Modeling Stochastic Volatility with Leverage and Jumps: A 'Smooth' Particle Filtering Approach

Sheheryar Malik and Michael K Pitt
Department of Economics, University of Warwick, Coventry CV4 7AL

October 8, 2008

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. 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.

DRAFT: Preliminary and Incomplete

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 specificiations 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.

In the subsequent subsections we first describe how the state-space form of the vanilla SV model can be adapted to allow for leverage, provide a brief literature survey. In Section (2) we analyse the 'smooth particle filter' and describe how it can be used to carry out likelihood-based inference of model parameters. Section (3) deals with a simulated examples of the SV with leverage and SV with leverage and jumps model. We provide monte carlo evidence on the performance of the estimator, in addition to a diagnostic check on filter performance. In Section (5) we investigate empirical...

Section (6) concludes.

1·1 *Stochastic Volatility with Leverage:*

Given that the standard stochastic volatility model with uncorrelated measurement and state equation disturbances is given by,

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

where,

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

Here y_t is the observed return, $\{h_t\}$ are the unobserved log-volatilities, μ is the drift in the state equation, σ_η is the volatility of log-volatility and ϕ is the persistence parameter. Allowing for the disturbances to be correlated implies the covariance matrix has the form,

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

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 in (1.1) can be re-written as,

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

By substituting , $\epsilon_t = y_t \exp(-h_t/2)$ into (1.2), the model adopts the following Gaussian non-linear state-space form where the parameter ρ measures the leverage effect.

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \rho \; y_t \exp(-h_t/2) + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t \end{aligned} \tag{1·3}$$

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 stochastic volatility model with leverage is given by, $y_t|\eta_t \sim N(\rho \exp(h_t/2)\eta_t \; ; (1-\rho^2)\exp(h_t))$ where $h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t$.

1·2 *Stochastic Volatility with Leverage and Jumps*

Modifying the stochastic volatility with leverage model to allow for a jumps in the returns process would yield,

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) + J_t \varpi_t \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t, \quad t = 1, ..., T \end{aligned} \tag{1·4}$$

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

$J_t = j$ 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 $f(\eta_t|\epsilon_t) = N(\rho\epsilon_t; 1 - \rho^2)$.In constrast to the case with just leverage, now the returns process can jump with a certain probability. This necessitates simulating ϵ_t from the mixture density $f(\epsilon_t|h_t, y_t)$, where,

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

If the unconditional probability of a jump (intensity) is p, it follows that the corresponding conditional probability is given by

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

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

hence, $\Pr(J_t = 0|h_t, y_t) = 1 - \Pr(J_t = 1|h_t, y_t)$.Given that,

$$f(\epsilon_t|J = 1; h_t, y_t) \propto f(y_t|J = 1, h_t, \epsilon_t)f(\epsilon_t)$$

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,

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

The resultant quardratic form facilitates completing the square to yield,

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

This implies that,

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

Please refer to the appendix for derivation of these moments. If the process does not jump, there is a dirac delta mass at the point,

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

Hence, our required denisty may be written as,

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

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 $N(\upsilon_{\epsilon_1}, \sigma_{\epsilon_1}^2)$. This associated distribution function, $F(\epsilon_t|h_t, y_t)$ can be split into three regions with boundaries delineated as follows[1].

[1] The continuous component is indicative of fact that identifying whether a jump has actually occured is confounded by the presence of noise.

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

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

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

1·3 *Survey of the Literature*

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.

Amongst the earliest contributions to 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 SV models (see Harvey, Ruiz and Shephard (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. A problem with the QML approach is that $\log \epsilon_t^2$ is a poor approximated by the normal distribution yielding a quasi-likelihood estimator with poor finite sample properties. When applied to daily CRSP (Centre for Research in Security Prices) and SP30 (Standard and Poors), the authors find evidence of a leverage effect.

In order to correct for this Kim, Shephard and Chib (1998) the develop an alternative approach for analysis of SV models employing Markov chain Monte Carlo 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. Following Omori et al. (2007), they specify $d_t = sign(y_t) = I(\epsilon_t > 0) - I(\epsilon_t \leq 0)$, $y_t^* = \log y_t^2 = h_t + \varsigma_t$ where $\varsigma_t = \log \epsilon_t^2$ and $y_t = d_t \exp(y_t^*/2)$. In the case where $\rho = 0$, the signs of $y = (y_{1,\ldots,}y_T)'$ and $y^* = (y_{1,\ldots,}^* y_T^*)'$ are independent hence we can neglect $d = (d_{1,\ldots,}d_T)'$, hence y^* is a linear process. When $\rho \neq 0$ the situation is complicated given that now, d_t can not be ignored since, $\eta_t|d_t, \varsigma_t \backsim N(d_t\rho\sigma \exp(\varsigma_t/2), \sigma^2(1-\rho^2))$ (equation (5) in Omori et al). Their approach relies on approximating the bivariate conditional density $f(\varsigma_t, \eta_t|d_t) = f(\varsigma_t)f(\eta_t|d_t, \varsigma_t)$ using a ten-component mixture of bivariate normal distributions. They apply this approach to fit a model to daily returns of TOPIX and find evidence of leverage.

Jacquier, Polson and Rossi (1994) also propose a Bayesian 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. In contrast to, for example, Kim, Shephard and Chib (1998) and Omori et al. (2007), the method advocated by these authors deals with basic SV model, i.e. not linearized by log-squared transformation.

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

[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).

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 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 Shephard et al. (1996, 1998, 2004) is preferred.

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 (2002) 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, ϵ_t. This paper uses simulated maximum likelihood approach to conduct inference.

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 methodolgy 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.

2 The Smooth Particle Filter

This paper is concerned with evaluation of state-space models via particle filter. We model time series $\{y_t, t = 1, ..., T\}$ using 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 can not 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 Shephard and Pitt (1997), Pitt and Shephard (1999) and reviewed 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 Pitt (2003).

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

2·1 *Particle filtering*

2·1.1 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,

$$f(h_{t+1}|Y_{t+1}) \propto f(y_{t+1}|h_{t+1}) \int f(h_{t+1}|h_t)dF(h_t|Y_t) \tag{2·1}$$

Prediction step :Passing these particles, $\left\{h_t^i\right\}, i = 1, ..., M$, through the transition density will yield the prior 'empirical' density of the state i.e. $h_{t+1}^i \backsim f(h_{t+1}|h_t^i)$. The state evolution is initialized by some density $f(h_0)$.

Updating step :The prior is combined the likelihood $f(y_{t+1}|h_{t+1})$ in order to update. This step relies on a result by Smith and Gelfand (1992) which states that Bayes theorem can be implemented as a weighted bootstrap, (see also Rubin (1987)).

Theorem 2.1 *Following Smith and Gelfand (1992), 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* $M \to \infty$.

This is the basis for the updating step, in that on receiving the measurement y_{t+1}, we evaluate the likelihood at each prior sample h_{t+1}^i.We proceed to calculate *normalized* weights,

$$\lambda_{t+1}^i = \frac{f(y_{t+1}|h_{t+1}^i)}{\sum_{i=1}^M f(y_{t+1}|h_{t+1}^i)}$$

and hence we obtain a discrete distribution over h_{t+1}^i with probability mass $\lambda_{t+1}^i, i = 1, ..., M$. The weighted bootstrap (Rubin (1998) refers to this as SIR; Sampling Importance Resampling) involves resampling h_{t+1}^i, N times using weights λ_{t+1}^i will yield an approximation of the desired posterior density, $f(h_{t+1}|Y_{t+1})$. This prediction-updating procedure is iterated through the data to in order to produce empirical filtering densities,

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

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.

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.

2·1.2 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 prediction decompostion (e.g. see Harvey(1993)). In order to estimate this function, we exploit the relationship,

$$f(y_{t+1}|\theta; Y_t) = \int f(y_{t+1}|h_{t+1};\theta) f(h_{t+1}|Y_t;\theta) dh_{t+1} \tag{2·3}$$

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. We have weights (discrete probabilities),

- $f(y_{t+1}|h^i_{t+1})$when consdering the stochastic volatility with leverage, or
- $f(y_{t+1}|h^i_{t+1}, \sigma^2_J)$[3] in the case of stochastic volatility with leverage and jumps,

associated with proposals $h^i_{t+1}, i = 1, ..., M$. The SIR technique as employed in the Gordon et al. (1993) algorithm works by replicating those particles with large weights and removes those with negligible weights. This allows the resultant particles to be more concentrated in domains of higher posterior probability.

