DOCUMENT
DE TRAVAIL
N° 318

MODELLING STOCHASTIC VOLATILITY WITH LEVERAGE AND JUMPS: A SIMULATED MAXIMUM LIKELIHOOD APPROACH VIA PARTICLE FILTERING

Sheheryar Malik and Michael K Pitt

February 2011

BANQUE DE FRANCE
EUROSYSTÈME

DIRECTION GÉNÉRALE DES ÉTUDES ET DES RELATIONS INTERNATIONALES

MODELLING STOCHASTIC VOLATILITY WITH LEVERAGE AND JUMPS: A SIMULATED MAXIMUM LIKELIHOOD APPROACH VIA PARTICLE FILTERING

Sheheryar Malik and Michael K Pitt

February 2011

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

Sheheryar Malik*and Michael K Pitt†‡

December 26, 2010

*Banque de France. Address: Banque de France, 31, Rue Croix des Petits Champs - 75049 Paris Cedex 01. E-mail: sheheryar.malik@banque-france.fr

†University of Warwick. Address: Department of Economics, University of Warwick, Coventry CV4 7AL. E-mail: m.pitt@warwick.ac.uk

‡We would like to thank Stephane Gregoir, Valentina Corradi and seminar participants and the University of Warwick, Banque de France and CREST for helpful suggestions and comments. All errors are our own. The views expressed in this paper do not necessarily reflect those of Banque de France or Eurosystem.

Abstract

In this paper we provide a unified methodology for conducting likelihood-based inference on the unknown parameters of a general class of discrete-time stochastic volatility (SV) models, characterized by both a leverage effect and jumps in returns. Given the nonlinear/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 standard 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 nontrivial task of model comparison. Furthermore, we introduce new volatility model, namely SV-GARCH which attempts to bridge the gap between GARCH and stochastic volatility specifications. In nesting the standard GARCH model as a special case, it has the attractive feature of inheriting the same unconditional properties of the standard GARCH model but being conditionally heavier-tailed; thus more robust to outliers. It is demonstrated how this model can be estimated using the described methodology. The technique is applied to daily returns data for S&P 500 stock price index for various spans. In assessing the relative performance of SV with leverage and jumps and nested specifications, we find strong evidence in favour of a including leverage effect and jumps when modelling stochastic volatility. Additionally, we find very encouraging results for SV-GARCH in terms of predictive ability which is comparable to the other models considered.
JEL classification: C01, C11, C14, C15, C32, E32

Key words: Stochastic volatility, Particle filter, Simulation, State space, Leverage effect, Jumps

Abstract

Nous fournissons dans ce papier une méthodologie unifiée pour réaliser une inférence en vraisemblance sur les paramètres inconnus d'une classe de modèles à volatilité stochastique (SV) en temps discret, caractérisée à la fois par un effet de levier et des sauts dans les rendements. Compte tenu de la forme espace-état non-linéaire / non-gaussienne, l'approximation de la vraisemblance pour les paramètres est réalisée au moyen d'un output généré par le filtre à particules. On s'assure que l'approximation de la vraisemblance est continue en tant que fonction des paramètres inconnus, ce qui permet l'utilisation d'algorithmes de maximisation standards de type Newton-Raphson. Notre approche est robuste, et efficace par rapport aux méthodes de Monte Carlo à chaînes de Markov utilisées dans ce type de contexte. De plus, elle fournit une base accessible pour entreprendre la tâche non triviale de comparaison de modèles. De plus, nous introduisons un nouveau modèle de volatilité, en particulier un SV-GARCH qui vise à unir des spécifications GARCH et de volatilité stochastique. En limitant le modèle GARCH standard à un cas isolé, notre approche présente la caractéristique intéressante d'hériter les propriétés d'inconditionnalité du modèle GARCH standard, mais tout en étant conditionnellement leptokurtique, et donc plus robuste aux points aberrants. On démontre comment un tel modèle peut être estimé en utilisant la méthodologie décrite. Cette technique est appliquée aux données relatives aux rendements quotidiens pour l'indice S&P 500 pour différentes périodes. En évaluant la performance relative du SV avec levier et sauts et spécifications isolées, nous trouvons des preuves fortes plaidant en faveur de l'inclusion d'un effet de levier et de sauts lorsqu'on modélise la volatilité stochastique. De plus, nous trouvons des résultats encourageants pour le SV-GARCH en termes de capacité prédictive, celle-ci étant comparable aux autres modèles considérés.
Classification JEL : C01, C11, C14, C15, C32, E32

Mots-clés : volatilité stochastique, filtre à particules, simulation, espace-état, effet de levier, sauts.

1 Introduction

The aim of this paper is to provide a unified methodology for conducting simulated maximum likelihood (SML) based inference using a particle filter on a general class of non-linear and non-Gaussian state-space models. Specifically, stochastic volatility models which take into account two well-known, stylised features of financial data i.e. leverage and jumps. In studying the relationship between volatility and asset price return, a leverage effect refers to the increase in future expected volatility following bad news. The underlying reasoning 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 return. In the finance literature, empirical evidence supportive of a leverage effect has been provided by Black (1976) and Christie (1982). Jumps, can basically be described as rare events; large, infrequent movements in returns which are an important feature of financial markets (see Merton, 1976). These have been widely documented to be important in characterizing the non-Gaussian tail behaviour of conditional distributions of returns.

To date, the state space models used in this context of modelling time varying conditional volatility have broadly fallen within two competing categories. These being, (i) Autoregressive Conditional Heteroscedasticity (ARCH) model, originally proposed by Engle (1982), various versions of which have been surveyed extensively by Bollerslev, Chou and Kroner (1992) and Shephard (1996); and (ii) Stochastic Volatility (SV) models as considered inter alia by Taylor (1994), Harvey, Ruiz and Shephard (1994) and Jacquier, Polson and Rossi (1994). Whereas the former category of models make conditional variance a deterministic function of past squared returns, SV models allow variance to evolve according to some latent stochastic process. These are natural discrete-time versions of continuous-time models on which much of modern financial economics, including generalizations of the Black-Scholes result in asset pricing has been developed; see for example Hull and White (1987). It can also be intuitively more appealing to consider information flow, especially at higher frequencies as being governed by a stochastic process. In a similar vein, the rapidly increasing usage of high frequency intraday data for constructing so-called, realized volatility measures is intimately linked to SV framework in financial economics (see Barndorf-Nielson and Shephard, 2002).

A major reason for the popularity of the ARCH family of models in describing the dynamics of financial market volatility is their tractable estimation. More specifically, given the deterministic dependence of the conditional variance on past observations, estimation and inference for ARCH-type models is greatly facilitated given that one-step ahead prediction densities are available in closed form. This enables the likelihood of parameters to be explicitly written via prediction decomposition (see Harvey, 1993 and Kim, Shephard and Chib, 1998). Estimation of SV models is however greatly complicated by the stochastic evolution of volatility which implies that, unlike ARCH counterparts, the likelihood here can not be obtained in closed form. 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.

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). The approach of estimating SV models with student-t errors has also been employed by, for example, Chib, Nedari and Shephard (2006) and Sandmann and Koopman (1998) inorder to capture conditionally heavier-tailed behaviour. For the same purposes, an alternative approach employed by Durham (2007) is to use a mixture of Gaussians for the measurement equation disturbance. His paper uses a simulated maximum likelihood approach to conduct inference based upon a Laplace approximation as the proposal density.

In this paper we add to the literature in two ways. We provide a unified and general methodology for carrying out simulated maximum likelihood (SML) estimation via particle filtering of the parameters of an SV model which incorporates both leverage and jumps (SVLJ). The approach is simple to implement and relatively fast on a standard PC or laptop. We demonstrate the speed and robustness of the methodology by examining simulated data arising from the specified data generating process. The generality of the method is highlighted by the fact that the standard SV or SV with leverage (SVL) specifications are nested within the SV with leverage and jumps specification (SVLJ), and can thus straightforwardly be recovered imposing restrictions on the latter complete model. We also show how diagnostics, filtered volatilities, quantile plots of filtered volatilities and filtered probability of jumps.The latter enables us to identify jump times. Furthermore, we introduce a new volatility model, one that is also characterized by a non-linear/non-Gaussian state space form. The essential point is that the proposed hybrid model, namely SV-GARCH henceforth, attempts to bridge elements of SV and GARCH specifications. This model nests the standard GARCH model as a special case. It has the attractive feature of inheriting the same extensively well-documented unconditional properties of the standard GARCH model but being conditionally heavier tailed. At the same time being no more complicated than GARCH, i.e. with the addition of just one more parameter. It is again demonstrated how SML via particle filtering can be employed to estimate this model. Its robustness to jumps/outliers realtive to GARCH is demonstrated and we also investigate its performace relative to the other three stochastic volatility models mentioned which have a comparatively deeper theoretical underpinning in the finanical economics literature.

The structure of the paper is as follows. In Section 2 we describe the standard SV model, the SV with leverage model and the SV with leverage with jumps model. We also introduce the SV-GARCH model. 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. This methodology of course allows for no jumps or leverage as special cases. We also describe the relevant diagnostic tests for the general case. Section 4 provides results for simulation experiments testing estimator performance in the case of both SVLJ and SV-GARCH. Section 5 provides empirical examples using daily returns data for S&P500. In Section 6 we conclude.

2 Volatility Models

2·1 *Stochastic Volatility*

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

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

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^2 is the volatility of log-volatility and ϕ is the persistence parameter. Typically we would impose that $|\phi| < 1$ so that we have a stationary process with the initial condition that,

$$h_1 \backsim N\{0, \sigma_\eta^2/(1-\phi^2)\}.$$

This is in fact the Euler-Maruyama discretization of the continuous-time Orstein-Uhlenbeck (log-OU) process. Within the financial econometrics literature, this model is seen as a generalization of the Black-Scholes model for option pricing that allows for volatility clustering in returns.

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. We retain the model as,[1]

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

where we now allow for the disturbances to be correlated (see e.g. Heston,1993) as,

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

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. Typically we would impose that $|\phi| < 1$ so that we have a stationary process with the initial condition that,

$$h_1 \backsim N\{0, \sigma_\eta^2/(1-\phi^2)\}.$$

Due to the timing of the, typically negative correlation in disturbances, the unconditional distribution is not affected by the leverage term (see Yu, 2005). For example, the unconditional skewness of the returns, y_t, remains zero. We note that the disturbances are conditionally Gaussian and so 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 then be reformulated as,

$$h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \rho \epsilon_t + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t. \tag{2·3}$$

[1]The case of SV with leverage has recently been considered by Christofferesen, Jacobs and Minouni (2010). 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, introduced by Pitt (2002) for the purpose of parameter estimation. The methodology of the smooth particle filter has also been recently applied by Duan and Fulop (2009) in the context of models for credit risk.

By substituting, $\epsilon_t = y_t \exp(-h_t/2)$ into (2·3), the model adopts the following Gaussian nonlinear state-space form, where the parameter ρ measures the leverage effect,

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) && (2\cdot4) \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \rho \, y_t \exp(-h_t/2) + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t. && (2\cdot5) \end{aligned}$$

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 $h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t$.

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,

$$\begin{aligned} y_t &= \epsilon_t \exp(h_t/2) + J_t \varpi_t && (2\cdot6) \\ h_{t+1} &= \mu(1-\phi) + \phi h_t + \sigma_\eta \eta_t, \quad t = 1, ..., T \end{aligned}$$

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 $f(\eta_t|\epsilon_t) = N(\rho\epsilon_t; 1-\rho^2)$.[2]

2·4 *SV-GARCH*

In the spirit of studying heavier tailed voltility models we propose a new propose a new nonlinear model for volatility, the SV-GARCH. If we denote the observed return y_t, and lagged conditional variance $\sigma_t^2 \equiv v_t$ then the Generalized ARCH (GARCH[3]) model as put forth by, Bollerslev (1986) and Taylor (1986) can be written as;

$$\begin{aligned} y_t &= \sqrt{v_t}\epsilon_t && (2\cdot7) \\ v_{t+1} &= \gamma + \alpha v_t + \beta y_t^2, \; t = 1, ..., T, \end{aligned}$$

where $\epsilon_t \backsim N(0,1)$. Parameter restrictions $\gamma > 0$ $\alpha \geq 0$, $\beta \geq 0$ are set in order to ensure conditional variances are uniformly positive and for the existence of stationarity of the process

[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.

[3] By including the lagged conditional variance, GARCH improves greatly upon the original ARCH specification in terms of being more parsimonious. Typically in empricical applications ARCH was found to require a relatively long lag lenght for squared returns in order to adequately caputure the behaviour of conditional variance (see Bollerslev,1986).

we require the condition $\alpha + \beta < 1$ to hold. The initial condition is typically given by the unconditional expectation of the variance process,

$$v_1 = \gamma/(1-\alpha-\beta).$$

The GARCH specification implies that conditional variance depends on the previous squared return, i.e. $y_t^2 = v_t \epsilon_t^2$. Let us define a disturbance term ζ_t as,

$$\zeta_t = \varphi \epsilon_t + \sqrt{(1-\varphi^2)}\xi_t \text{ where } \xi_t \backsim N(0,1). \tag{2·8}$$

By replacing ϵ_t^2 by ζ_t^2 in the GARCH specification yields the non-linear transition function,

$$\begin{aligned} v_{t+1} &= \gamma + \alpha v_t + \beta v_t \zeta_t^2 \\ &= \gamma + \alpha v_t + \beta v_t \{\varphi \epsilon_t + \sqrt{(1-\varphi^2)}\xi_t\}^2. \end{aligned} \tag{2·9}$$

Here, as in GARCH, parameter restrictions $\gamma > 0$ $\alpha \geq 0$, $\beta \geq 0, \alpha + \beta < 1$ apply and additionally $0 \leq \varphi \leq 1$. In the case of $\varphi = 1$ the model collapses to the standard GARCH specification with linear transition function as in (2·7); whereas $\varphi = 0$ yields a specification which is 'stochastic' in nature, in that the feedback effect via the observed standardized return $\epsilon_t = y_t v_t^{-\frac{1}{2}}$ is eliminated. Hence under boundary restrictions for φ the SV-GARCH model is given by,

$$\begin{aligned} y_t &= \sqrt{v_t}\epsilon_t \\ v_{t+1} &= \gamma + \alpha v_t + \beta v_t \epsilon_t^2 \text{ if } \varphi = 1 \text{ or} \\ &= \gamma + \alpha v_t + \beta v_t \xi_t^2 \text{ if } \varphi = 0, \end{aligned} \tag{2·10}$$

for $t = 1, ..., T$. The specification is very flexible and the process can be potentially governed by values as, $\varphi \longrightarrow 1$ or $\varphi \longrightarrow 0$.

