_t^1, y_t) + N(\upsilon_{\epsilon_1}^1, \sigma_{\epsilon_1}^{21}). \Pr(J_t = 1|h_t^1, y_t).$$

For notational simplicity we set $x^* = y_t \exp(-h_t^1/2)$ and the conditional probability of a jump to be $\Pr^J = \Pr(J_t = 1|h_t^1, y_t)$. The associated distribution function $F(\epsilon_t^1|h_t^1, y_t)$ is thus of the form.

$F(\epsilon_t^1|h_t^1, y_t) = \Pr^J . \int_{-\infty}^{\epsilon_t^1} f(\epsilon_t^1|J_t = 1, h_t, y_t)d\epsilon_t^1 \;\; \text{for } \epsilon_t^1 < x^*$

$F(\epsilon_t^1|h_t^1, y_t) = \Pr^J . \int_{-\infty}^{x^*} f(\epsilon_t^1|J_t = 1, h_t^1, y_t)d\epsilon_t^1 + (1 - \Pr^J) \;\; \text{for } \epsilon_t^1 = x^*$

$F(\epsilon_t^1|h_t^1, y_t) = \Pr^J . \int_{-\infty}^{x^*} f(\epsilon_t^1|J_t = 1, h_t^1, y_t)d\epsilon_t^1 + (1 - \Pr^J) \; + \Pr^J . \int_{\epsilon_t^1}^{+\infty} f(\epsilon_t^1|J_t = 1, h_t^1, y_t)d\epsilon_t^1 \;\; \text{for } \epsilon_t^1 > x^*$

As is evident from the form of $F(\epsilon_t^1|h_t^1, y_t)$, the height of this distribution function can be split into three distinct regions. First generate a uniform random variate $u_1 \backsim UID(0, 1)$, then record within which region u_1 falls. Conditional on the recorded region we then invert in accordance with the following scheme.

- If $u_1 \leq \Phi(\frac{x^* - \upsilon_{\epsilon_1}^1}{\sigma_{\epsilon_1}^1}). \Pr^J$, we sample $\epsilon_t^1 = \upsilon_{\epsilon_1}^1 + \sigma_{\epsilon_1}^1 \Phi^{-1}(\frac{u_1}{\Pr^J})$.

- If $\Phi(\frac{x^* - \upsilon_{\epsilon_1}^1}{\sigma_{\epsilon_1}^1}). \Pr^J < u_1 \leq \Phi(\frac{x^* - \upsilon_{\epsilon_1}^1}{\sigma_{\epsilon_1}^1}). \Pr^J + (1 - \Pr^J)$, we sample $\epsilon_t^1 = y_t \exp(-h_t^1/2)$.

- If $u_1 > \Phi(\frac{x^* - \upsilon_{\epsilon_1}^1}{\sigma_{\epsilon_1}^1}). \Pr^J + (1 - \Pr^J)$, we sample $\epsilon_t^1 = \upsilon_{\epsilon_1}^1 + \sigma_{\epsilon_1}^1 \Phi^{-1}(\frac{u_1 - (1 - \Pr^J)}{\Pr^J})$.

$\Phi(.)$ denotes the standard normal distribution function. The above probability integral transform procedure is repeated for each of the generated uniforms $u_{1,} ......, u_M \backsim UID(0, 1)$ in order to obtain the required sample $\epsilon_t^i \backsim f(\epsilon_t^i|h_t^i, y_t)$, $i = 1, ..., M$.