As noted in Pitt (2003), if particles $h^i_t, i = 1, ..., M$ drawn from the filtering density $f(h_t|Y_t;\theta)$ are slightly altered then the proposal samples, $h^i_{t+1}, i = 1, ..., M$ 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 huge implications for using gradient based maximization and computation of standard errors using conventional techniques since the likelihood surface will not be smooth.

2·1.3 Smooth bootstrapping:

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 need 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. The discrete cdf used in SIR is given by $\widehat{G}(h) = \sum_{i=1}^{M} \lambda^i I(h < h^i)$ approximates the true cdf $G(h)$. In order to obtain a continuous interpolation for, $\widehat{G}(h)$ 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$(M-1)$=$\frac{1}{2}(\lambda^{M-1} +$

[3] The exact form of these weights will be provided in the subseqent section. Recall, these weights are normalized before resampling .

$2\lambda^M$), such that these probabilities sum to unity. Within each region we have conditional densities given by,

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

By following the above procedure we attain a continuous interpolation for the discrete cdf , and this ‘continuous’ cdf $\widetilde{G}(h)$ will pass through the mid-point of each step. As $M \to \infty$, $\widehat{G}(h) \to \widetilde{G}(h) \rightharpoonup G(h)$. We sample from the continuous density by selecting region i with $\Pr(i)$ and sample from $g(h|i)$. We detail the 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[4]. 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[5],

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

[4]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.

[5]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 $u_1^*, ..., u_M^*$.

Smooth resampling Algorithm:
set s=0,j=1;
for (i=1 to M-1)
{
s=s+$\widetilde{\lambda}^i$;
while ($u_j \leq s$ AND $j \leq M$)
  {
    $r^j = i$;
    u_j^* = ($u_j - (s - \widetilde{\lambda}^i$)) / $\widetilde{\lambda}^i$;

2·1.4 Constructing the likelihood function:

Once we are able to resample in a smooth manner, the log-likelihood function associated with the particle filtering scheme used in this paper becomes straight forward to construct. See Pitt (2002) for a detailed discussion of other possible schemes. We record at each time step the Monte Carlo estimator of the empirical prediction density, i.e. in the case of **stochastic volatility with leverage,**

$$\widehat{l}^{L}_{t+1} = \log \widehat{f}(y_{t+1}|\theta; Y_t) = \frac{1}{M}\sum_{i=1}^{M} \log f(y_{t+1}|h^i_{t+1}) \qquad (2\cdot4)$$

where the *non-normalized* weights have the following Gaussian form,

$$f(y_{t+1}|h^i_{t+1}) = \frac{1}{\sqrt{2\pi \exp(h^i_{t+1})}} \exp\left(-\frac{1}{2}\frac{y^2_{t+1}}{\exp(h^i_{t+1})}\right) \qquad (2\cdot5)$$

For **stochastic volatility with leverage and jumps**,

$$\widehat{l}^{LJ}_{t+1} = \log \widehat{f}(y_{t+1}|\theta; Y_t) = \frac{1}{M}\sum_{i=1}^{M} \log f(y_{t+1}|h^i_{t+1}, \sigma^2_J) \qquad (2\cdot6)$$

where the *non-normalized* weights[6] take the form,

$$\begin{aligned} f(y_{t+1}|h^i_{t+1}, \sigma^2_J) &= \left(\frac{1}{\sqrt{2\pi \exp(h^i_{t+1})}} \exp\left(-\frac{1}{2}\frac{y^2_{t+1}}{\exp(h^i_{t+1})}\right)\right)(1-p) \\ &+ \left(\frac{1}{\sqrt{2\pi(\exp(h^i_{t+1}) + \sigma^2_J}} \exp\left(-\frac{1}{2}\frac{y^2_{t+1}}{\exp(h^i_{t+1}) + \sigma^2_J}\right)\right)(p) \qquad (2\cdot7) \end{aligned}$$

After running through time we calculate[7],

$$\log \widehat{L}(\theta) = \sum_{t=1}^{T} \widehat{l}_{t+1} \qquad (2\cdot8)$$

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.

3 Simulation Experiments:

3·1 *Stochastic Volatility with Leverage*

We initialize the states using unconditional density $f(h_1) \backsim N(\mu, \frac{\sigma^2_\eta}{1-\phi^2})$. After running the smooth particle filter we maximise the estimated log-likelihood function with respect to $\theta =$

```
            j = j + 1;
        }
    }
```

[6]Observe that setting the jumps compnents $\sigma^2_J = p = 0$, we recover the stochastic volatility with leverage (or without leverage) specification for the weights.

[7]Superscripts L and LJ on $\widehat{l}_{t+1}$ specifiying either case are suppressed for notational convenience.

$(\mu, \phi, \sigma_\eta^2, \rho)$. The associated variance estimates are obtained by taking the negative of the inverse of the hessian for θ at the mode.

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 write the distribution function as,

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

This cdf can be estimated by;

$$\widehat{\mathrm{u}}_t = \frac{1}{M} \sum_{i=1}^{M} F(y_t|h_t^i) \quad \text{where } i = 1, ..., M \tag{3·1}$$

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 $M \to \infty$. (Rosenblatt (1952)).

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 each run, maximizing the resulting estimated log-likelihood estimates with respect to θ for each run. 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 estimates are are obtained by taking the negative of the inverse of the hessian matrix for θ at the mode. Results are given in Table1.

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.*

It is informative to consider the ratio of the variance of the ML 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, 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 Appendix (Fig: 1 and 2).

M=200			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5107	0.6045	2.3506
ϕ	0.9726	0.0062	0.0073
σ_η^2	0.0206	0.0043	0.0045
ρ	-0.7859	0.2248	0.6067
M=500			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5154	0.6744	2.3482
ϕ	0.9728	0.0051	0.0057
σ_η^2	0.0204	0.0033	0.0044
ρ	-0.7895	0.1569	0.5722
M=3000			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.5164	0.6824	2.353
ϕ	0.9728	0.0056	0.0053
σ_η^2	0.0205	0.0034	0.0040
ρ	-0.7911	0.2126	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.*

Next, we generated 50 different time series each of length $T = 1000$, with fixed values of parameters $\theta = (\mu, \phi, \sigma_\eta^2, \rho) = (0.5, 0.975, 0.02, -0.8)$ as before. Keeping the random number seed fixed and running the smooth particle filter in turn for each of the time series,we maximize the estimated log-likelihoods 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 the Appendix (Fig 3,4 and 5) 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[8] and the true values of the parameters lie well within their 95% confidence limits. The procedure does not throw up any outliers and we have no problem with convergence to the mode.

3·2 *Stochastic Volatility with Leverage and Jumps*

In comparison to the SV with leverage case, the incorporation of a jumps in the returns process requires our procedure to include an additional step, i.e. simulating $\epsilon_t^i \backsim f(\epsilon_t^i|h_t^i, y_t)$, $i = 1, ..., M$. The method for obtaining the required sample is detailed in the Appendix. Once these samples are obtained we can evaluate the integral $f(\eta_t|y_t, h_t) = \int f(\eta_t|\epsilon_t)f(\epsilon_t|y_t, h_t)d\epsilon_t$ where $f(\eta_t|\epsilon_t) = \rho\ \epsilon_t + \sqrt{(1-\rho^2)}\xi_t$, and $\xi_t \backsim N(0, 1)$. The states are again intialized using unconditional density $f(h_1) \backsim N(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$. We run the smooth particle filter[9] and maximize the

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

[9] Our implementation of the particle filter in the context of the leverage and jumps case allows us to estimate the posterior probability of a jump, $\Pr(J_t = 1|Y_{t-1}) = \int \Pr(J_t = 1|y_t; h_t) f(h_t|Y_{t-1})dh_t$ by

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

estimated log-likelihood with respect to the parameter vector $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$. In order to test the hypothesis that the prior and model are true in the case of SV with leverage and jumps we estimate the distribution function for this case, i.e. $\mathrm{u}_t^J = F(y_t|Y_{t-1}) = \int F(y_t|h_t)f(h_t|Y_{t-1})dh_t$ as,

$$\widehat{\mathrm{u}}_t^J = \left(\frac{1}{M}\sum_{i=1}^{M} F(y_t|h_t^i)\right)(1-p)+ \quad \left(\frac{1}{M}\sum_{i=1}^{M} F(y_t|h_t^i+\sigma_J^2)\right)(p) \ , \quad i = 1, ..., M \qquad (3{\cdot}2)$$

and proceed to test if $\widehat{\mathrm{u}}_t^J \backsim UID(0,1),\ t = 1, ..., T.$

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.*

We again run the smooth particle filter 50 times keeping the dataset fixed, setting paramters $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p) = (0.5, 0.975, 0.02, -0.8, 10, 0.10)$ and using a different random number seed for the filter for each run. In Table 3, 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 different cases considered. We compute the variance covariance matrix in using the variance of the scores, i.e. the outer product of gradients (OPG) estimator.