The SV-GARCH model has some attractive features in that it inherits all the same unconditional properties of the well established standard GARCH model, i.e. skewness, kurtosis and autocorrelation structure (see Bollerslev, 1986); but the stochastic nature of the transition equation (2.9) renders the conditional distribution of returns a mixture,

$$f(y_{t+1}|Y_t) = \int f(y_{t+1}|v_{t+1}) f(v_{t+1}|Y_t) dv_{t+1}. \tag{2·11}$$

The implication of this is that the model displays conditional leptokurtosis, so long as $\varphi \neq 1$.In the standard GARCH the predictive density $f(v_{t+1}|Y_t)$ would be (degenerate) dirac delta. This suggests that in principle the SV-GARCH model is more robust to jumps/outliers relative to conditionally Gaussian counterparts.[4]Authors such as Bollerslev (1987) and Nelson (1991) have

[4]It should be montioned that our model is in contrast to the GARCH-SV model of Franses et al. (2008) in which the standard GARCH specification is augmented by a moving average term in order to capture SV properties. The model in Frances et al. (2008), with slight abuse of notation is given as,

$$\begin{aligned} y_t &= \sqrt{v_t}\epsilon_t \\ v_t &= \gamma + \alpha v_{t-1} + \beta y_{t-1}^2 + \mu_t + \theta \mu_{t-1} \\ \epsilon_t &\backsim N(0,1) \text{ and } \mu_t \backsim N(0, \sigma_\mu^2). \end{aligned}$$

Here $\gamma > 0$ $\alpha \geq 0$, $\beta \geq 0$, $\alpha + \beta < 1$ and $-\alpha \leq \theta < 1$. In noting the estimation difficulties associated with models of stochastic volatility they need to restrict the parameters, $-\alpha = \theta$. Essentially, their test between GARCH and stochastic volatility specifications is reduced to testing whether $\sigma_\mu^2 = 0$, when the model collapses to a GARCH, as opposed to $\sigma_\mu^2 > 0$, i.e. when stochastic moving average components are maintained.

assumed heavier-tailed distributions such as standardized Student's t (GARCH-t) and Generalized Error distributions (GED) respectively in order to provide robustness to outliers. The advantage of employing the SV-GARCH approach in incorporating heavier-tailed behaviour is that, unlike GARCH-t and GED, which postulate (fixed) heavier-tailed unconditional (and conditional) distribution for the returns process, this formulation with a latent stochastic process driving volatility is far less dependent to possible mispecification brought about by assuming a fixed distribution. Essentially, the path of SV-GARCH volatility can thus adjust after encountering an outlier, since in essence it remains centred on the GARCH volatility path in normal times. This feature also enables us to quantify the contribution to volatility of deviations brought about by abnormal (jumps) returns. Furthermore, the shochastic volatility nature of SV-GARCH allows it to be extended to a continuous-time setting.[5]

3 Particle Filter Estimation

This paper is concerned with evaluation of state-space models via the 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 filtering 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 filtering density or weighted samples from under this density.

There has been considerable work done on the development of simulation based methods to perform filtering for nonlinear Gaussian state space models. The particle filter was first introduced by Gordon, Salmond and Smith (1993). Additional references include Kitagawa (1996) who introduced ideas of stratification and Pitt and Shephard (1999) who introduce the auxiliary particle filter. A review is provided by Doucet et al. (2000). Most of the literature revolves around on-line filtering of the states with less work done in the parameter estimation within this framework; see Liu and West (2000), Pitt (2002) and Polson, Stroud and Muller (2008).

The great advantage of particle filtering is that in many implementations we simply have to simulate forward in time from our data generating process for the state (prior to a multinomial sampling step). This is typically straightforward whereas Bayesian imputation via Markov chain Monte Carlo (MCMC) is usually much more complicated. Like MCMC however, the particle filter can effectively reduce the dimension over which we integrate. This is in marked contrast to SML via importance sampling, see for example Geweke (1999), where the dimension of integration is over the entire length of the time series.

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. The structure of this section is as follows. In Section 3·1 we describe the implementation of the standard particle filter of Gordon et al. (1993) for general latent time series models. In Section 3·2 we show how the likelihood may be estimated using particle filters for general models. The adjustments of Pitt (2002) are introduced which allow for continuity in the estimated likelihood surface. Some of the details will be relegated to the Appendix. In Section 3·3, the implementation is outlined for our class of volatility models with jumps and leverage. The implementation for no jumps or no leverage arises quite straightforwardly from this general

[5] We refer the reader to Nelson (1990) and Corradi (2000) for futher discussion on the continous-time limit of GARCH processes.

formulation. We discuss model diagnostics in Section 3·4, using the output from the particle filter. In Section 3·5 the estimation of Bayes factors for model comparison is outlined.

3·1 *Preliminaries*

For particle filtering we will assume a known measurement density $f(y_t|h_t)$ and the ability to simulate from the Markov 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', $h_t^1, \ldots.., h_t^M$ with associated discrete probability masses $\lambda_t^1, \ldots., \lambda_t^M$, 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,

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

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 $h_0^i \sim f(h_0)$, $i = 1, \ldots., M$ 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* $i = 1, \ldots, M$.

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* $L(x)G(x)$ *and that we have samples* $x^i \backsim G(x), i = 1, \ldots., M$. *If* $L(x)$ *is a known function then 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 proof may be found in Smith and Gelfand (1993). Essentially our required density is the "empirical filtering density",

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

for each time step. We are sampling, in **Step 1** of **Algorithm : SIR**, from the standard mixture, the summation above in (3·2). This mixture plays the role of $G(x)$ in Theorem 3.1. In **Step 2** of **Algorithm : SIR** we reweight each of these sample with respect to the normalised version of $f(y_{t+1}|h_{t+1}^i)$, which plays the role of $L(x)$ in Therorem 3.1. We then apply resampling in **Step 3** of Algorithm:SIR to achieve an equally weighted sample. Various approaches for approximating (3·2) as closely as possible are considered via the auxiliary particle filter approach of Pitt and Shephard (1999).

It is worth noting that we need to know $f(y_{t+1}|h_{t+1})$ only up to 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**. The latter estimator is typically more efficient for the estimation of moments. However, for other quantities (for instance the estimation of quantiles) the equally weighted sample from **Step 3** may be more convenient to use. Next we look at how this simple SIR particle filter framework can be 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;

$$\log L(\theta) = \log f(y_{1,\ldots,}.y_T|\theta) = \sum\nolimits_{t=1}^T \log f(y_{t+1}|\theta; Y_t), \tag{3·3}$$

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

$$f(y_{t+1}|\theta; Y_t) = \int f(y_{t+1}|h_{t+1};\theta)f(h_{t+1}|Y_t;\theta)dh_{t+1}. \tag{3·4}$$

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. In **Algorithm : SIR** we may estimate the predictive density (3·4) unbiasedly as,

$$\widehat{f}(y_{t+1}|\theta; Y_t) = \frac{1}{M}\sum\nolimits_{k=1}^M f(y_{t+1}|\widetilde{h}_{t+1}^k;\theta) = \frac{1}{M}\sum\nolimits_{k=1}^M \omega_{t+1}^k.$$

The terms ω_{t+1}^k are simply the unormalised weights computed in **Step 2** of **Algorithm : SIR**. The estimation of the likelihood is therefore a by-product of a single run of the particle filter. The estimator for the log-likelihood would therefore be[6]

$$\log \widehat{L}_M(\theta) = \sum\nolimits_{t=1}^T \log \widehat{f}(y_t|\theta; Y_{t-1}) = \sum\nolimits_{t=1}^T \log\left(\frac{1}{M}\sum\nolimits_{k=1}^M \omega_t^k\right). \tag{3·5}$$

There are various methods for implementing the resampling step, **Step 3** of the SIR algorithm. Instead of using multinomial sampling for sampling the indices we use a continuous (in the states) resampling procedure with stratification. The reason for this is as follows.

[6]**Bias correction:** We use a standard first order correction for the bias in the estimation of the log of the prediction density, $\log f(y_{t+1}|\theta; Y_t)$. This is detailed in Appendix A.

As noted in Pitt (2002), 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. Following Pitt (2002), we have chosen the distribution function $G(x) = x$ corresponding to a Uniform distribution although other choices are possible. Importantly as $M \to \infty$, $\widetilde{F}(h_{t+1}) \to \widehat{F}(h_{t+1}) \to F(h_{t+1}|Y_t)$. In practice the difference between $\widetilde{F}(h_{t+1})$ and $\widehat{F}(h_{t+1})$ becomes negligible for moderate M. It is also straightforward and quick to invert this function. The computational overhead is in principle $O(M \times \log M)$ due to the necessary sorting of the sampled h_{t+1} though in practice we found this term to be largely irrelevant. This is because the sample is close to being sorted prior to being sorted. As a consequence, taking the length of the time series into account, the computational burden is in practice $O(T \times M)$. We fix the random numbers (or equivalently the random number seeds) used in **Step 1** of **Algorithm : SIR**. This fixes the innovations we propagate through the state equation. We also keep the uniforms associated with the stratified bootstrap method fixed. The method of continuous resampling, from the distribution function $\widetilde{F}(h_{t+1})$ is described in further detail in **Appendix A.**

This can be related to the work of Gourieroux and Monfort (1990), summarised in Gourieroux and Monfort (1996, Chapter 3). Consider optimising the estimator of the log-likelihood,

$$\begin{aligned}\widehat{\theta}_{SML} &= \arg\max\left\{\log \widehat{L}_M(\theta)\right\} \\ &= \arg\max\left\{\sum\nolimits_{t=1}^{T} \log\left(\frac{1}{M}\sum\nolimits_{k=1}^{M} \omega_t^k\right)\right\}.\end{aligned}$$

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[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

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

likelihood function $L(\theta)$, such that $\widehat{L}_M(\theta) \stackrel{a.s.}{\to} L(\theta)$ as $M \to \infty$ and $E[\widehat{L}_M(\theta)] = L(\theta)$. The second of these results is surprising and important as the estimator is unbiased regardless of the particle filter size M. The resulting simulated maximum likelihood (SML) 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.[8] The following central limit theorem holds,

$$\sqrt{T}(\widehat{\theta}_{SML} - \theta_0) \Longrightarrow N(0, I^{-1}(\theta_0)),$$

where $I(\theta_0)$ is the expected information matrix at the true parameter value θ_0. The practical implications of this result are that we can maximize the estimated likelihood from the particle filter as long as M increases at a rate of at least $\sqrt{T}$ (for instance at rate $\sqrt{T}\log T$).

For the corresponding estimator of the log-likelihood Chopin (2004) and Del Moral (2004, Section 9.4) show that

$$\sqrt{M}\{\log \widehat{L}_M(\theta) - \log L(\theta)\} \Longrightarrow N(0, \sigma^2_{SMC,T}), \tag{3·6}$$

where the sequential monte carlo (SMC) variance $\sigma^2_{SMC,T}$ also depends upon the length of the time series, T. It is useful to contrast this with standard importance sampling, e.g. Danielsson and Richard (1993) and Shephard and Pitt (1997). For standard importance sampling there is no resampling step as there is in particle filters; the weights are propagated until the final time. For importance sampling (IS) estimators of the log-likelihood, we obtain a similar expression to (3·6) but with the variance being $\sigma^2_{IS,T}$. In particular Chopin (2004) demonstrates that for the particle filter the variance $\sigma^2_{SMC,T}$ in (3·6) is upper bounded uniformly in time. For the importance sampler (where no resampling is performed) the corresponding variance $\sigma^2_{IS,T}$ goes to infinity.

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 standard 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 using the expression (2·5).

Let us now consider the SV with leverage and jumps model in (2·6). **Step 1** is now slightly more complicated and nests two additional steps (**1a** and **1b**) which will be described below. We alter the data generating process (DGP) into an equivalent system.

We may alter the DGP of (2·6) by having the return innovation ϵ_t arising from the density $f(\epsilon_t|h_t, y_t)$, after h_t and y_t have occurred. We can then think of this innovation ϵ_t the feeding

[8] 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).

into the state equation $f(h_{t+1}|h_t; y_t; \epsilon_t)$,

$$h_{t+1} = \mu(1-\phi) + \phi h_t + \sigma_\eta \rho \epsilon_t + \sigma_\eta \sqrt{(1-\rho^2)}\xi_t, \quad (3{\cdot}7)$$

where ξ_t is again independent standrad Gaussian, as for the leverage alone model. Having recast our DGP in this, equivalent, form we now have the two sub-steps for **Step 1** of **Algorithm : SIR**.

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

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

We have that the density of **Step (1a)** is a mixture of the form,

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

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,

$$f(\epsilon_t|h_t, y_t) = \delta \{y_t \exp(-h_t/2) - \epsilon_t\} \times \Pr(J_t = 0|h_t, y_t) + N(\epsilon_t|\upsilon_{\epsilon_1}, \sigma_{\epsilon_1}^2) \times \Pr(J_t = 1|h_t, y_t). \quad (3{\cdot}8)$$

It is evident that this density is characterized by point mass at a unique point, $y_t \exp(-h_t/2)$, allowing for no jump, and continuity elsewhere, allowing for a jump. The derivation of this mixture in addition to computation of the moments υ_{ϵ_1}, $\sigma_{\epsilon_1}^2$, probability $\Pr(J_t = 1|h_t, y_t)$ and the associated distribution function is detailed in the **Appendix B**. This distribution function can be inverted easily allowing simple continuous simulation by using fixed uniform random variates. The simulation from the density $f(h_{t+1}|h_t^i; y_t; \epsilon_t^i)$ for **Step (1b)** may be performed straightforwardly by applying (3·7).

The non-normalized weights for **Step 2** in the SIR algorithm are of the form,

$$\begin{aligned} f(y_{t+1}|\widetilde{h}_{t+1}^i, \sigma_J^2) &= (1-p)\left\{2\pi \exp\left(\widetilde{h}_{t+1}^i\right)\right\}^{-\frac{1}{2}} \exp\left(-\frac{1}{2} y_{t+1}^2 \exp\left(-\widetilde{h}_{t+1}^i\right)\right) \\ &\quad + p\left\{2\pi\left[\exp(\widetilde{h}_{t+1}^i) + \sigma_J^2\right]\right\}^{-\frac{1}{2}} \exp\left(-\frac{1}{2} y_{t+1}^2 \left[\exp(\widetilde{h}_{t+1}^i) + \sigma_J^2\right]^{-1}\right), \end{aligned}$$

for $i = 1, ..., M$.