Figures

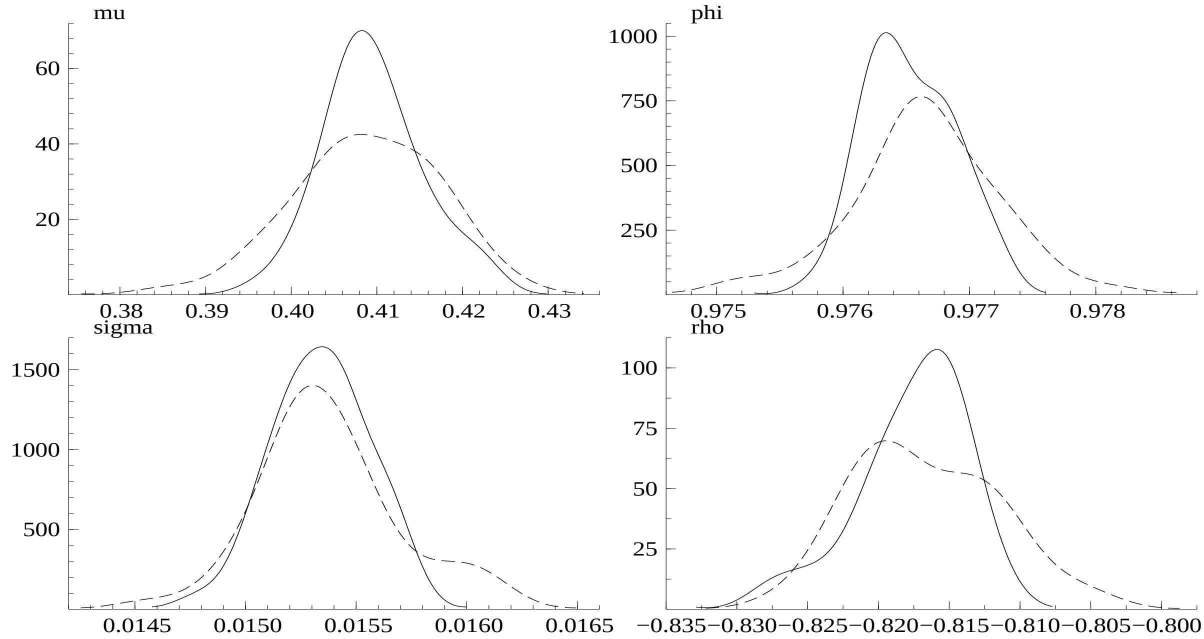

Figure 1: Fixed dataset. Dashed line:Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model; $T = 2000$ and $M = 300$. Solid line: Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model; $T = 2000$ and $M = 600$. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$.

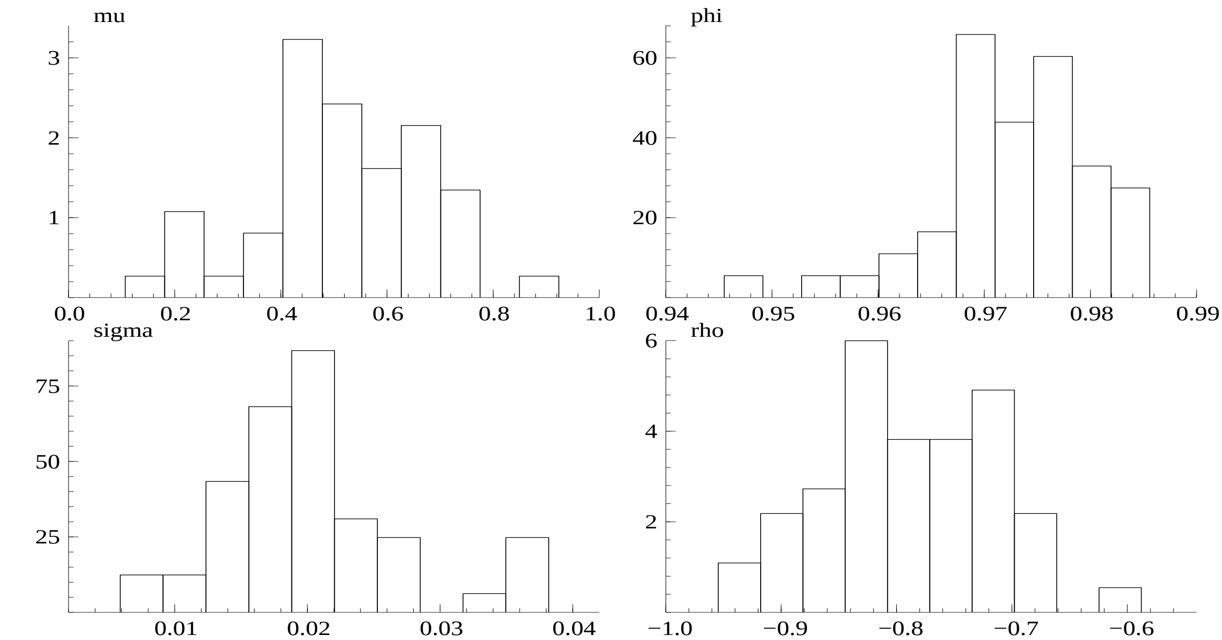

Figure 2: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$. $M = 500$ and $T = 1000$.