We consider the ratio of the variance of the ML 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 the variance of the ML estimates ratios as M increases is illustrated by kernel density estimates in Appendix (Fig: 6 and 7).

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 and run the smooth particle filter in turn for each of the time series, we maximize the estimated log-likelihoods 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 covanriance matrix.

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 dataset. Analysis of the maximum likelihood estimator for stochastic volatility with leverage model for cases, M=200, 500 and 900. T=1000 in all cases.*

The corresponding histograms in Appendix (Figs.8,9 and 10) suggest convergence towards the mode and 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% confindence limits. 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 arrival 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 inspite of having large jumps with high intensity, our procedure delivers highly reliable estimates for all the parameters.

Next we investigate how the error in estimation is affected by varying the intensity and jumps size.

The results in Table 5 suggest that having small jumps occuring 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

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	9516.9	60.103
p	0.23544	43614	6.8037

Table 5: *50 different dataset. 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 .*

low frequency, i.e.setting $p = 0.01$, the accuracy of our estimates is greatly enhanced (Table 6.). All parameters fall well within their 95% confidence limits with only moderate bias in the estimate of leverage. See histograms in Appendix (Fig. 11 and 12).

4 Empirical examples:

We now employ our methodology to estimate the following 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, FTSE100, Dow Jones and Nasdaq. For each of the series', the parameter estimates along with standard errors and log-likelihood values for the three specifications are reported in Tables 7,8,9 and10

S&P 500		
	ML estimate	Standard error
SV		
μ	0.1717	0.1872
ϕ	0.9832	0.0056
σ_η^2	0.0218	0.0048
Log-likelihood value $= -3044.1$		
SVL		
μ	0.2432	0.0983
ϕ	0.9739	0.0040
σ_η^2	0.0307	0.0044
ρ	-0.7944	0.0426
Log-likelihood value $= -2996.4$		
SVLJ		
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
Log-likelihood value $= -2993.7$		

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

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. Table 11. provided likelihood ratio statistics when comparing different specifications.

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 Dow Jones and Nasdaq. Results of the diagnostic check on the SVLJ specification reveal that the pior and model are true in all cases. See Appendix (Fig 13,14,15 and 16).

FTSE 100		
	ML estimate	Standard error
SV		
μ	0.0751	0.2093
ϕ	0.9859	0.0052
σ_η^2	0.0176	0.0046
Log-likelihood value $= -3004.4$		
SVL		
μ	0.1135	0.1257
ϕ	0.9842	0.0037
σ_η^2	0.0201	0.0040
ρ	-0.7825	0.0509
Log-likelihood value $= -2972.8$		
SVLJ		
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
Log-likelihood value $= -2972.2$		

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

Dow Jones		
	ML estimate	Standard error
SV		
μ	-0.2379	0.1717
ϕ	0.9830	0.0061
σ_η^2	0.0183	0.0043
Log-likelihood value $= -2623.5$		
SVL		
μ	-0.1745	0.0963
ϕ	0.9805	0.0035
σ_η^2	0.0213	0.0037
ρ	-0.8282	0.0410
Log-likelihood value $= -2586.7$		
SVLJ		
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
Log-likelihood value $= -2579.2$		

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

Nasdaq		
	ML estimate	Standard error
SV		
μ	0.7193	0.5488
ϕ	0.9973	0.0016
σ_η^2	0.0054	0..0156
Log-likelihood value = -3457.6		
SVL		
μ	0.4877	0.1834
ϕ	0.9942	0.0014
σ_η^2	0.0077	0.0016
ρ	-0.8291	0.0543
Log-likelihood value = -3429.3		
SVLJ		
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
Log-likelihood value = -3423.3		

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

Likelihood Ratio Test Statistic		
	SV vs SVL	SVL vs SVLJ
S&P 500	95.4**	5.4
FTSE 100	63.2**	1.2
Dow Jones	73.6**	15**
Nasdaq	56.6**	12**

Table 11: (*) indicates statistical significance at 5% critical level.

5 Conclusion:

References

Bates, D., (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Duetsch 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 (1999). Analysis of high dimensional multivariate stochastic volatility models. Unpublished paper, John M. Olin School of Business, Washington University, St. Louis. Revised June 2001.

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.(Unpublished working paper).

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 priceing 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.

Gweke, J., (1992). Evaluating the accuracy of sampling-based approaches to calculationof 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.

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.

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 appraoch. 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.

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

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

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 measurment time series. Biometrika 84, 653-67.

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

6 Appendix

Computation of moments of $f(\epsilon_t|J = 1; h_t, y_t)$:

We begin by noting that $f(\epsilon_t|J = 1; h_t, y_t) \propto f(y_t|J = 1, h_t, \epsilon_t)f(\epsilon_t)$,

$$\implies \quad f(\epsilon_t|J = 1; h_t, y_t) \propto N(\ y_t|\ \epsilon_t \exp(h_t/2); \sigma_J^2) \times N(\epsilon_t|0; 1)$$

On taking logarithmic transformation of the conditional density $f(\epsilon_t|J = 1; h_t, y_t)$ we obtain,

$$= const - \frac{1}{2}\frac{(y_t - \epsilon_t \exp(h_t/2))^2}{\sigma_J^2} - \frac{1}{2}\epsilon_t^2$$

We need to complete squares on this expression to obtain the general form,

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

In order to do this first collect the squared terms corresponding to $-\frac{1}{2}\epsilon_t^2$;

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

$$\Longrightarrow \sigma^2_{\epsilon_1} = \frac{\sigma^2_J}{\exp(h_t) + \sigma^2_J}$$

Next those corresponding to ϵ_t;

$$\frac{\upsilon_{\epsilon_1}}{\sigma^2_{\epsilon_1}} = \frac{y_t \exp(h_t/2)}{\sigma^2_J}$$

$$\Longrightarrow \upsilon_{\epsilon_1} = \frac{y_t \exp(h_t/2)}{\exp(h_t) + \sigma^2_J}$$

Therefore we establish that $f(\epsilon_t|J = 1; h_t, y_t) = N(\upsilon_{\epsilon_1}, \sigma^2_{\epsilon_1})$.

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

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

$$f(\epsilon^1_t|h^1_t, y_t) = \delta(y_t \exp(-h^1_t/2)).\Pr(J_t = 0|h^1_t, y_t) + N(\upsilon^1_{\epsilon_1}, \sigma^{2^1}_{\epsilon_1}).\Pr(J_t = 1|h^1_t, y_t).$$

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

$$F(\epsilon^1_t|h^1_t, y_t) = \Pr^J . \int^{\epsilon^1_t}_{-\infty} f(\epsilon_t|J_t = 1, h_t, y_t)d\epsilon_t \quad \text{for } \epsilon^1_t < x^*$$
$$F(\epsilon^1_t|h^1_t, y_t) = \Pr^J . \int^{x^*}_{-\infty} f(\epsilon^1_t|J_t = 1, h^1_t, y_t)d\epsilon^1_t + (1 - \Pr^J) \quad \text{for } \epsilon^1_t = x^*$$
$$F(\epsilon^1_t|h^1_t, y_t) = \Pr^J . \int^{x^*}_{-\infty} f(\epsilon^1_t|J_t = 1, h^1_t, y_t)d\epsilon^1_t + (1 - \Pr^J) + \Pr^J . \int^{+\infty}_{\epsilon^1_t} f(\epsilon^1_t|J_t = 1, h^1_t, y_t)d\epsilon^1_t \quad \text{for } \epsilon^1_t > x^*$$