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 filtered 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,

$$\widehat{\Pr}(J_t = 1|Y_{t-1}) = \frac{1}{M} \sum_{i=1}^{M} \Pr(J_t = 1|y_t, h_t^i). \quad (3{\cdot}9)$$

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,

$$u_t = F(y_t|Y_{t-1}) = \int F(y_t|h_t)f(h_t|Y_{t-1})dh_t.$$

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

$$\widehat{u}_t = (1-p)\left\{\frac{1}{M}\sum_{i=1}^{M}\Phi\left(y_t \exp\left(-\widetilde{h}_t^i/2\right)\right)\right\} + p\ \left\{\frac{1}{M}\sum_{i=1}^{M}\Phi\left(y_t\left[\exp(\widetilde{h}_{t+1}^i)+\sigma_J^2\right]^{-1/2}\right)\right\}\ ,$$

where $\Phi(.)$ denotes the standard normal distribution function and $\widetilde{h}_t^i$ arise from **Step 1b** of **Algorithm : SIR**. If the parameters and model are true, then the estimated distribution functions should be independently uniformly distributed through time so $\widehat{u}_t \backsim UID(0,1)$, for $t = 1, ..., T$, as $M \to \infty$ (see Rosenblatt,1952).

3·5 *Model Comparison*

We have concentrated on maximum likelihood approaches for inference. However, we can also conduct model comparison, in a Bayesian context, by computing marginal likelihoods of competing models. We have the marginal likelihood,

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

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,

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

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 and the states, through particle filtering, when computing the marginal likelihoods, we do not fall victim to the nuisance parameter problem encountered in similar contexts using likelihood ratio tests.

3·6 *Implementation of SV-GARCH model*

We apply the same general methodology described above for the estimation for the SV-GARCH model. The procedure is similar and is conducted relatively straightforwardly, within the standard **Algorithm : SIR** framework. In this case, we require only to simulate forward from the transition equation (2·9) in conjunction with continuous resampling at **Step 3.** Thus no additional modifications (i.e. sub-algorithms) are required; as in the case of SV with leverage and jumps. As before, output such as filtered volatilites, quantiles and diagnostics are again obtained as a by-product of the procedure. The method is detailed in **Appendix C**.

4 Simulation Experiments

4·1 *Stochastic Volatility with Leverage and Jumps*

	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 1: *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.*

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 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. 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 1, 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,

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 2: *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=2000 in all cases.*

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 Figures 1 and 2.

Next, we generate 50 different time series each of length $T = 2000$, 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 2, for each of three cases considered. Variance estimates are computed using the OPG estimator for the variance covariance matrix.

The corresponding histograms in Figure 3 suggest convergence towards the mode and that we are not far from normality. In testing for bias we find very encouraging results.[9] 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

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

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 3: *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=500 and T=2000* .

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 4: *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=500 and T=2000* .

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 settings to render the estimation of the stochastic volatility components less accurate (see Eraker et al., 2003). 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 3 suggest that having smaller jumps occurring with high intensity induces a slight amount of bias is estimating of σ_η^2, ρ and p. In sharp 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 4. In this case, all parameters fall well within their 95% confidence limits (Figures 4 and 5). 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 6. [10] The diagnostic test illustrated by the QQ plot and autocorrelation function (ACF) indicate the prior and model are correct.

4·2 *SV-GARCH*

We now consider the performace of the estimator in the case of the SV-GARCH model. We generated 50 different time series each of length $T = 2000$. 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 $\theta = (\mu, \alpha, \beta, \varphi)$ for each run. We conduct four different experiments keeping the values of μ, α, β fixed at 0.010, 0.925 and 0.069 respectively and taking the values of $\varphi \in \{0.055, 0.10, 0.50, 0.90\}$. 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 5 for each of three cases considered. The histograms in Figures 7,8,9 and 10 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 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. In unreported results we repeated this experiment taking $M = 300$ and 600.There was no substantial variability in the results and the finding of unbiasedness remained unaltered an all cases. Futhermore, inorder to demonstrate the robustness of the estimation methodology in the case of SV-GARCH we also conduct another set of experiments where we run the smooth particle filter 50 times using different random number seeds for the smooth particle filter for each run, keeping the time series fixed. In considering the case of setting $\varphi = 0.5$, after running the smooth particle filter we maximize the estimated log-likelihood function with respect to $\theta = (\mu, \alpha, \beta, \varphi)$.We note a reduction in the variance of maximum likelihood estimator as M and T increase which is illustrated by kernel density estimates; Figures 11 and 12.[11]

$\varphi = 0.055$			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.056979	1.7851	2.2189
α	0.88332	1.4911	2.2969
β	0.076309	0.20840	0.15377
φ	0.088225	4.1954	1.3380
$\varphi = 0.10$			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.01136	0.0026629	0.0033242
α	0.92525	0.022125	0.024973
β	0.067481	0.016447	0.018369
φ	0.15287	9.988	4.2387
$\varphi = 0.50$			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.013730	0.0033056	0.0049365
α	0.92076	0.024331	0.020418
β	0.070135	0.017954	0.014082
φ	0.42088	8.1521	7.5707
$\varphi = 0.90$			
	$\overline{ML_s}$	$\overline{Var} \times 100$	$Var(ML_s) \times 100$
μ	0.013149	0.0032951	0.0035935
α	0.92182	0.019846	0.018877
β	0.069746	0.015915	0.014076
φ	0.86951	5.2917	3.2639

Table 5: *50 different datasets. Analysis of the maximum likelihood estimator for SV-GARCH model. M=500 and T=2000 in all cases.* True parameters: $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ and φ is allowed to vary.

	ML Estimate	Standard Error
SV: Log-lik value $=-2654.6$		
μ	-0.24769	0.10517
ϕ	0.94924	0.011944
σ_η^2	0.063963	0.0098940
SVL: Log-lik value $=-2645.4$		
μ	-0.17810	0.10185
ϕ	0.94360	0.009757
σ_η^2	0.069233	0.0093204
ρ	-0.31698	0.064704
SVLJ: Log-lik value $=-2621.1$		
μ	-0.13763	0.13635
ϕ	0.98046	0.0064549
σ_η^2	0.014700	0.0042593
ρ	-0.33150	0.095694
σ_J^2	34.749	15.037
p	0.0060659	0.0025706
SV-GARCH: Log-lik value $=-2632.5$		
γ	0.070330	0.012253
α	0.67445	0.034516
β	0.25683	0.028008
φ	0.056531	0.22346

Table 6: *Parameter estimates for SandP500 daily returns data for period, 02/02/1982 - 29/12/1989. M=500. GARCH: log-likelihood = -2738.1*

5 Empirical Examples

We now employ the described methodology to estimate four models; (i) stochastic volatility (SV), (ii) stochastic volatility with leverage (SVL), (iii) stochastic volatility with leverage and jumps (SVLJ) and (iv) SV-GARCH model, using daily returns S&P 500 over three different spans. Returns are continuously compounded and scaled by 100; holidays and weekends are excluded. This is a prominent index with actively traded futures and European option contracts. The spans we consider cover the well-documented episodes of market stress, October 1987, October 1997, late Summer-Fall 1998 as well as the most recent episode in Fall 2008. For each of the series, the parameter estimates along with standard errors[12] and log-likelihood values and for these four specifications are reported in Tables 6, 7,8 and 9. We illustrate the actual returns data, along with the quantiles of filtered standard deviation and filtered jump probabilities for SVLJ specification for the spans considered in Figure 13, 14, 15 and 16. These figures suggest

	ML Estimate	Standard Error
SV: Log-lik value $=-3044.7$		
μ	0.13181	0.18190
ϕ	0.98211	0.0059105
σ_η^2	0.022618	0.0048037
SVL: Log-lik value $=-2994.0$		
μ	0.24248	0.097671
ϕ	0.97367	0.0045830
σ_η^2	0.030461	0.0049492
ρ	-0.8106	0.04346
SVLJ: Log-lik value $=-2991.5$		
μ	0.25477	0.10002
ϕ	0.97651	0.0040141
σ_η^2	0.026944	0.0048600
ρ	-0.82879	0.043263
σ_J^2	6.1967	0.44835
p	0.0088753	0.0034897
SV-GARCH: Log-lik value $=-3045.5$		
γ	0.00981	0.00334
α	0.88777	0.01226
β	0.10412	0.01097
φ	0.01126	0.84638

Table 7: *Parameter estimates for S&P 500 daily returns data for period, 16/05/1995 - 24/04/2003. M=500. GARCH: log-likelihood = -3074.5*

that the path of the estimated filtered standard deviation capture adequately the underlying

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

[11] Futhermore, inorder to demonstrate the robustness of the estimation methodology we also conduct another set of experiments where we and run the smooth particle filter 50 times using different random number seeds for the smooth particle filter for each run. In considering the case of $\varphi = 0.5$, after running the smooth particle filter we maximize the estimated log-likelihood function with respect to $\theta = (\mu, \alpha, \beta, \rho)$.We note a reduction in the variance of maximum likelihood estimator as M and T increase. This illustrated by kernel density estimates; Figures 7 and 8.

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

volatility of the returns process in addition to identifying periods which may be described as market stress, i.e. short periods of time with clusters of large movements in returns. In addition, the filtered probabilities adequately identify jump times.

Estimates of the jump probabilities (times) and average jumps size allow us to better understand the contribution of these components to volatility, especially during periods of market stress. Understanding this contribution is extremely important because jump risk can typically not be hedged away and thus investors demand higher premia in order to carry this risk.[13] From our estimates of the jump components (p and σ_J^2) it is revealed that jumps over the three spans considered jumps can indeed be considered rare events which occur (approximately) between 1.3 and 2 times per year. The average jumps sizes across the three spans in Table 6, 7 and 8 do tend to differ in magnitude, with the largest being over the span containing the October 19,1987 crash (Table 6). The diagnostics do not reveal any evidence of potential misspecification of the SVLJ model for any series.

	ML Estimate	Standard Error
SV: Log-lik value = -2866.0		
μ	0.50059	0.48211
ϕ	0.99372	0.002733
σ_η^2	0.01683	0.003530
SVL: Log-lik value =-2806.5		
μ	0.58585	0.14734
ϕ	0.98776	0.003864
σ_η^2	0.02292	0.002276
ρ	-0.8438	0.037732
SVLJ: Log-lik value = -2800.2		
μ	0.58516	0.14901
ϕ	0.98772	0.00248
σ_η^2	0.02452	0.00422
ρ	-0.8634	0.03825
σ_J^2	3.8493	0.03825
p	0.00790	0.00309
SV-GARCH: Log-lik value =-2844.3		
γ	0.00760	0.001841
α	0.89639	0.012946
β	0.10093	0.012597
φ	0.009993	0.88838

Table 8: *Parameter estimates for S&P 500 daily returns data for period, 19/12/2000 - 12/12/2008. M=500. GARCH: log-likelihood = -2909.1*

For the spans considered in Tables 7 and 8 we find that the magnitude of the estimated leverage parameter is high, $\rho > |0.8|$. In contrast we find an relatively lower estimate of $\rho = -0.33$ for the earlier span (see Table 6). Furthermore, it is for this span that we find the highest gain of SVLJ in log-likelihood terms over the SVL model. We find that the inclusion of a leverage effect in general is extremely important when modelling stochastic volatility. This is indicated by the substantial gain in the log-likelihood over the standard SV model in all cases.

In Table 9 a longer time series of from 31/03/1981 to 13/01/2011 is considered which includes covers all the spans analysed above and including the financial turmoil in Fall 2008. We find

[13]Evidence of large jump risk premia is found by Pan (2002) (see also Eraker et al., 2003).

	ML Estimate	Standard Error		ML Estimate	Standard Error
SV: Log-lik value $=-8162.6$			SV: Log-lik value $=-7970.0$		
μ	-0.09941	0.14884	μ	-0.21525	0.11124
ϕ	0.98648	0.00251	ϕ	0.98017	0.003273
σ_η^2	0.02829	0.00217	σ_η^2	0.03369	0.002578
SVL: Log-lik value $=-8078.0$			SVL: Log-lik value $=-7904.8$		
μ	-0.00880	0.084021	μ	-0.10804	0.075098
ϕ	0.97712	0.002400	ϕ	0.97249	0.002967
σ_η^2	0.03768	0.002660	σ_η^2	0.038011	0.002632
ρ	-0.63807	0.032374	ρ	-0.56161	0.037398
SVLJ: Log-lik value $=-8046.1$			SVLJ: Log-lik value $=-7865.7$		
μ	0.01115	0.090532	μ	-0.10804	0.080454
ϕ	0.98307	0.002030	ϕ	0.97249	0.002349
σ_η^2	0.026877	0.002738	σ_η^2	0.023349	0.002557
ρ	-0.67240	0.032746	ρ	-0.67938	0.037476
σ_J^2	16.992	2.3339	σ_J^2	27.330	4.3646
p	0.005553	0.001457	p	0.003635	0.001148
SV-GARCH: Log-lik value $=-8148.6$			SV-GARCH: Log-lik value $=-7951.9$		
γ	0.01026	0.001337	γ	0.01572	0.002221
α	0.86198	0.006821	α	0.83693	0.009138
β	0.13379	0.005455	β	0.15427	0.008086
φ	0.01116	0.015510	φ	0.00121	0.24609

Table 9: *LEFT:Parameter estimates for S&P 500 daily returns data for period , 31/03/1987 - 13/01/2011. M=500. GARCH log-likelihood = -8318.1 RIGHT:Parameter estimates for Dow Jones Composite daily returns data for period , 31/03/1987 - 13/01/2011. M=500. GARCH log-likelihood =* -8133.7.

that the SVLJ model describes well the evolution of S&P 500 volatility over this longer span of 6000 observations. The estimate of leverage $\rho = -0.67$, falls between those found in the previous examples. Although jumps occur with roughly the same frequency found in the smaller samples, the average jump size is higher than that found in Tables 7 and 8; but lower as compared to that in Table 6. As a comparison we also fit the model to daily Dow Jones Composite returns (Figure 17). We find that the leverage effect is of a comparable magnitude to that found for S&P 500 over the same span, but jumps arrive 1.5 times less often and tend to be larger, i.e. the average jumps size was found to be 1.7 times larger relative to S&P 500. The ranking in terms of log-likelihood points of the four models remains unaltered.