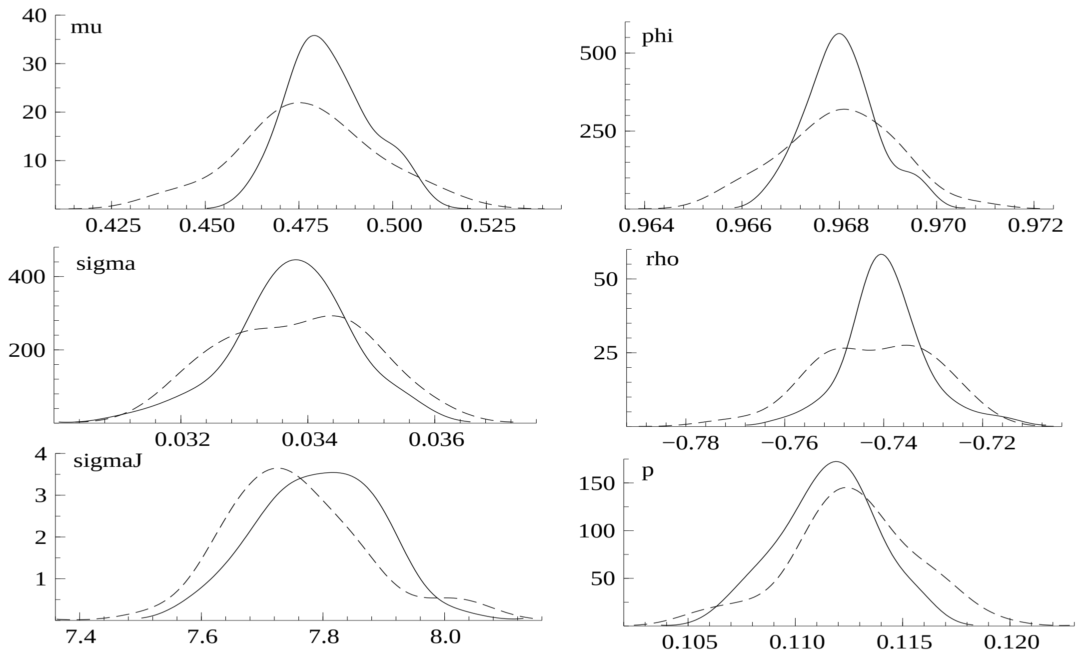

Figure 3: Fixed datasets. Dashed line:Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model; $T = 2000$ and $M = 300$. Solid line: Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model; $T = 2000$ and $M = 600$. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.10$.

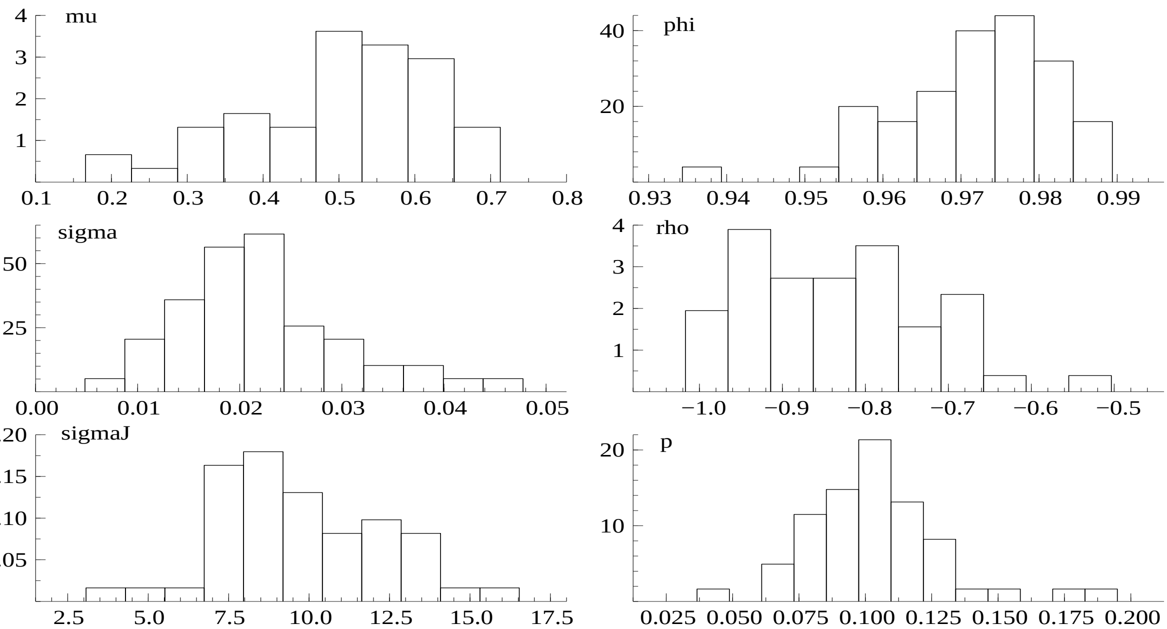

Figure 4: 50 different dataset. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.10$. $M = 500$ and $T = 1000$.