As is evident from the form of $F(\epsilon^1_t|h^1_t, y_t)$, 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 $u_1 \leq \Phi(\frac{x^* - \upsilon^1_{\epsilon_1}}{\sigma^1_{\epsilon_1}}).\Pr^J$, we sample $\epsilon^1_t = \upsilon^1_{\epsilon_1} + \sigma^1_{\epsilon_1}\Phi^{-1}(\frac{u_1}{\Pr^J})$

If $\Phi(\frac{x^* - \upsilon^1_{\epsilon_1}}{\sigma^1_{\epsilon_1}}).\Pr^J < u_1 \leq \Phi(\frac{x^* - \upsilon^1_{\epsilon_1}}{\sigma^1_{\epsilon_1}}).\Pr^J + (1 - \Pr^J)$, we sample $\epsilon^1_t = y_t \exp(-h^1_t/2)$

If $u_1 > \Phi(\frac{x^* - \upsilon^1_{\epsilon_1}}{\sigma^1_{\epsilon_1}}).\Pr^J + (1 - \Pr^J)$, we sample $\epsilon^1_t = \upsilon^1_{\epsilon_1} + \sigma^1_{\epsilon_1}\Phi^{-1}(\frac{u_1 - (1 - \Pr^J)}{\Pr^J})$

$\Phi(.)$ denotes the standard normal distribution function. The above probablity integral tranform procedure is repeated for each of the generated uniforms $u_1, \ldots\ldots, u_M \backsim UID(0, 1)$. in order to obtain the sample $\epsilon^i_t \backsim f(\epsilon^i_t|h^i_t, y_t)$, $i = 1, \ldots\ldots, M$.

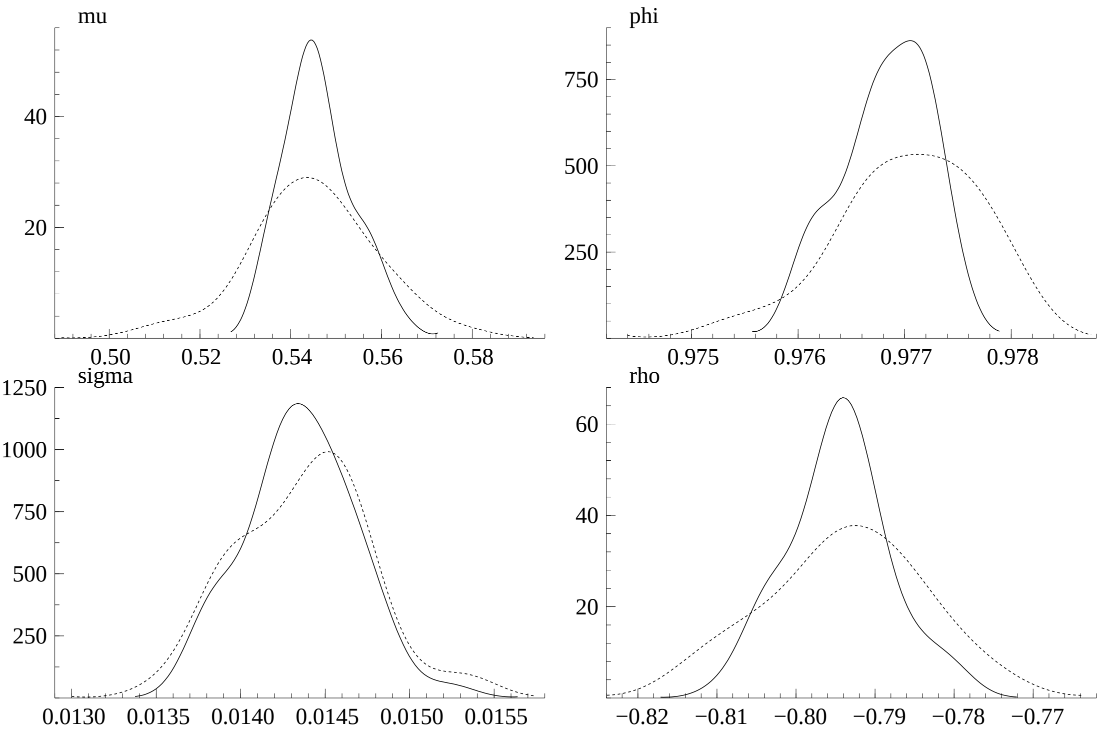

Figure 1: Fixed dataset. Dotted line:Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model; $T = 1000$ and $M = 300$. Solid line: Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model; $T = 1000$ and $M = 600$. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$.

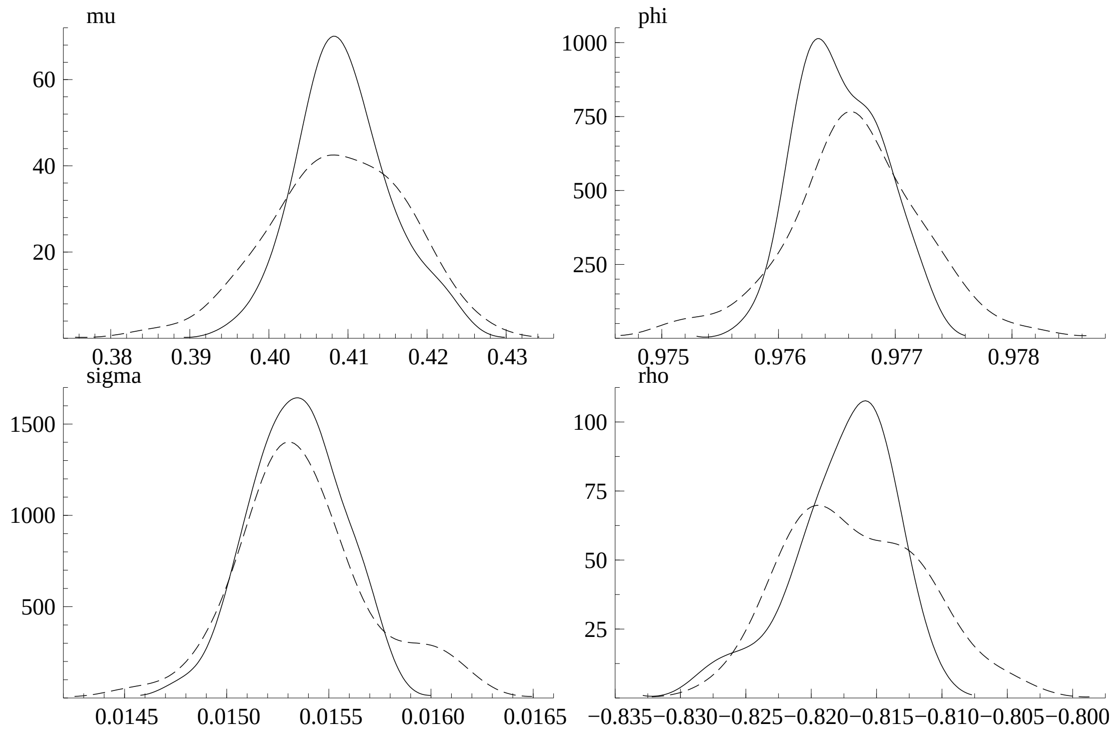

Figure 2: Fixed dataset. Dashed line:Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model; $T = 2000$ and $M = 300$. Solid line: Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with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