Focusing on SV-GARCH, if we consider the relative ranking of this model then we find that it is generally outperformed by the SVL (by approximately a 30 log-likelihood point gain) when leverage is relatively high; whereas interestingly, it outperforms SVL for the example of the span considered in Table 6. In this case we found a relatively less pronounced estimate of leverage and significantly larger contribution of incorporating jump components. In all cases SV-GARCH decisively gains over the standard GARCH model. This is reinforced furthermore by the finding that φ is found to be close to zero in all cases, thus favouring a stochastic as opposed to GARCH-type evolution for volatility. For the span considered in Table 6, we illustrate the returns, filtered standard deviation paths and quantiles in Figure 18. Moreover the robustness of SV-GARCH to jumps/outliers relative to the standard GARCH model is demonstrated in terms of the log-likelihood error which captures the predictive gain of the SV-GARCH when such events occur. We highlight for this particular example its diagnostic performace relative to standard GARCH (see Figure 19). It appears that both SV-GARCH and GARCH are able to model the memory of the returns process equally as well but the QQ plots indicate the superior performace of SV-GARCH in capturing conditionally heavier-tailed behaviour.

6 Conclusion

This paper has attempted to provide a unified methodology in order to conduct likelihood-based inference on the unknown parameters of discrete-time, stochastic volatility models incorporating both a leverage effect and jumps in the returns process. It was demonstrated how the likelihood can be approximated using output generated by the particle filter for this class of model. Given the inclusion of leverage and jumps components in the SV model, it was demonstrated how the basic **Algorithm: SIR** can be modified (with a specific sub-algorithm) such that it facilitates smooth resampling to be undertaken. The latter is done in order ensure that the likelihood estimator is continuous as a function of the unknown parameters and enables the use of gradient-based (Newton-Raphson type) maximization algorithms. A great advantage of this unified methodology is that it delivers 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 computation time and more general than many alternatives in the literature. The computation time being linear in T. With regards to generality, note that the standard SV and SV with leverage models (SVL) are restricted forms of the SV with leverage and jumps model (SVLJ). It was 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. The Monte Carlo experiments indicate that the method is both robust and statistically efficient. When examining finite sample bias in parameters, very encouraging results are found; even considering very high jump intensity. On simulated datasets, the methodology was statistically efficient and computationally fast even for very long time series. In unreported results, $T = 20,000$ was also tried which confirmed these findings.

Furthermore, a new volatility model, SV-GARCH is introduced. Whilst combining elements of both SV and standard GARCH specifications, it has the attractive feature of inheriting all the same unconditional, in addition to most of the dynamic properties of the standard GARCH model. It is more robust to jumps (outliers) given that it displays conditionally heavier tailed behaviour. By incorporating only one parameter more than GARCH (denoted by φ), it remains a relatively parsimonious specification. The non-linear/non-Gaussian state space form of SV-GARCH facilitates approximating the likelihood using output generated by the particle filter in conjunction with smooth resampling. As in the case of SVLJ, the estimation framework straightforwardly yields objects of interest such as quantiles of filtered volatility as by-products of the procedure. Results of various Monte Carlo experiments showed that the method in this case was robust, computationally fast and statistically efficient. Investigating finite sample bias yielded very encouraging results.

Lastly, the proposed methodology was used to estimate four models (SV, SVL, SVLJ and SV-GARCH) for daily S&P 500 returns and compare their relative performance over various time spans. The SVLJ model was found to systematically outperform all the other models considered and did very well in identifying jumps times and adequately detecting periods of market stress. Of particular interest in these applications was to assess how leverage, frequency of jumps and average jump size differed over the various spans. The inclusion of leverage was found to be very important in modelling stochastic volatility in all cases. The inclusion of the jump components provided a further substantial gain in log-likelihood which varied in magnitude over the different spans we considered. Moreover, considering a long span ($T = 6,000$) covering all the well-documented episodes of market stress (i.e. 1987, 1997, 1998 and 2008). It was found that incorporating jump components lead to a substantial gain in excess of 30 log-likelihood points after having incorporated leverage. A comparative example using daily returns on the Dow Jones Composite index was also considered where this gain was found to be close to 40 log-likelihood points.

The SV-GARCH model consistently outperformed both the standard GARCH and SV models. By considering the error in the predictive log-likelihood components, the robustness of the SV-GARCH model to outliers relative to GARCH was also illustrated. It was found that the estimated value of φ is generally closer to zero than unity. Given the structure of the model, this would imply a more stochastic evolution for volatility rather than purely deterministic process implied by extreme case of $\varphi = 1$.

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.

Barndorff-Nielsen, O-E and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. Journal of the Royal Statistical Society, Series B, 63, 2002, 253–280.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics,* 31, 307–327.

Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics,* vol.69, No.3, pp.542-547, 1987.

Bollerslev, T., Chou, R. Y., and Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics,* vol. 52, issue 1-2, pages 5-59.

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.

Chopin, N. (2004). Central limit theorem for sequential Monte Carlo methods and its application to Bayesian inference, The Annals of Statistics, 32, 2385-2411.

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

Christoffersen, P., Jacobs, K., and Mimouni, K. (2010). Models for S&P dynamics: Evidence from realized volatility, daily returns, and options prices. (Forthcoming) in the Review of Financial Studies.

Corradi, V. (2000). Reconsidering the continuous time limit of the GARCH(1,1) process. Journal of Econometrics. 96, 145-153.

Danielsson, J. and J. F. Richard (1993). Accelerated Gaussian importance sampler with application to dynamic latent variable models. Journal of Applied Econometrics, 8.

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

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

Durham, G. B. (2007). SV mixture models with application to S&P 500 index returns. Journal of Financial Economics, 85, 822-856.

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). The impact of jumps in volatility and returns. 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.

Franses, P. H., Leij, V.-D. M. and Paap, R. (2008). A simple test for GARCH against stochastic volatility models. Journal of Financial Econometrics, 1–16.

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.

Geweke, 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.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: Inference, development, and communication. Econometric Reviews, 18, 1-73.

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. (1990). Simulation based inference in models with heterogeneity, Annales d'Economie et de Statistique, 69-107.

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

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.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. Journal of Finance, 42, 281-300.

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.

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.

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

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. (1990). ARCH models as diffusion approximations. Journal of Econometrics, 45, 7-38.

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.

Pitt, M. K. (2002). 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.

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

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.

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.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In O.E. Barndorff-Nielsen, D. R. Cox and D. V.Hinkley (eds.), *Statistical Models in Econometrics, Finance and Other Fields*, pp. 1–67. London: Chapman & Hall.

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

This section of the Appendix, Appendix A, deals with the general implementation of the particle filter. Specifically we detail the continuous resampling scheme, the smooth resampling algorithm, stratification and bias correction in the following sub-sections.

Continuous Resampling This section describes the continuous sampling of **Step 3** in **Algorithm : SIR**. We may think of this operating on the index itself as a multinomial or weighted bootstrap procedure. Alternatively we may, equivalently, think of sampling the unobserved state itself. Originally we have the weighted empirical cumulative distribution function (cdf), from which we invert, as

$$\widehat{G}(x) = \sum\nolimits_{k=1}^{M} \lambda^k I(x^{(k)} < x),$$

where the $x^{(k)}$ are sorted in ascending order and $\sum_{k=1}^{M} \lambda^k = 1$. We may approximate this, following Pitt (2002), by

$$\begin{aligned}\widetilde{G}(x) &= \pi^0 I(x^{(1)} < x) + \pi^M I(x^{(M)} < x) \\ &\quad + \sum\nolimits_{k=1}^{M-1} \pi^k G_k \left(\frac{x - x^{(k)}}{x^{(k+1)} - x^{(k)}} \right),\end{aligned}$$

where $\pi^0 = \lambda^1/2$, $\pi^M = \lambda^M/2$ and $\pi^k = \left(\lambda^{k+1} + \lambda^k\right)/2$ for $k = 1, \ldots, M-1$. The function $G_k\,(z)$ is chosen as a cdf on [0,1] so that $G_k\,(z)$ is monotonically non-decreasing $G_k\,(z) = 0$ for $z < 0$ and $G_k\,(z) = 1$ for $z > 1$. It is clear that $\widetilde{G}(x)$ is a valid cdf.

We choose to have all the $G_k(z) = z$ for $z \in [0,1]$. These uniform cdfs are simple to invert. It is straightforward to show as $M \to \infty$, $\widetilde{G}(z) \to \widehat{G}(z) \to G(z)$. This modification means for a fixed set of random numbers in both step 3 and step 2 of the SIR algorithm, we obtain continuous samples of the states as we vary θ. We also obtain a continuous likelihood surface, in θ, making Newton-Raphson methods straightforward to apply.

Algorithm: Smooth resampling

We are assuming that we have a set of uniforms that we use to invert the cdf $\widetilde{G}(x)$ above. So we have the set of M, sorted in ascending order, uniform variates $u_1 < \ldots < u_M$. The generation of the, possibly stratified uniforms is discussed later in this Appendix.