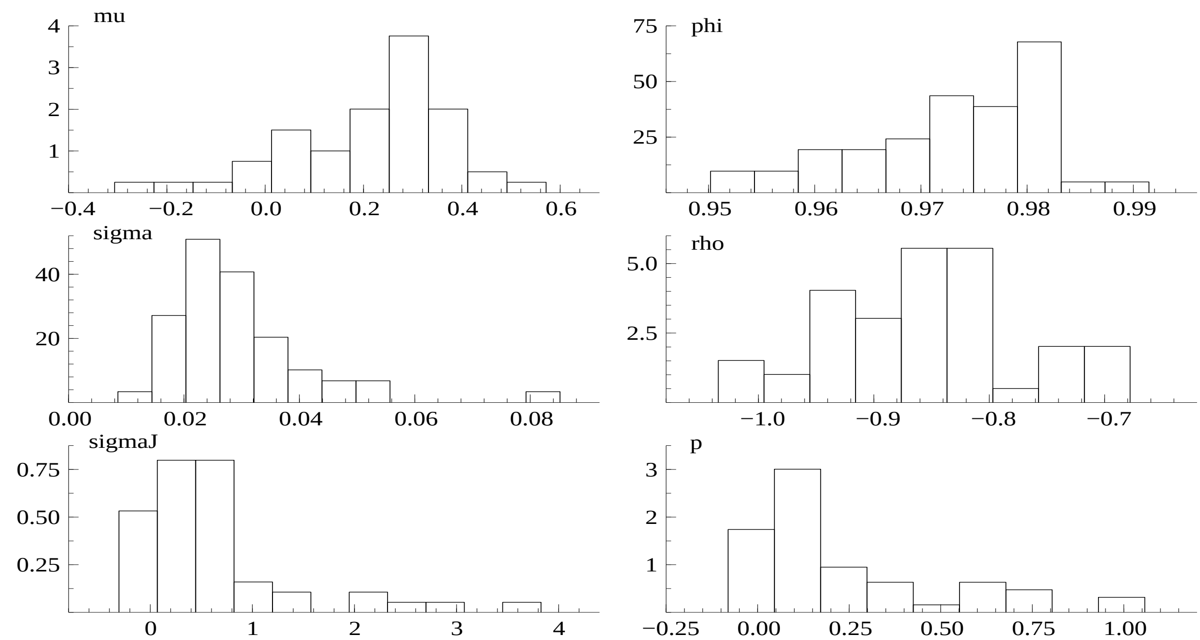

Figure 5: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.25, \phi = 0.975, \sigma_\eta^2 = 0.025$ and $\rho = -0.8, \sigma_J^2 = 0.5$ and $p = 0.10$. $M = 300$ and $T = 1000$.