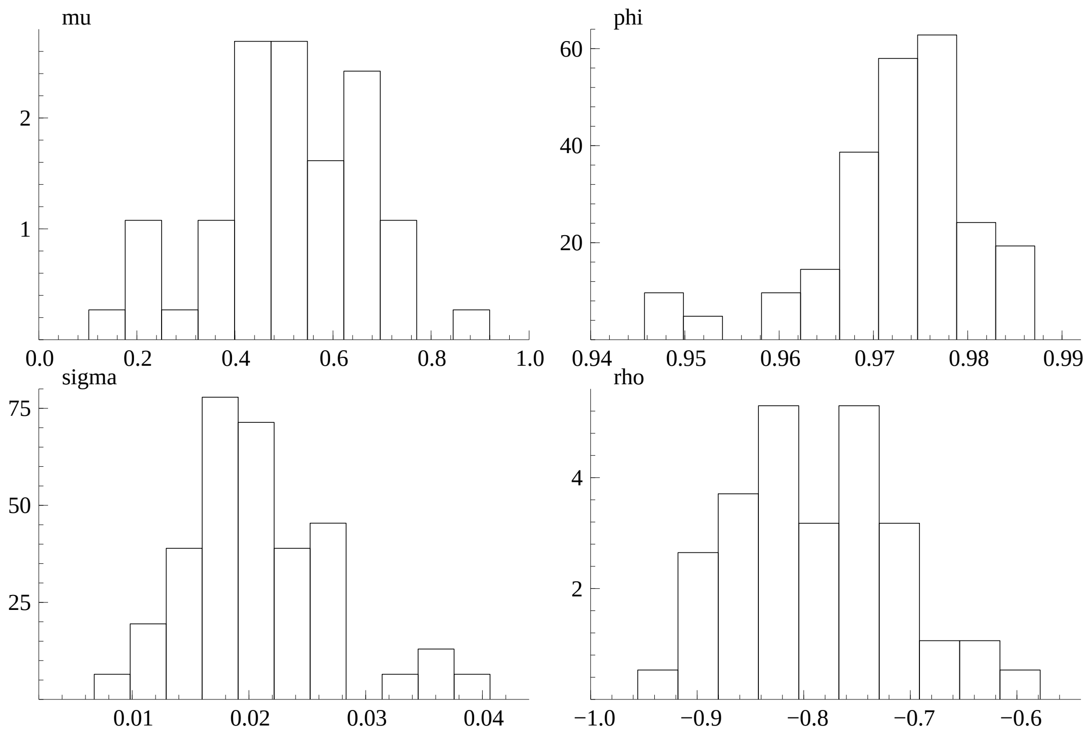

Figure 3: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$. $M = 200$ and $T = 1000$.

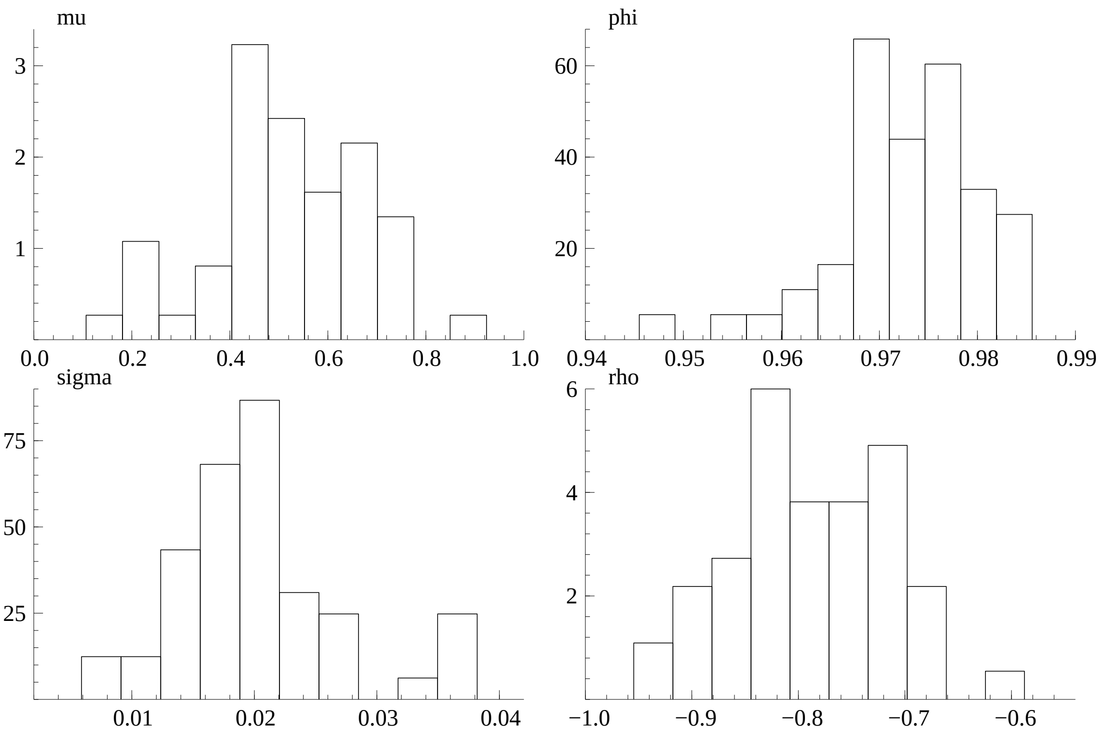

Figure 4: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$. $M = 500$ and $T = 1000$.

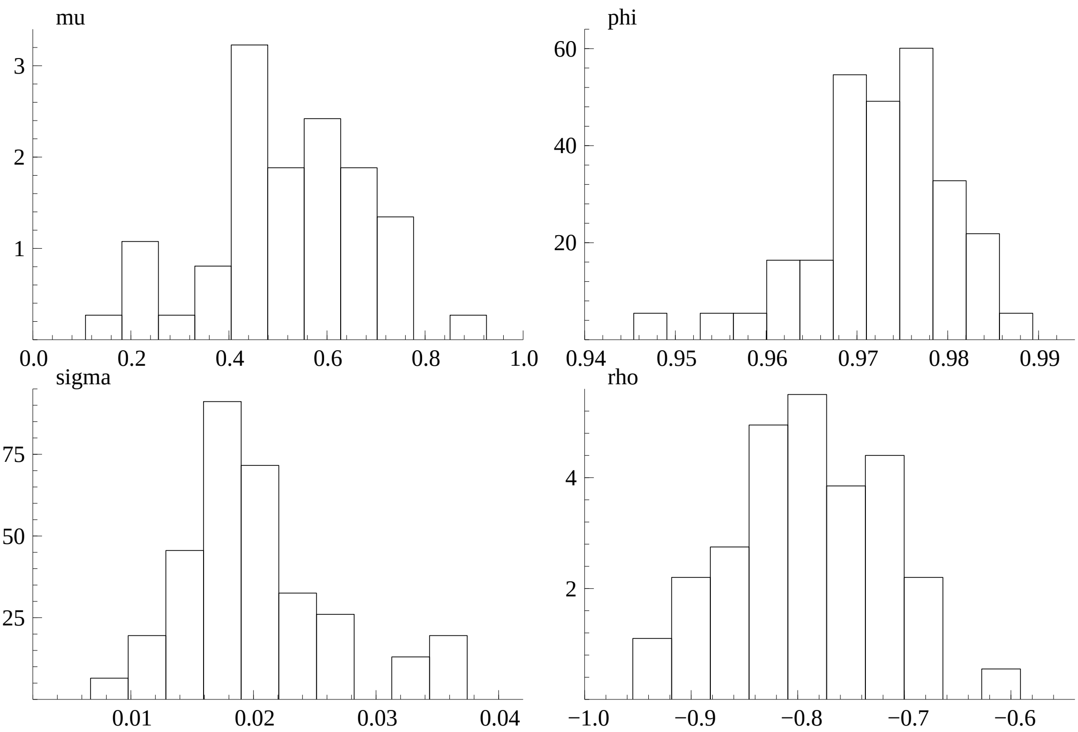

Figure 5: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho)$, for SV with leverage model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8$. $M = 3000$ and $T = 1000$.