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

```
set s=0, j=1;
for {i=0:M}
{
    s=s+π^i;

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

For the selected regions where $r^j = 0$ we set $x_j^* = x^{(1)}$ and when $r^j = M$ we set $x_j^* = x^{(M)}$ otherwise setting

$$x_j^* = (x^{(r^j+1)} - x^{(r^j)}) \times u_j^* + x^{(r^j)}.$$

This produces as sample $x_1^*, \ldots, x_M^*$ from $\widetilde{G}(x)$ and is $O(M)$ computationally.

Stratification The above method uses a set of M uniform variates $u_1 < \ldots < u_M$. We use a stratified sampling scheme to generate these for purposes of this paper. Stratification reduces sample impoverishment and has been suggested in the context of particle filtering 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[14]. In contrast to the standard SIR method 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 $u_j = (j-1)/M + u/M$ for $j = 1, ..., M$. This is the approach adopted by Carpenter et al. (1999).

A closely related approach, used by Kitagawa (1996), is to generate the uniforms as,

$$u_k = \frac{(k-1)+u^k}{M}, \quad \text{where} \quad u^k \overset{iid}{\sim} UID(0,1).$$

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. Abstracting 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,

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

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

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

8 Appendix B

This section of the Appendix, Appendix B, deals with the specific implementation of the particle filter for the case of leverage and jumps. This relates to Section 3·3. Specifically we are concerned with **Step (1a)** of **Algorithm: SIR**.

Deriving the functional form of density $f(\epsilon_t|h_t, y_t)$ We describe continuous simulation methods (via inversion of the cumulative distribution function) for sampling from the mixture

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

The conditional probability of a jump is given by,

$$\begin{aligned} \Pr(J_t = 1|h_t, y_t) &= \tfrac{\Pr(y_t|J=1)\Pr(J=1)}{\Pr(y_t|J=1)\Pr(J=1)+\Pr(y_t|J=0)\Pr(J=0)}, \\ &= \tfrac{N(y_t|0;\exp(h_t)+\sigma_J^2)p}{N(y_t|0;\exp(h_t)+\sigma_J^2)p+N(y_t|0;\exp(h_t))(1-p)}. \end{aligned}$$

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

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

[14] 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.

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,

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

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 $\epsilon_t = y_t \exp(-h_t/2)$, as we observe the return innovation exactly. We therefore have the expression (3·8). If we denote $p_t^* \equiv \Pr(J_t=j|h_t,y_t)$ then this mixture is,

$$f(\epsilon_t|h_t,y_t) = \delta\left\{y_t \exp(-h_t/2) - \epsilon_t\right\} \times (1-p_t^*) + N(\epsilon_t|\upsilon_{\epsilon_1},\sigma_{\epsilon_1}^2) \times p_t^*.$$

We may invert the corresponding distribution function $F(\epsilon_t|h_t,y_t)$ straightforwardly allowing for draws which are continuous as a function of our parameters.

Assume we have generated a uniform random variate $U \backsim UID(0,1)$. We show how to generate a single sample $\epsilon_t = F^{-1}(U|h_t,y_t)$ accordingly, where $\epsilon_t^* = y_t \exp(-h_t/2)$,

$$K = \Phi\left(\frac{\epsilon_t^* - \upsilon_{\epsilon_1}^1}{\sigma_{\epsilon_1}^1}\right) p_t^*,$$

$p_t^* \equiv \Pr(J_t=j|h_t,y_t)$ again, and $\Phi(.)$ denotes the standard normal distribution function. The following scheme is applied,

- If $U \le K$, set $\epsilon_t = \upsilon_{\epsilon_1} + \sigma_{\epsilon_1}\Phi^{-1}\left(\frac{u}{p_t^*}\right)$.
- If $K < U \le K + (1-p_t^*)$, set $\epsilon_t = y_t \exp(-h_t/2)$.
- If $U > K + (1-p_t^*)$, set $\epsilon_t = \upsilon_{\epsilon_1} + \sigma_{\epsilon_1}\Phi^{-1}\left(\frac{U-(1-p_t^*)}{p_t^*}\right)$.

The above probability integral transform procedure is repeated for each of the (fixed and stratified) uniforms $u_{1,......,}u_M$ in order to obtain the required sample $\epsilon_t^i \backsim f(\epsilon_t^i|h_t^i,y_t)$, $i = 1,.....,M$.

9 Appendix C