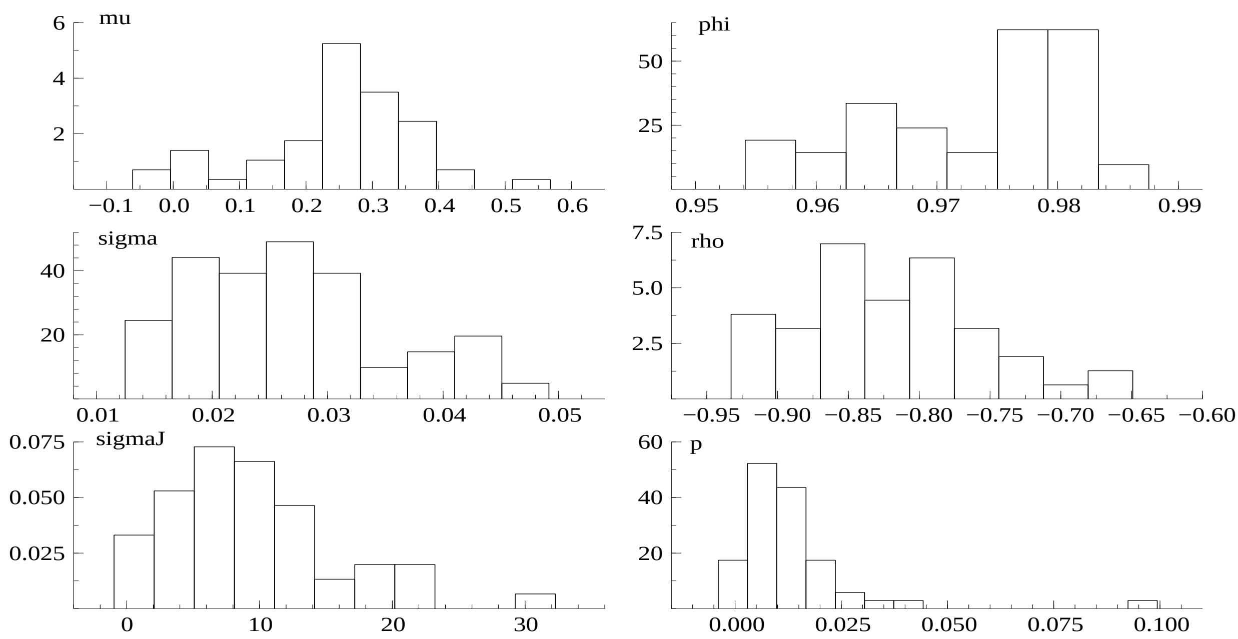

Figure 6: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.25, \phi = 0.975, \sigma_\eta^2 = 0.025$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.01$. $M = 300$ and $T = 1000$.

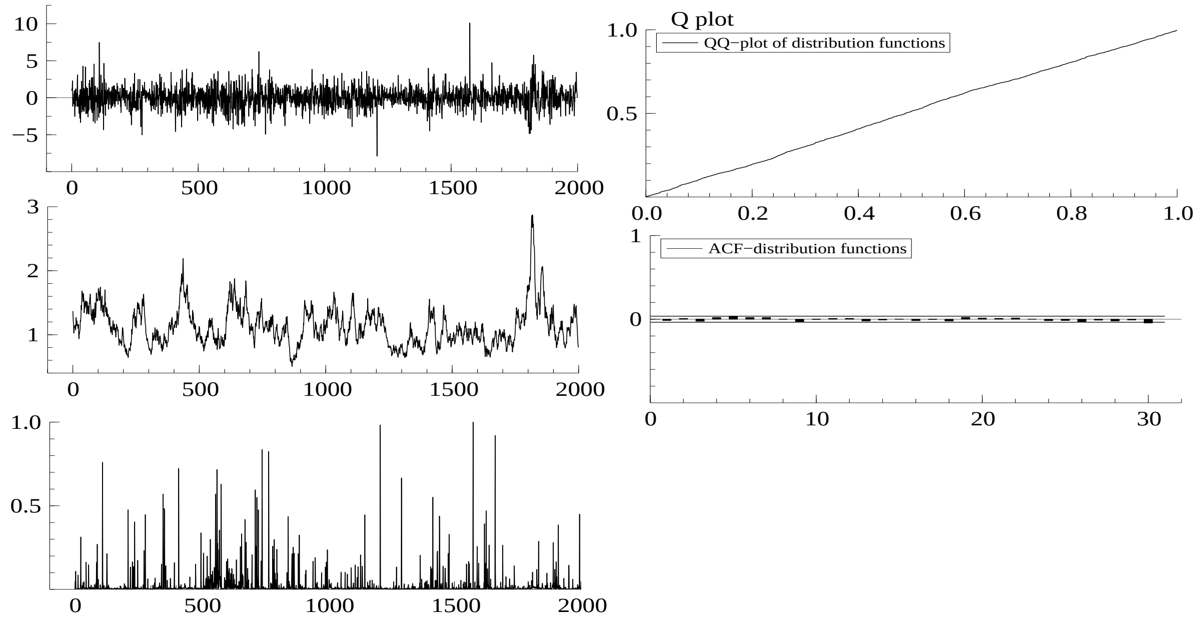

Figure 7: Using simulated data with $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p) = (0.25, 0.975, 0.025, -0.8, 10, 0.01)$ and a single run of the smooth particle filter. LEFT PANEL: (i) Plot of data, (ii) filtered standard deviation, (iii) estimated jump probabilities. RIGHT PANEL: (i)QQ-plot of estimated distribution functions, $\widehat{u}_t^J$ (ii) correlogram of $\widehat{u}_t^J$. $M = 500, T = 2000$.