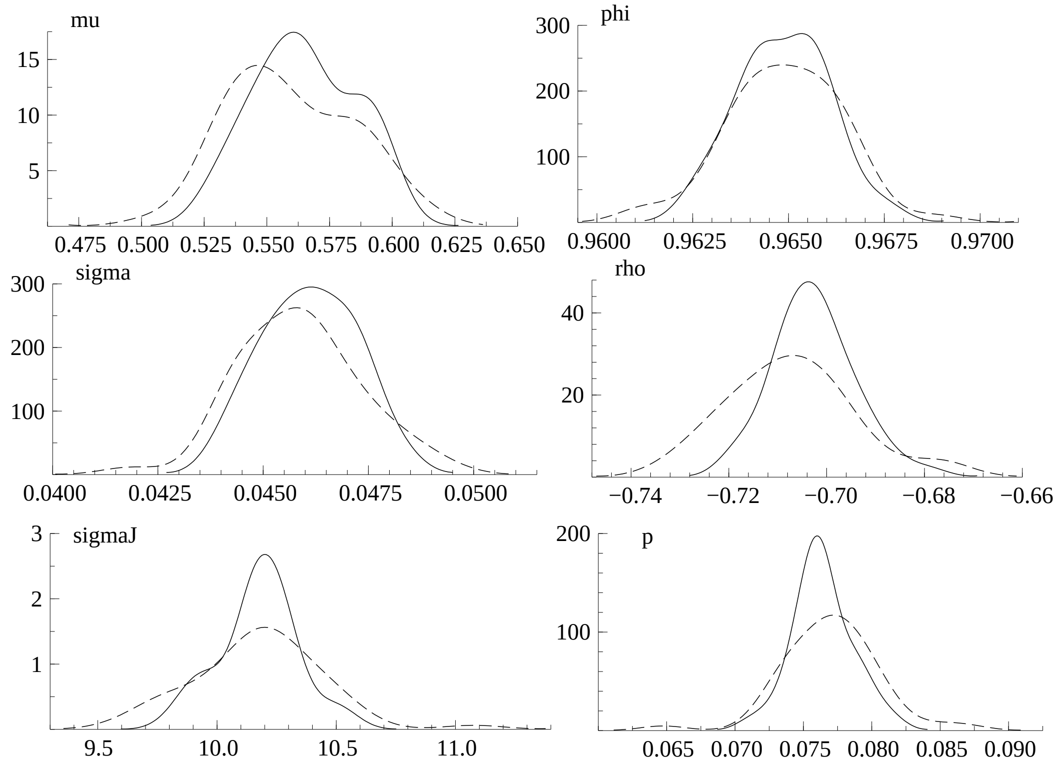

Figure 6: Fixed datasets. Dashed line:Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model; $T = 1000$ and $M = 300$. Solid line: Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model; $T = 1000$ 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 $p = 0.10$.

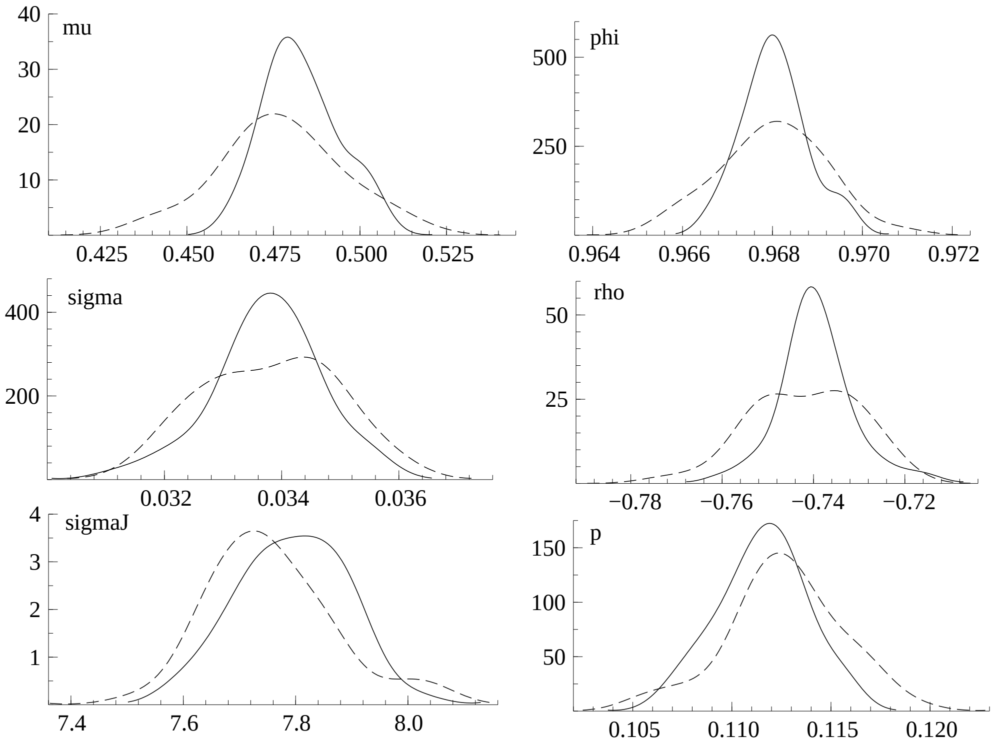

Figure 7: Fixed datasets. Dashed line:Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model; $T = 2000$ and $M = 300$. Solid line: Kernel density estimate of the ML estimator for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, 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 $p = 0.10$.

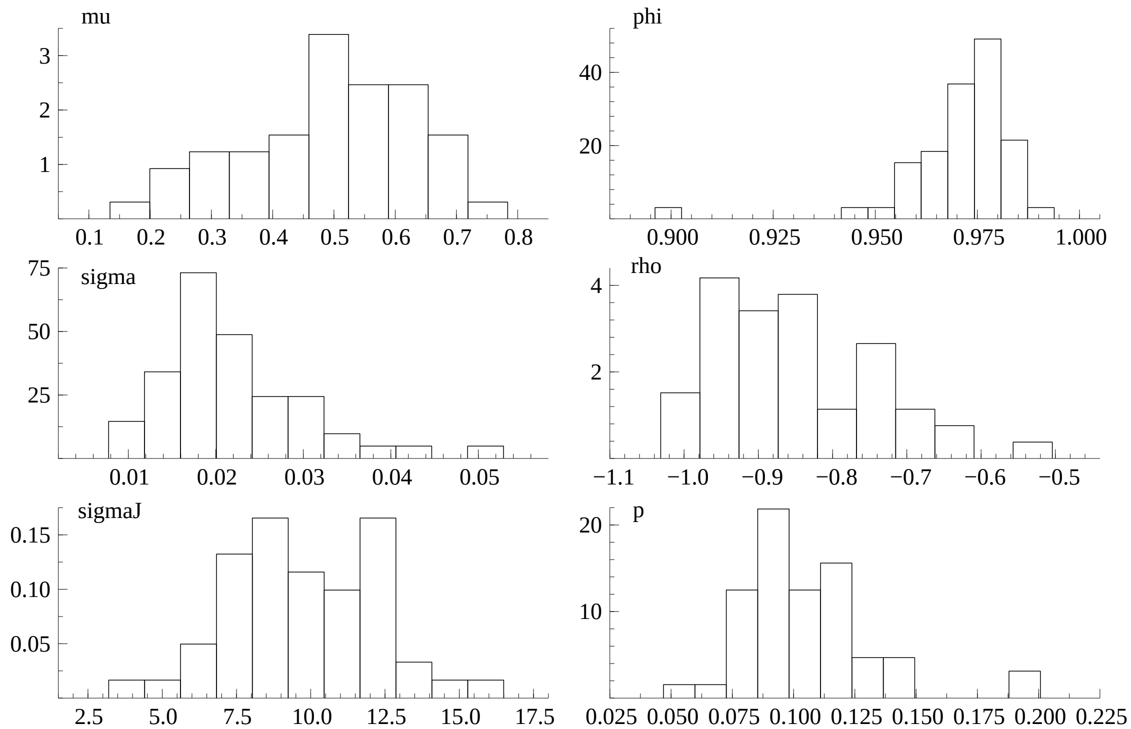

Figure 8: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.10$. $M = 200$ and $T = 1000$.