Particle filter estimation of SV-GARCH

We start at $t=0$ with samples from the stationary distribution of GARCH, $v_0^i \sim f(v_0)$, $i=1,....,M$.

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

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

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

$$\begin{aligned}\lambda_{t+1}^{i} &= \frac{\omega_{t+1}^{i}}{\sum_{k=1}^{M}\omega_{t+1}^{k}},\\ \text{where } \omega_{t+1}^{i} &= f(y_{t+1}|\widetilde{v}_{t+1}^{i})\\ &= \left\{2\pi\widetilde{v}_{t+1}^{i}\right\}^{-\frac{1}{2}}\exp\left(-\frac{1}{2}\frac{y_{t+1}^{2}}{\sqrt{\widetilde{v}_{t+1}^{i}}}\right).\end{aligned}$$

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

   Replacing **Step 3** with continous resampling as in Section 3.2 and demonstrated in **Appendix A**.

Parameters of the SV-GARCH $\theta = (\mu, \alpha, \beta, \varphi)$ can be estimated by maximizing the log-likelihood function,

$$\begin{aligned}\widehat{\theta}_{SML} &= \arg\max\left\{\log\widehat{L}_{M}(\theta)\right\}\\ &= \arg\max\left\{\sum_{t=1}^{T}\log\left(\frac{1}{M}\sum_{k=1}^{M}\omega_{t}^{k}\right)\right\}.\end{aligned}$$

Similar to the case described for SV with leverage and jumps, for diagnostics we require the distribution function,

$$u_t = F(y_t|Y_{t-1}) = \int F(y_t|v_t)f(v_t|Y_{t-1})dv_t.$$

In the case of SV-GARCH the distribution function is estimated by,

$$\widehat{u}_t = \frac{1}{M}\sum_{i=1}^{M}\Phi(\frac{y_t}{\sqrt{\widetilde{v}_t^{i}}}),$$

where $\Phi(.)$ denotes the standard normal distribution function and $\widetilde{v}_t^{i}$ arise from **Step 1** of **Algorithm : SIR**. If the parameters and model are true, then the estimated distribution functions should be independently uniformly distributed through time so $\widehat{u}_t \backsim UID(0,1)$, for $t = 1, ..., T$, as $M \rightarrow \infty$.

FIGURES

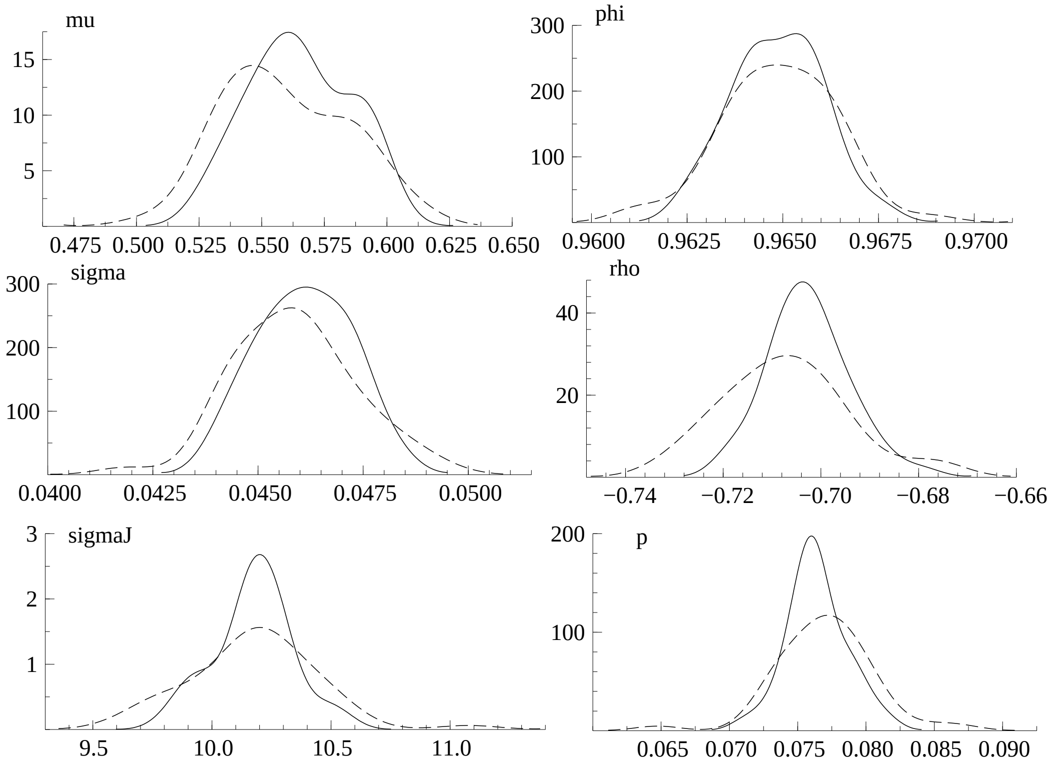

Figure 1: *Fixed simulated datasets. Dashed line:Kernel density estimate of the ML estimator for* $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, *for SV with leverage and jumps model;* $T = 1000$ *and* $M = 300$. *Solid line: Kernel density estimate of the ML estimator for* $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$, *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* $p = 0.10$.

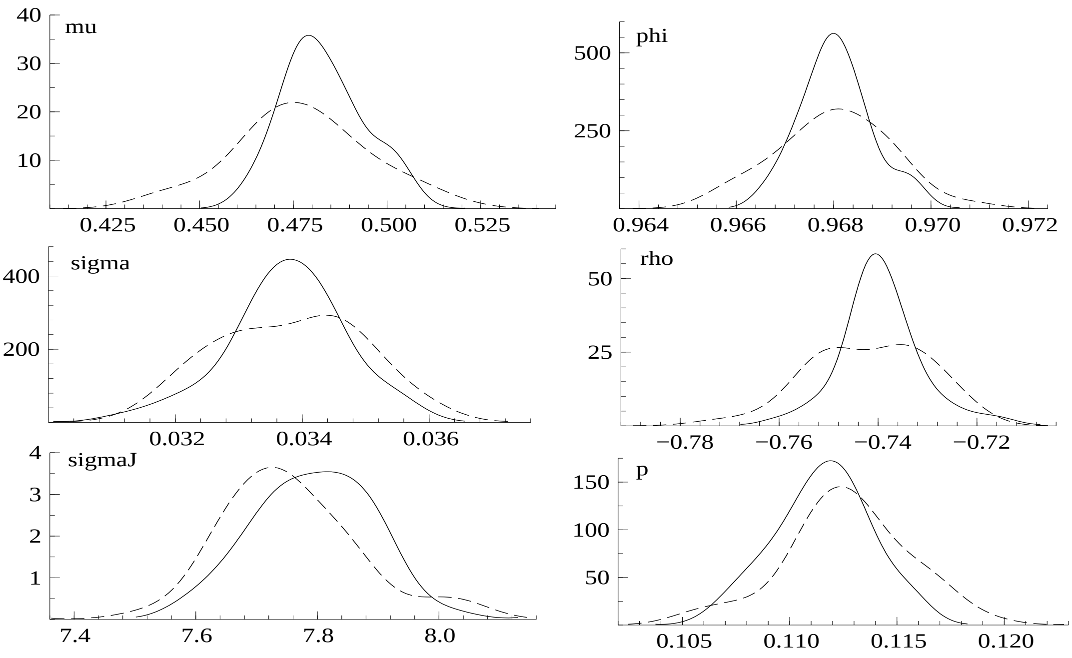

Figure 2: *Fixed simulated datasets. Dashed line:Kernel density estimate of the ML estimator for* $\theta = (\mu, \phi, \sigma^2_\eta, \rho, \sigma^2_J, p)$*, for SV with leverage and jumps model;* $T = 2000$ *and* $M = 300$*. Solid line: Kernel density estimate of the ML estimator for* $\theta = (\mu,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

.

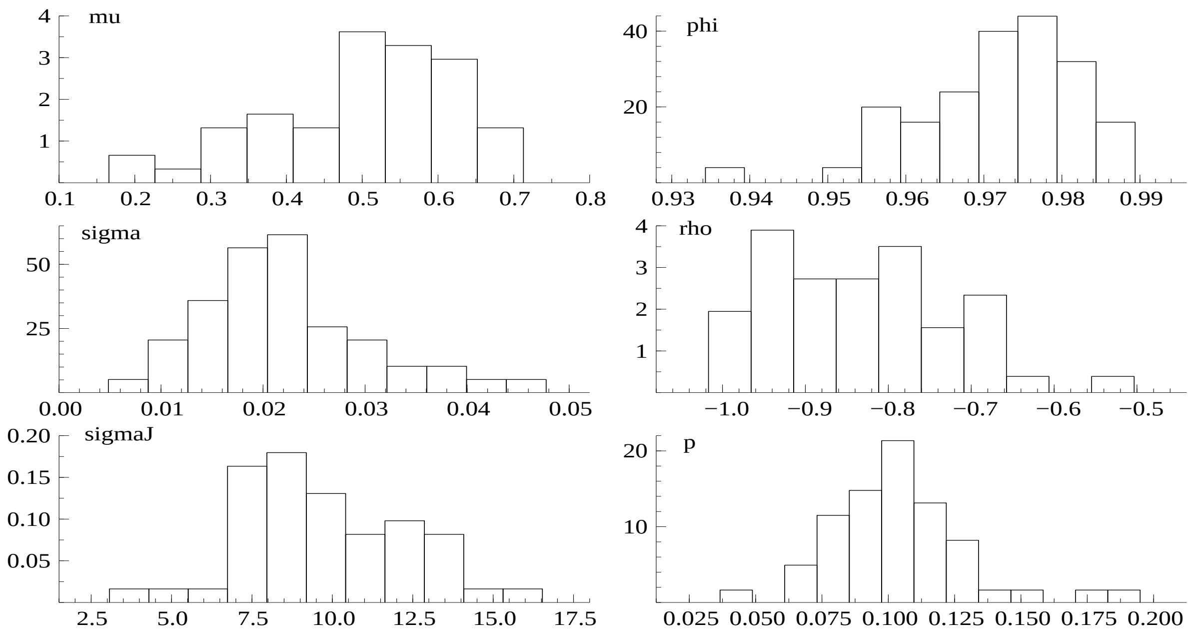

Figure 3: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$*, for SV with leverage and jumps model. True parameters,* $\mu = 0.5, \phi = 0.975, \sigma_\eta^2 = 0.02$ *and* $\rho = -0.8, \sigma_J^2 = 10$ *and* $p = 0.10$. $M = 500$ *and* $T = 2000$.

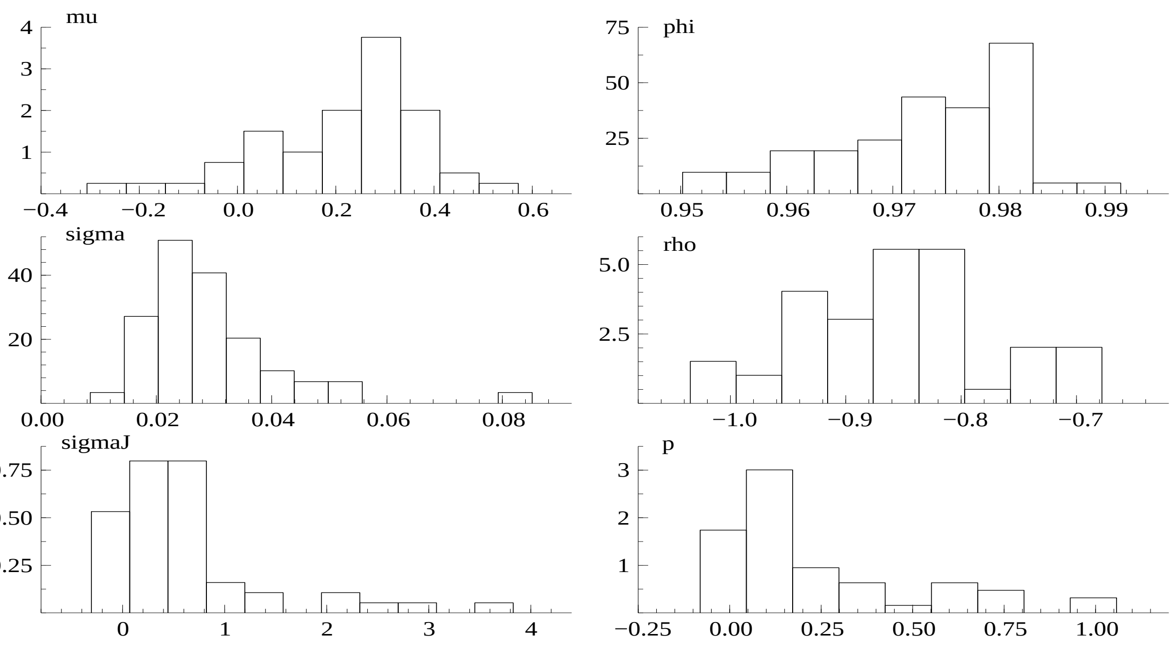

Figure 4: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$*, for SV with leverage and jumps model. True parameters,* $\mu = 0.25, \phi = 0.975, \sigma_\eta^2 = 0.025$ *and* $\rho = -0.8, \sigma_J^2 = 0.5$ *and* $p = 0.10$. $M = 500$ *and* $T = 2000$.

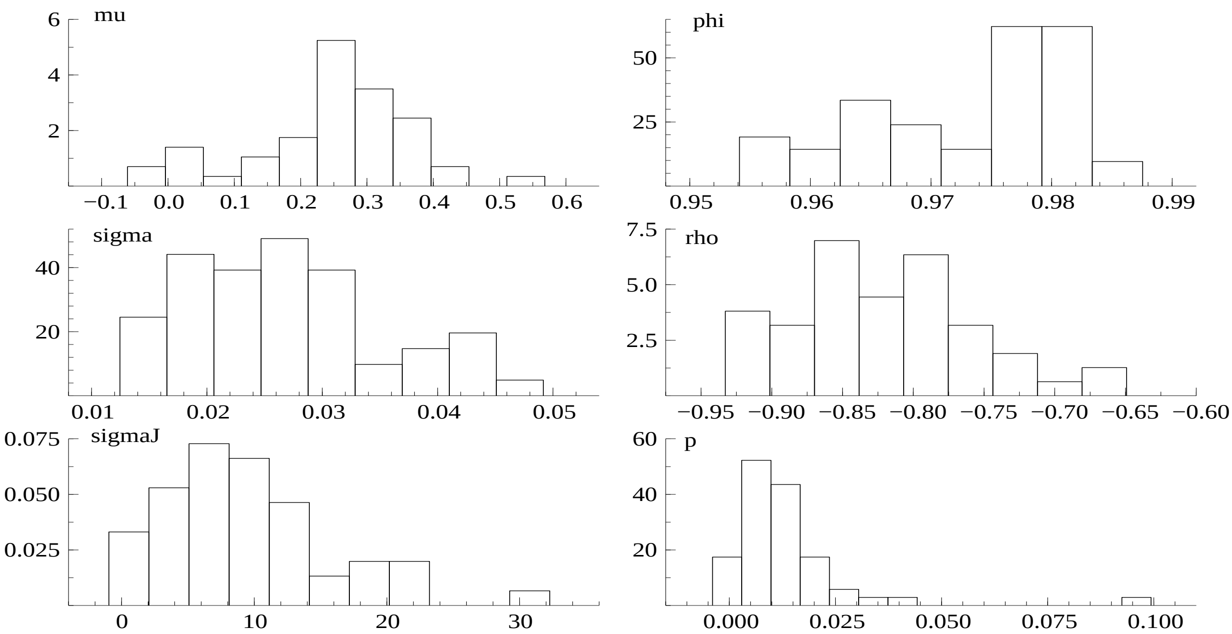

Figure 5: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \phi, \sigma_\eta^2, \rho, \sigma_J^2, p)$*, for SV with leverage and jumps model. True parameters,* $\mu = 0.25, \phi = 0.975, \sigma_\eta^2 = 0.025$ *and* $\rho = -0.8, \sigma_J^2 = 10$ *and* $p = 0.01$. $M = 500$ *and* $T = 2000$.

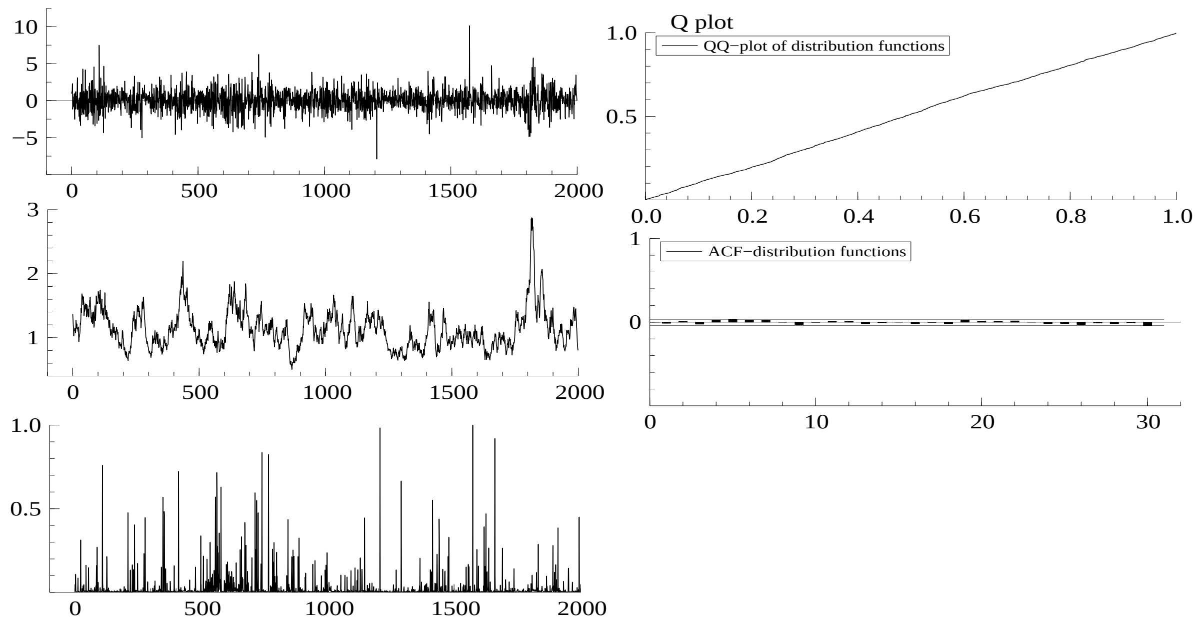

Figure 6: *Fixed simulated dataset. Parameters* $\mu = 0.25$, $\phi = 0.975$, $\sigma_\eta^2 = 0.025$, $\rho = -0.8$, $\sigma_J^2 = 10$ *and* $p = 0.01$. *and a single run of the smooth particle filter. LEFT PANEL: (i) Plot of data, (ii) filtered standard deviation, (iii) estimated jump probabilities. RIGHT PANEL: (i)QQ-plot of estimated distribution functions,* $\widehat{\mathrm{u}}_t$ *(ii) correlogram of* $\widehat{\mathrm{u}}_t$. $M = 500, T = 2000$.