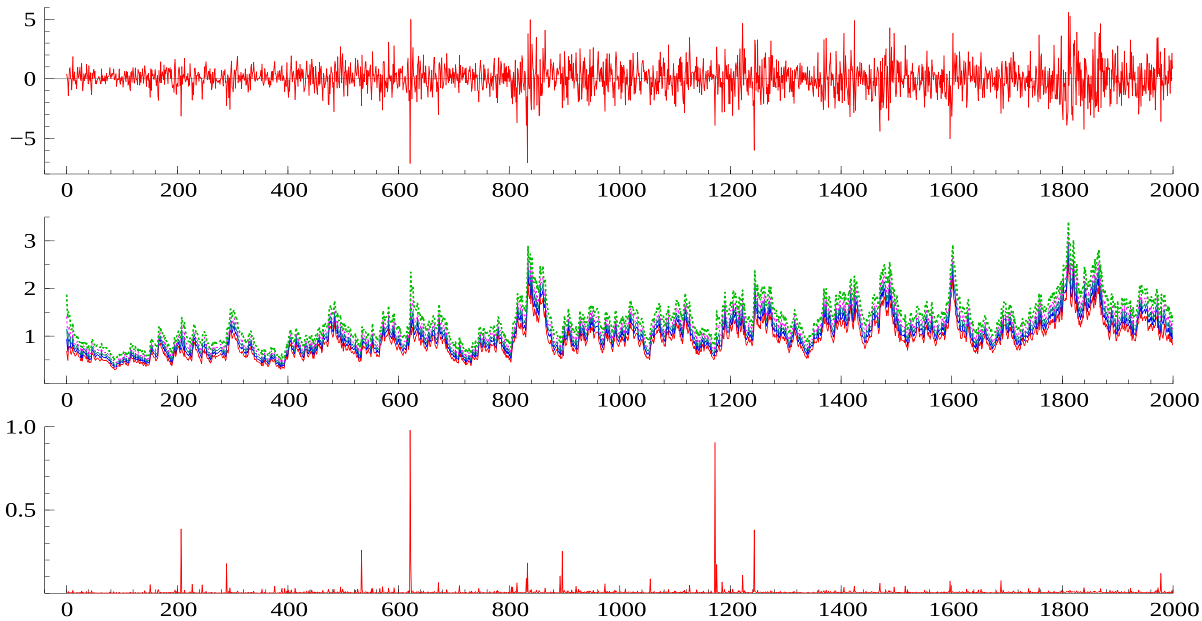

Figure 8: SV with leverage and jumps model. Daily S&P 500 returns over the period 16/05/1995 - 24/04/2003. (i) returns data, (ii) quantiles of filtered standard deviation and (iii) estimated jump probabilities. $M = 500$.