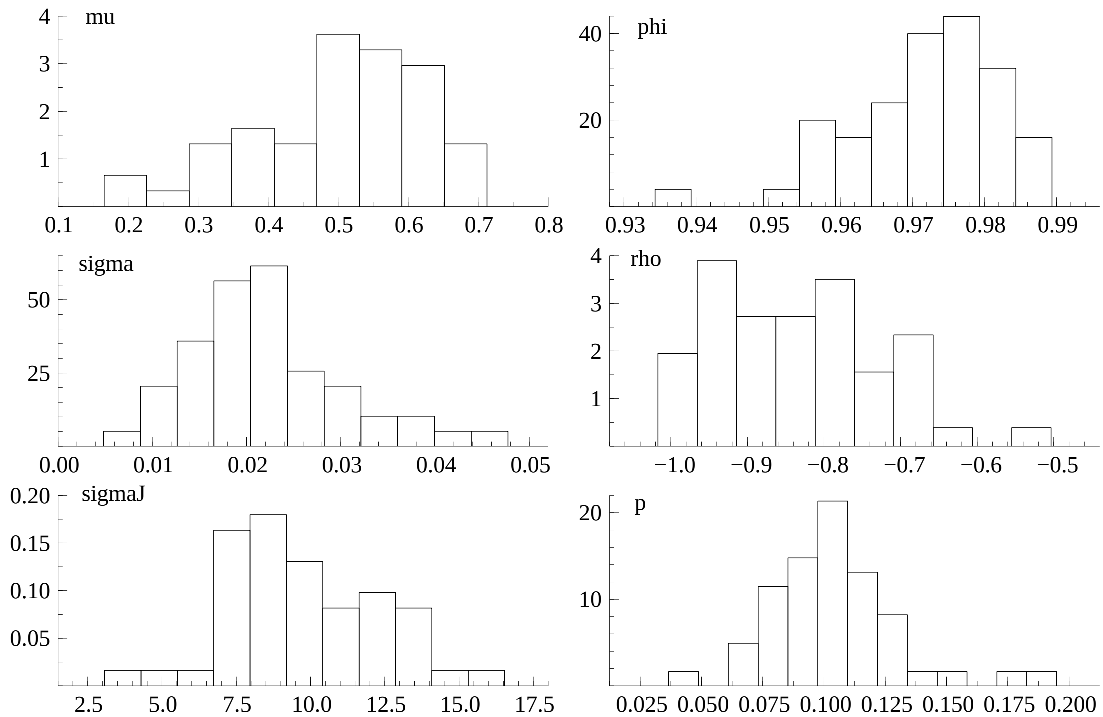

Figure 9: 50 different dataset. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.10$. $M = 500$ and $T = 1000$.

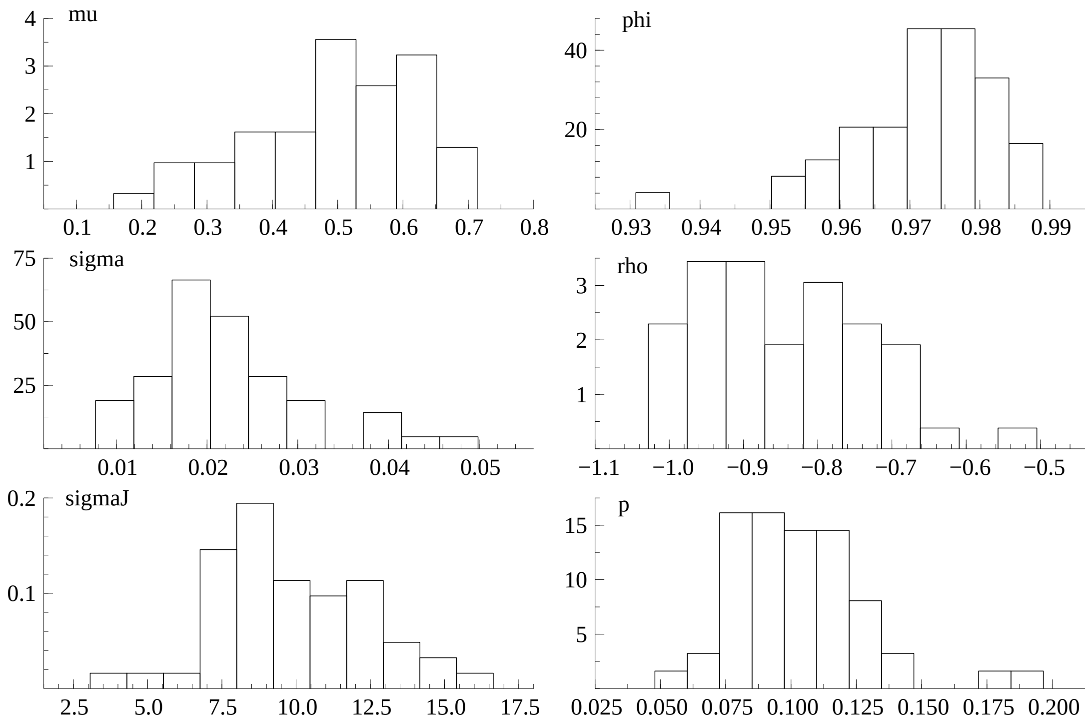

Figure 10: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.10$. $M = 900$ and $T = 1000$.

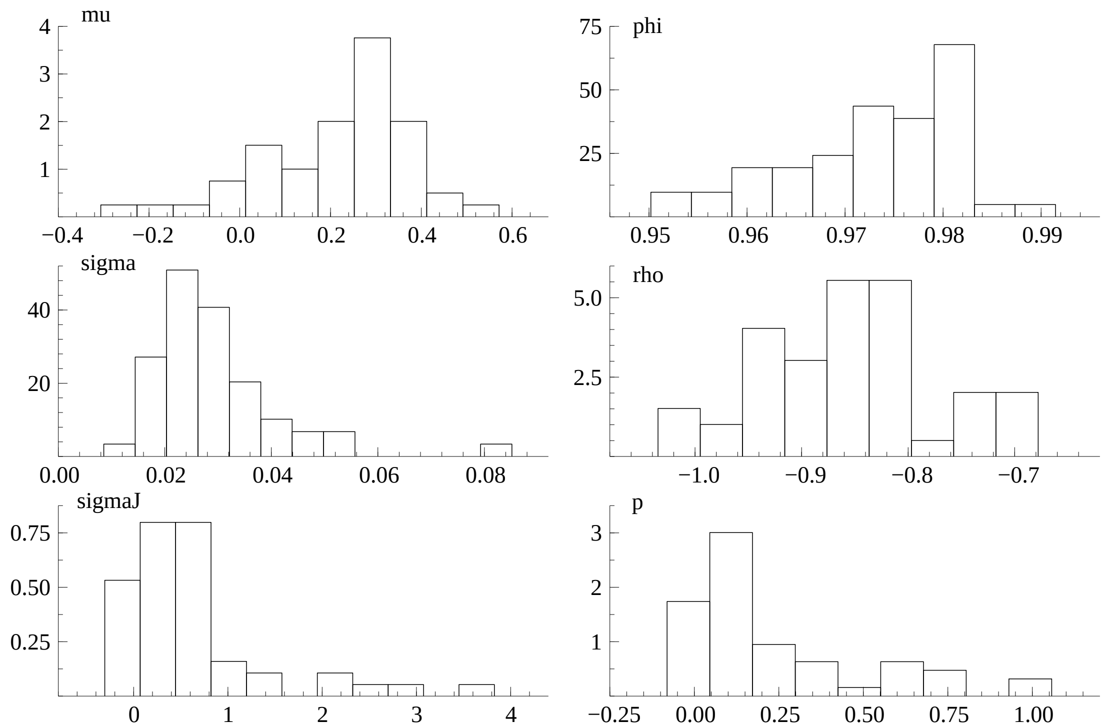

Figure 11: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.25, \phi = 0.975, \sigma_\eta^2 = 0.025$ and $\rho = -0.8, \sigma_J^2 = 0.5$ and $p = 0.10$. $M = 300$ and $T = 1000$.