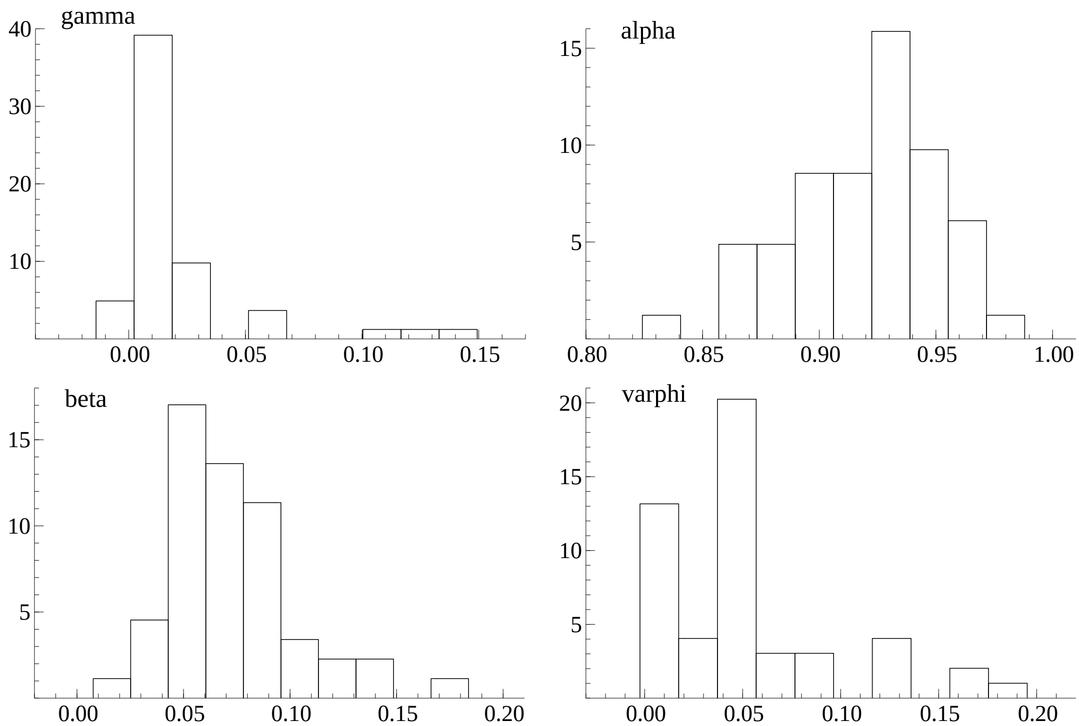

Figure 7: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \alpha, \beta, \varphi)$, *for SV-GARCH model. True parameters,* $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ *and* $\varphi = 0.05$. $M = 500$ *and* $T = 2000$.

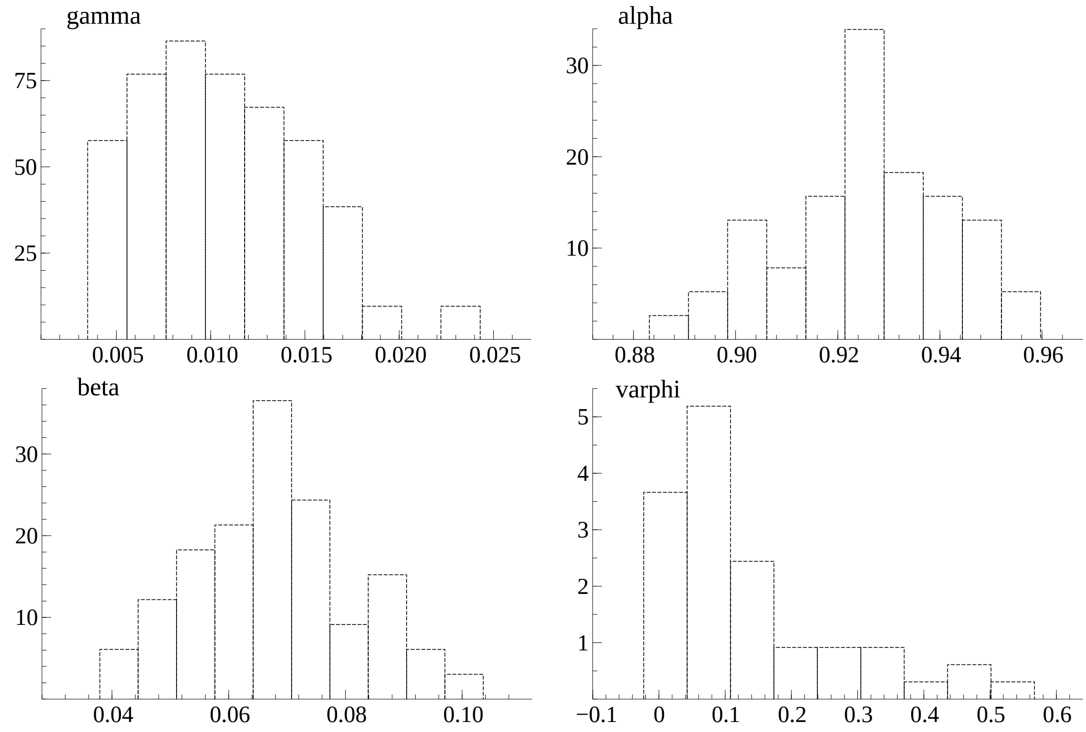

Figure 8: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \alpha, \beta, \varphi)$*, for SV-GARCH model. True parameters,* $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ *and* $\varphi = 0.10$*.* $M = 500$ *and* $T = 2000$

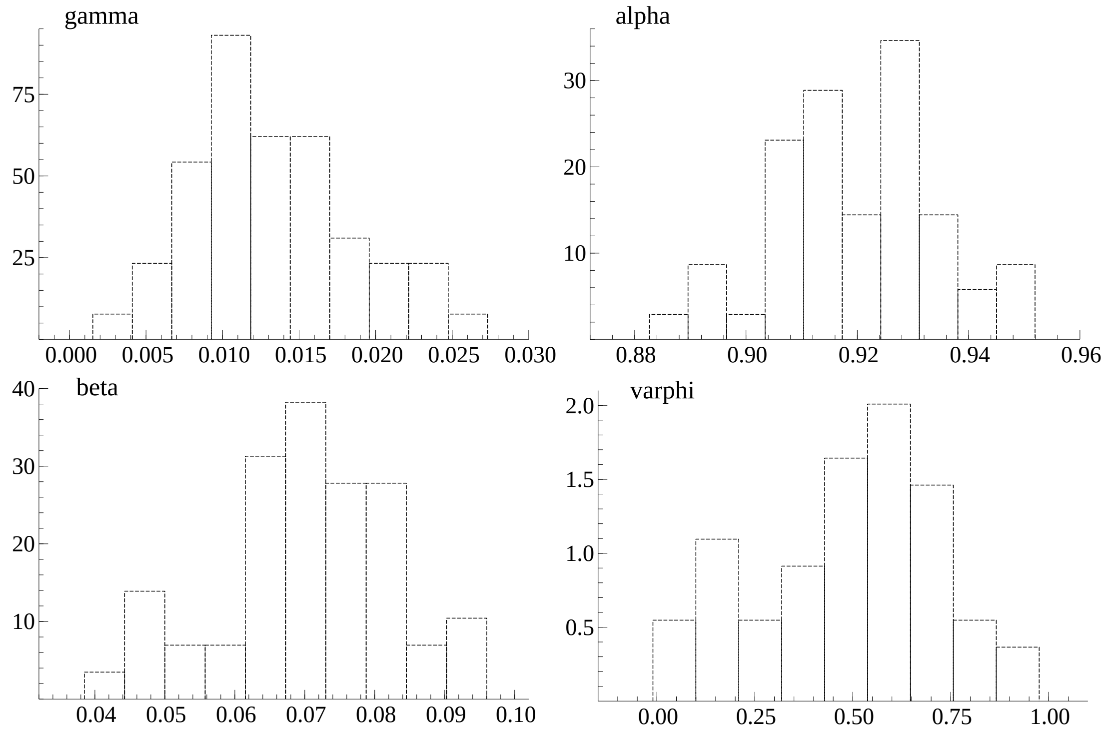

Figure 9: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \alpha, \beta, \varphi)$*, for SV-GARCH model. True parameters,* $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ *and* $\varphi = 0.5$. $M = 500$ *and* $T = 2000$

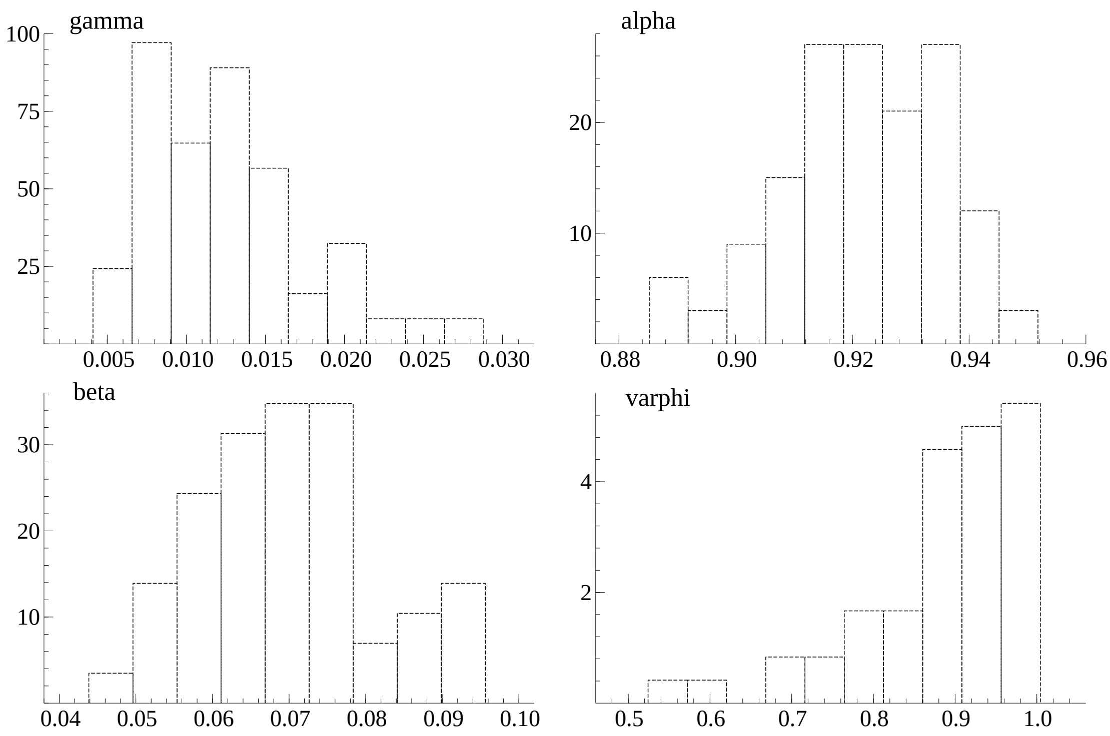

Figure 10: *50 different simulated datasets. Histogram of the Monte Carlo samples of the ML estimates for* $\theta = (\mu, \alpha, \beta, \varphi)$*, for SV-GARCH model. True parameters,* $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ *and* $\varphi = 0.90$. $M = 500$ *and* $T = 2000$

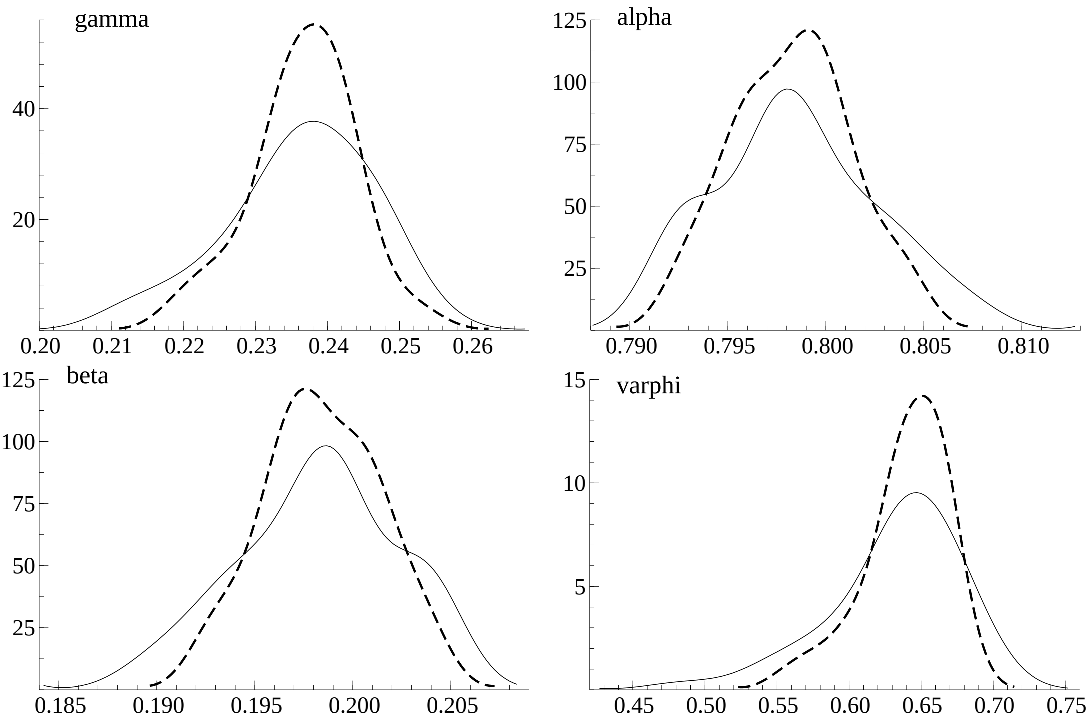

Figure 11: *Fixed simulated datasets. Dashed line:Kernel density estimate of the ML estimator for* $\theta = (\mu, \alpha, \beta, \rho)$, *for SV-GARCH model;* $T = 1000$ *and* $M = 300$. *Dashed line: Kernel density estimate of the ML estimator for* $\theta = (\mu, \alpha, \beta, \varphi)$, *for SV-GARCH model;* $T = 1000$ *and* $M = 600$. *True parameters,* $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ *and* $\varphi = 0.5$.

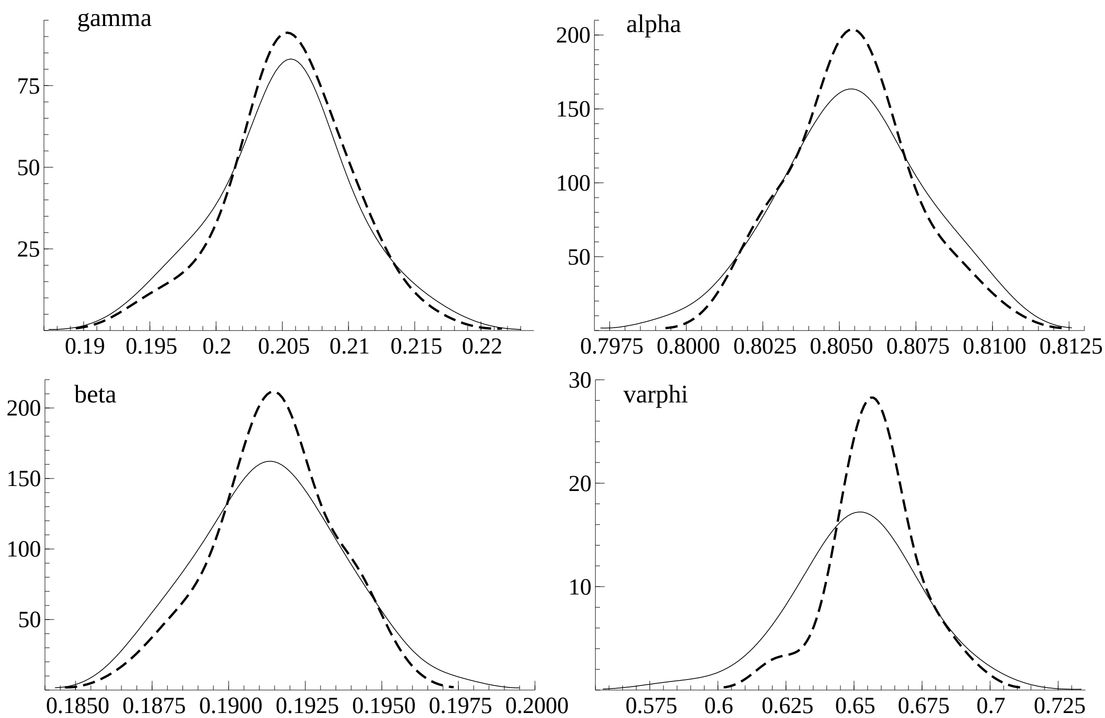

Figure 12: *Fixed simulated datasets. Dashed line:Kernel density estimate of the ML estimator for* $\theta = (\mu, \alpha, \beta, \rho)$*, for SV-GARCH model;* $T = 2000$ *and* $M = 300$*. Dashed line: Kernel density estimate of the ML estimator for* $\theta = (\mu, \alpha, \beta, \varphi)$*, for SV-GARCH model;* $T = 2000$ *and* $M = 600$*. True parameters,* $\mu = 0.010, \alpha = 0.925, \beta = 0.069$ *and* $\varphi = 0.5$*.*

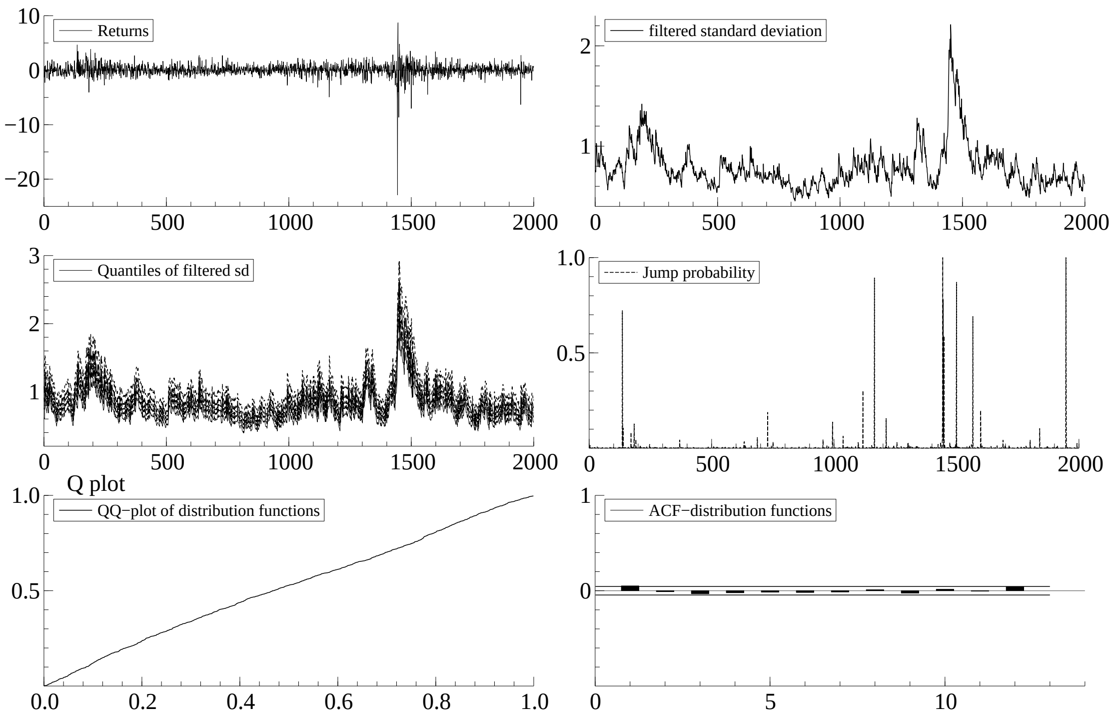

Figure 13: *Daily S&P 500 returns over the period 02/02/1982 - 29/12/1989. SV with leverage and jumps model. (i) returns data, (ii) quantiles of filtered standard deviation and (iii) estimated jump probabilities (iv) QQ-plot of estimated distribution functions,* $\widehat{u}_t$ *and (v) associated correlograms of* $\widehat{u}_t$. $M = 500$.

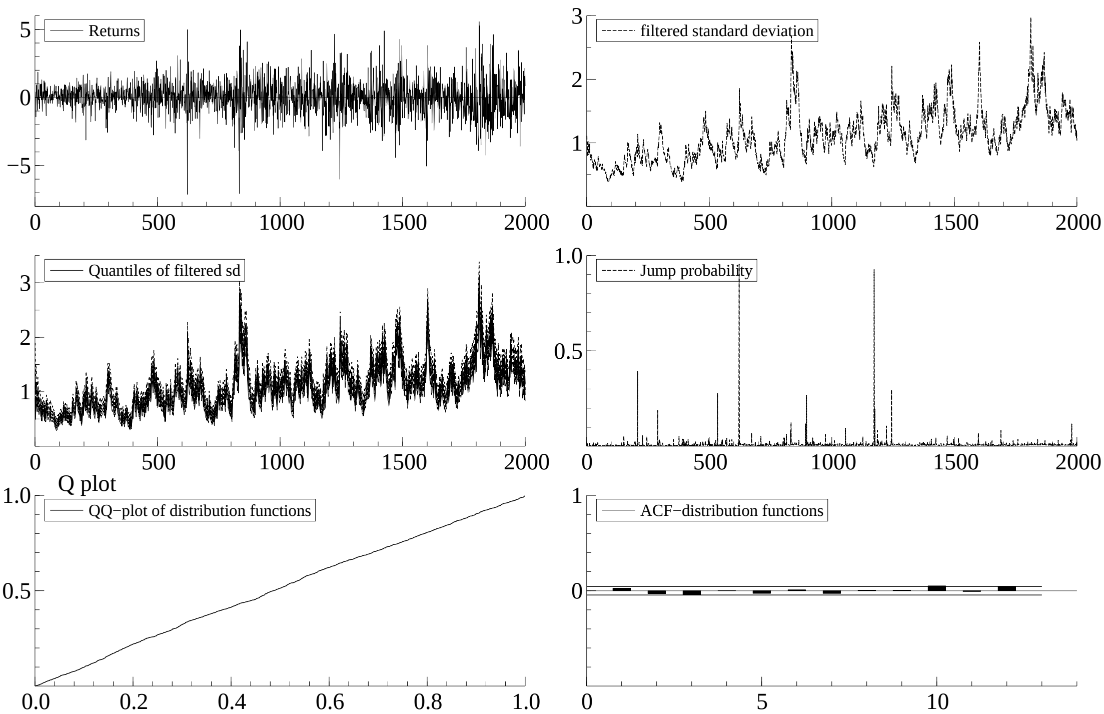

Figure 14: *Daily S&P 500 returns over the period 16/05/1995 - 24/04/2003. SV with leverage and jumps model. (i) returns data, (ii) quantiles of filtered standard deviation, (iii) estimated jump probabilities, (iv) QQ-plot of estimated distribution functions,* $\widehat{u}_t$ *and (v) associated correlograms of* $\widehat{u}_t$ $M = 500$.

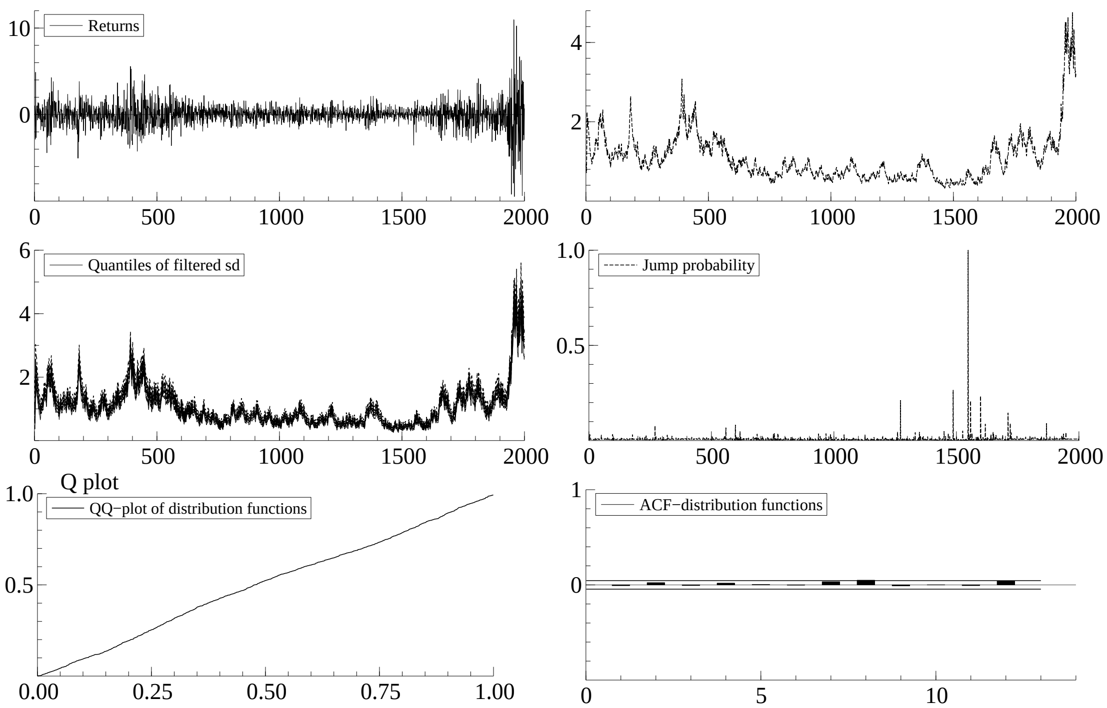

Figure 15: *Daily S&P 500 returns over the period 19/05/1995 - 12/12/2008. SV with leverage and jumps model. (i) returns data, (ii) quantiles of filtered standard deviation,(iii) estimated jump probabilities, (iv) QQ-plot of estimated distribution functions,* $\widehat{u}_t$ *and* (v) *associated correlograms of* $\widehat{u}_t$ $M = 500$.

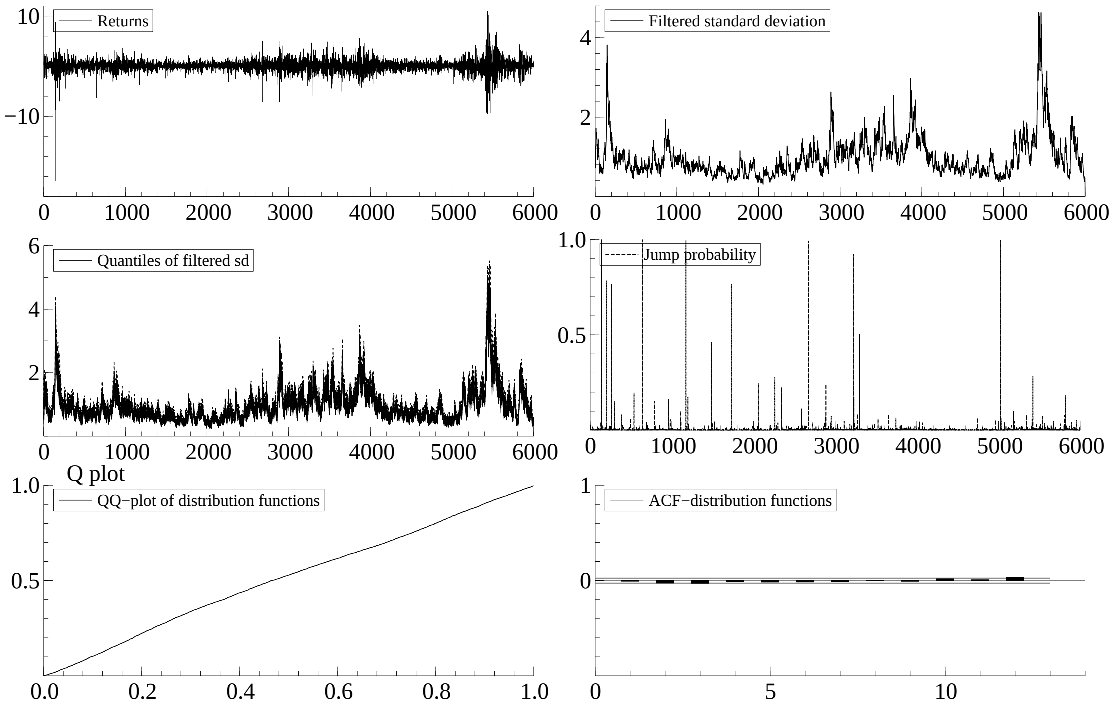

Figure 16: *Daily S&P 500 returns over the period 31/03/1987 - 13/01/2011. SV with leverage and jumps model. (i) returns data, (ii) quantiles of filtered standard deviation and (iii) estimated jump probabilities, (iv) QQ-plot of estimated distribution functions,* $\widehat{u}_t$ *and* (v) *associated correlograms of* $\widehat{u}_t$ $M = 500$.

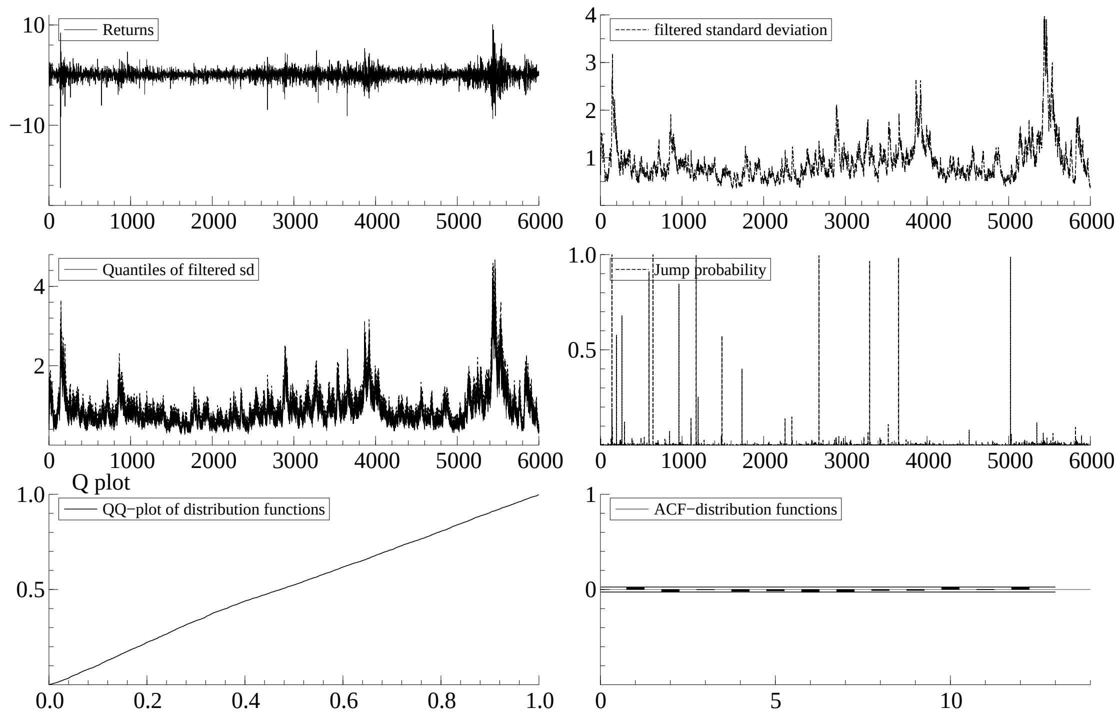

Figure 17: *Daily Dow Jones Composite returns over the period 31/03/1987 - 13/01/2011. SV with leverage and jumps model. (i) returns data, (ii) quantiles of filtered standard deviation and (iii) estimated jump probabilities, (iv) QQ-plot of estimated distribution functions,* $\widehat{u}_t$ *and (v) associated correlograms of* $\widehat{u}_t$. $M = 500$.

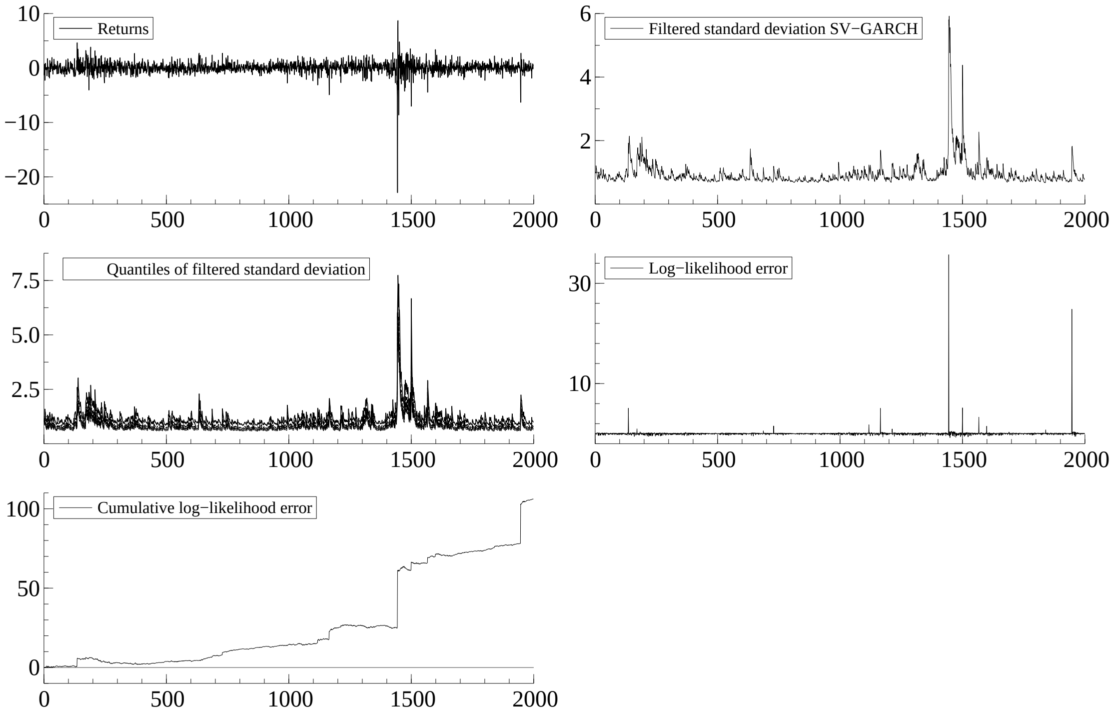