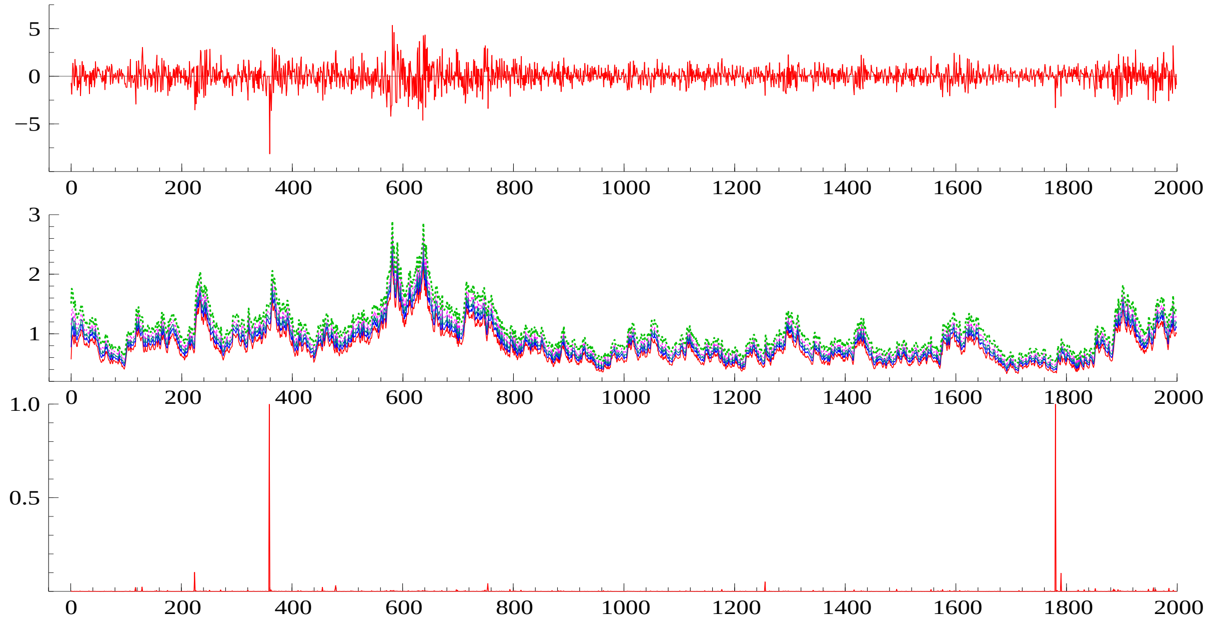

Figure 9: SV with leverage and jumps model. Dow Jones Composite 65 Stock Average returns over the period 01/05/2000 - 31/12/2007.(i) returns data, (ii) quantiles of filtered standard deviation and (iii) estimated jump probabilities. $M = 500$.

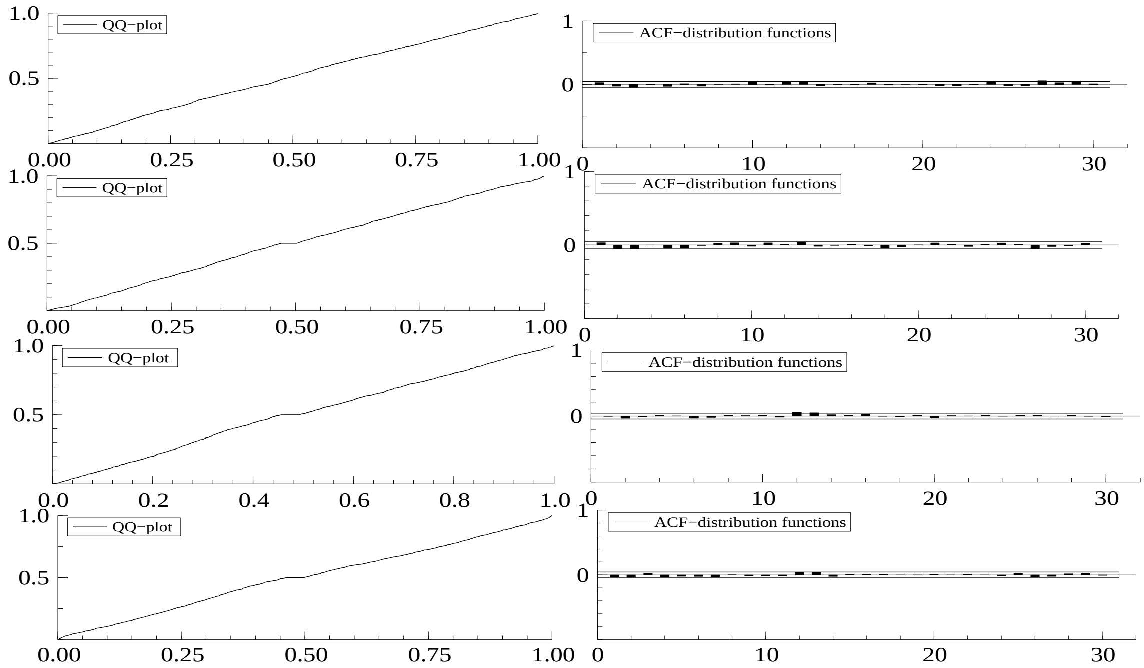

Figure 10: Diagnostic tests in the context of modelling stochastic volatility with leverage and jumps. Left panel: QQ-plot of estimated distribution functions, $\widehat{u}_t^J$. Right panel: Associated correlogram of $\widehat{u}_t^J$. The first row provides diagnostics for the case of S&P 500, the second row FTSE 100, the third row, Dow Jones and fourth row, Nasdaq.