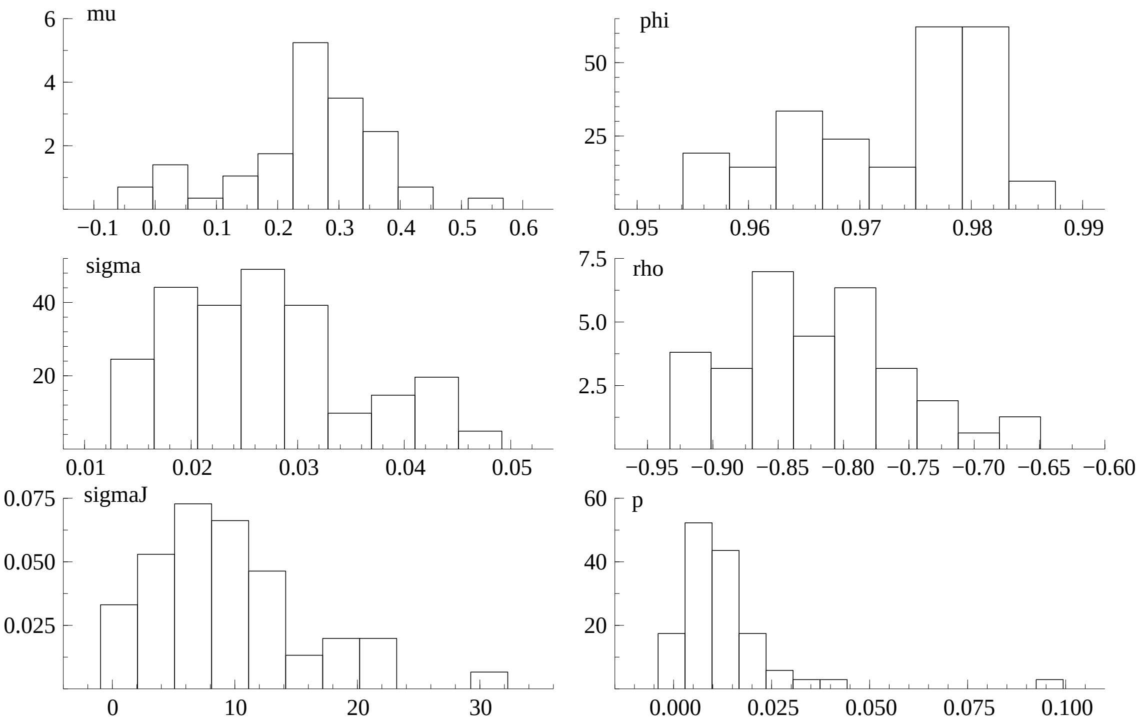

Figure 12: 50 different datasets. Histogram of the Monte Carlo samples of the ML estimates for $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, for SV with leverage and jumps model. True parameters, $\mu = 0.25, \phi = 0.975, \sigma_\eta^2 = 0.025$ and $\rho = -0.8, \sigma_J^2 = 10$ and $p = 0.01$. $M = 300$ and $T = 1000$.

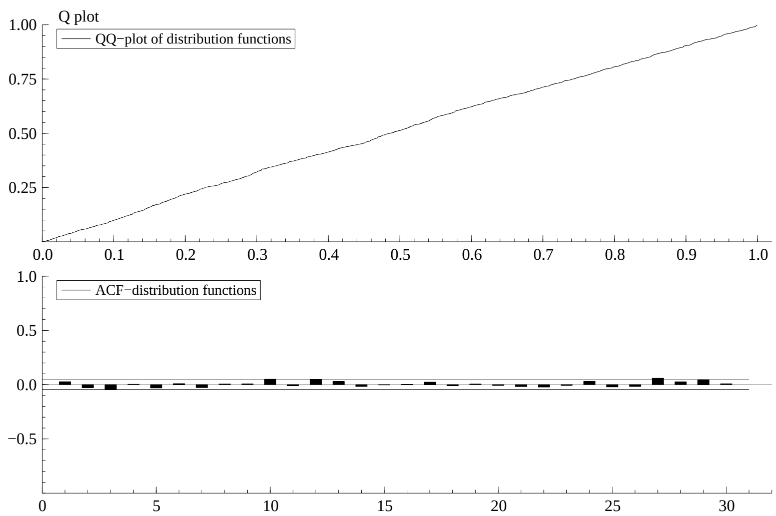

Figure 13: Filter diagnostics in the context of modelling stochastic volatility with leverage and jumps using daily S&P 500 returns over the period 16/05/1995 - 24/04/2003.(Above) QQ-plot of estimated distribution functions, $\widehat{u}_t^J$. (Below) Correlogram of $\widehat{u}_t^J$.$M = 500$.

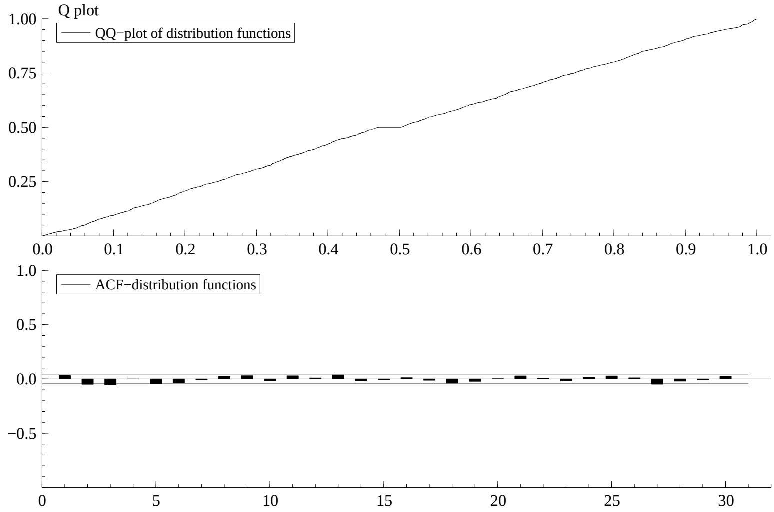

Figure 14: Filter diagnostics in the context of modelling stochastic volatility with leverage and jumps using daily FTSE 100 returns over the period 01/07/1996 - 01/03/2004.(Above) QQ-plot of estimated distribution functions, $\widehat{\mathrm{u}}_t^J$. (Below) Correlogram of $\widehat{\mathrm{u}}_t^J$. $M = 500$.

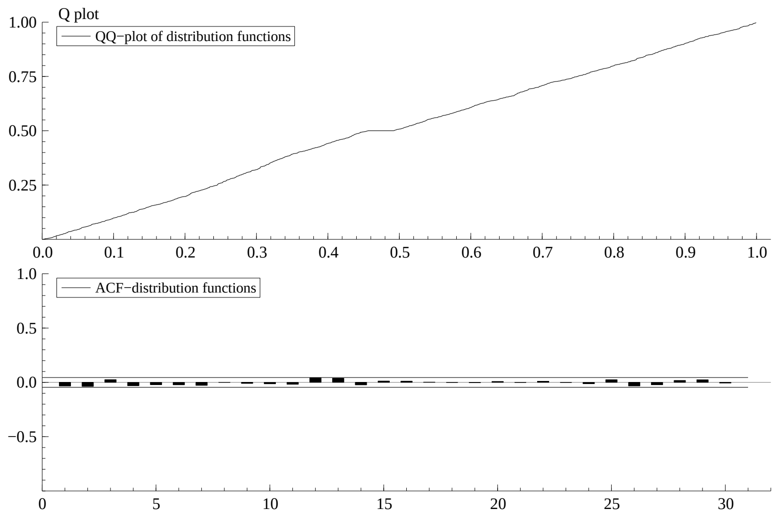

Figure 15: Filter diagnostics in the context of modelling stochastic volatility with leverage and jumps using daily Dow Jones Composite 65 Stock Av erage returns over the period 01/05/2000 - 31/12/2007.(Above) QQ-plot of estimated distribution functions, $\widehat{\mathrm{u}}_t^J$. (Below) Correlogram of $\widehat{\mathrm{u}}_t^J$. $M = 500$.

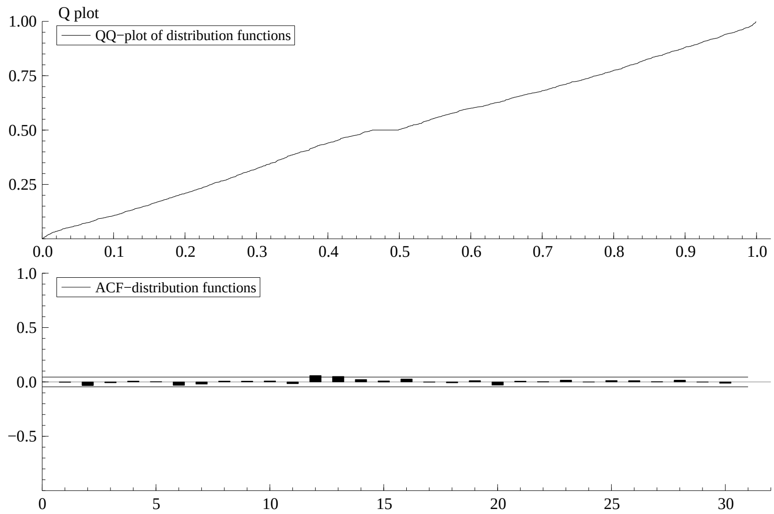

Figure 16: Filter diagnostics in the context of modelling stochastic volatility with leverage and jumps using daily Nasdaq Composite returns over the period 01/05/2000 - 31/12/2007.(Above) QQ-plot of estimated distribution functions, $\widehat{\mathrm{u}}_t^J$. (Below) Correlogram of $\widehat{\mathrm{u}}_t^J$. $M = 500$.