Figure 18: *Daily S&P 500 returns over the period 02/02/1982 - 29/12/1989. SV-GARCH model. (i) returns data, (ii) filtered standard deviation, (iii) quantiles of filtered standard deviation,(iv) error in log-likelihood components between SV-GARCH and GARCH and (v) cumulative error.* $M = 500$.

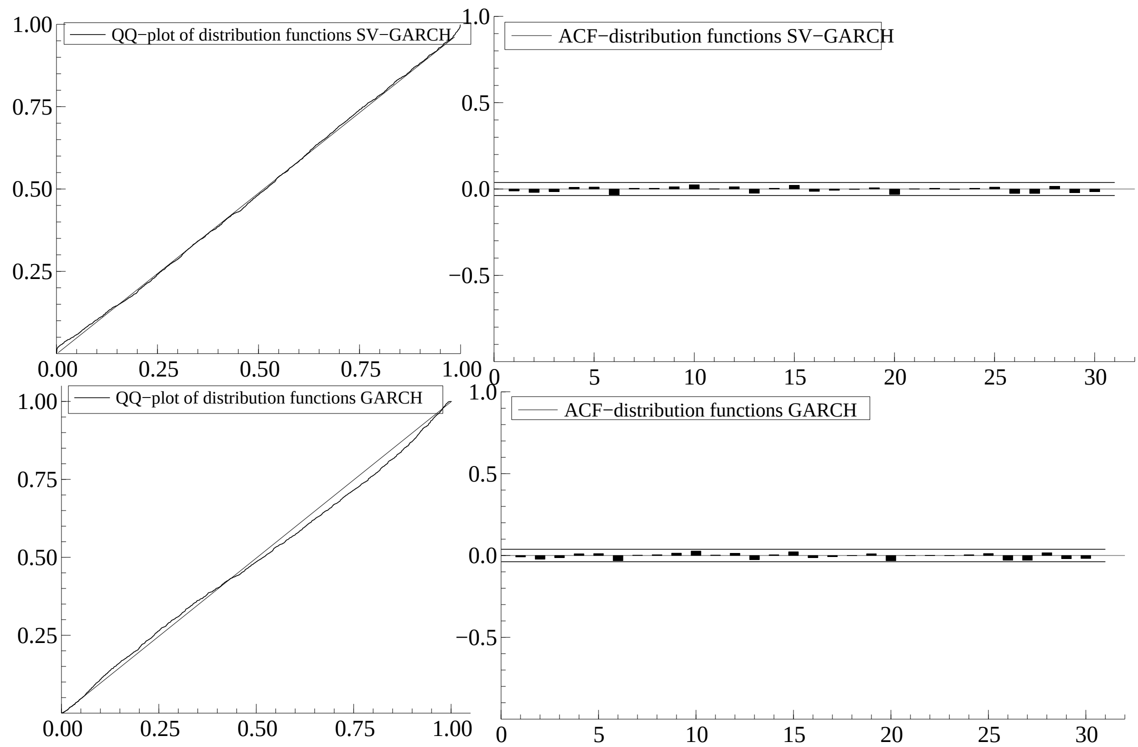

Figure 19: *Diagnostics for SV-GARCH and standard GARCH. Daily S&P 500 returns over the period 02/02/1982 - 29/12/1989. Left panel: QQ-plot of estimated distribution functions, $\widehat{u}_t$. Right panel: Associated correlograms of $\widehat{u}_t$. M=500.*

Documents de Travail

300. X. Ragot, "The Case for a Financial Approach to Money Demand," October 2010

301. E. Challe, F. Le Grand and X. Ragot, "Incomplete markets, liquidation risk, and the term structure of interest rates," October 2010

302. F. Le Grand and X. Ragot, "Prices and volumes of options: A simple theory of risk sharing when markets are incomplete," October 2010

303. D. Coulibaly and H. Kempf, "Does Inflation Targeting decrease Exchange Rate Pass-through in Emerging Countries?," November 2010

304. J. Matheron, « Défiscalisation des heures supplémentaires : une perspective d'équilibre général », Décembre 2010

305. G. Horny and P. Sevestre, "Wage and price joint dynamics at the firm level: an empirical analysis," December 2010

306. J. Coffinet and S. Lin, "Stress testing banks' profitability: the case of French banks," December 2010

307. P. Andrade and H. Le Bihan, "Inattentive professional forecasters," December 2010

308. L. Clerc, H. Dellas and O. Loisel, "To be or not to be in monetary union: A synthesis," December 2010

309. G. Dufrénot and S. Malik, "The changing role of house price dynamics over the business cycle," December 2010

310. M. Crozet, I. Méjean et S. Zignago, "Plus grandes, plus fortes, plus loin…Les performances des firmes exportatrices françaises," Décembre 2010

311. J. Coffinet, A. Pop and M. Tiesset, "Predicting financial distress in a high-stress financial world: the role of option prices as bank risk metrics," Décembre 2010

312. J. Carluccio and T. Fally, "Global Sourcing under Imperfect Capital Markets," January 2011

313. P. Della Corte, L. Sarnoz and G. Sestieri, "The Predictive Information Content of External Imbalances for Exchange Rate Returns: How Much Is It Worth?," January 2011

314. S. Fei, "The confidence channel for the transmission of shocks," January 2011

315. G. Cette, S. Chang and M. Konte, "The decreasing returns on working time: An empirical analysis on panel country data," January 2011

316. J. Coffinet, V. Coudert, A. Pop and C. Pouvelle, "Two-Way Interplays between Capital Buffers, Credit and Output: Evidence from French Banks," February 2011

317. G. Cette, N. Dromel, R. Lecat, and A-Ch. Paret, "Production factor returns: the role of factor utilisation," February 2011

318. S. Malik and M. K Pitt, "Modelling Stochastic Volatility with Leverage and Jumps: A Simulated Maximum Likelihood Approach via Particle Filtering," February 2011

Pour accéder à la liste complète des Documents de Travail publiés par la Banque de France veuillez consulter le site :
http://www.banque-france.fr/fr/publications/documents_de_travail/documents_de_travail_11.htm

For a complete list of Working Papers published by the Banque de France, please visit the website:
http://www.banque-france.fr/fr/publications/documents_de_travail/documents_de_travail_11.htm

Pour tous commentaires ou demandes sur les Documents de Travail, contacter la bibliothèque de la Direction Générale des Études et des Relations Internationales à l'adresse suivante :

For any comment or enquiries on the Working Papers, contact the library of the Directorate General Economics and International Relations at the following address :

BANQUE DE FRANCE
49- 1404 Labolog
75049 Paris Cedex 01
tél : 0033 (0)1 42 97 77 24 ou 01 42 92 62 65 ou 48 90 ou 69 81
email : marie-christine.petit-djemad@banque-france.fr
jeannine.agoutin@banque-france.fr
michael.brassart@banque-france.fr
veronique.jan-antuoro@banque-france.fr
nathalie.bataille-salle@banque-france.f