Statistics and Its Interface Volume 13 (2020) 209–219

Non-Gaussian stochastic volatility model with jumps via Gibbs sampler

Arthur T. Rego* and Thiago R. dos Santos

In this work we propose a model for estimating volatility from financial time series, extending the non-Gaussian family of space-state models with exact marginal likelihood proposed by [6]. On the literature there are models focused on estimating financial assets risk, however, most of them rely on MCMC methods based on Metropolis algorithms, since full conditional posterior distributions are not known. We present an alternative model capable of automatically estimating the volatility, since all full conditional posterior distributions are known, and it is possible to obtain an exact sample of volatility parameters via Gibbs Sampler. The incorporation of jumps in returns allows the model to capture speculative movements of the data so that their influence does not propagate to volatility. We evaluate the performance of the algorithm using synthetic and real data time series and the results are satisfactory.

AMS 2000 subject classifications: Primary 62M10, 62P20; secondary 91B70.
Keywords and phrases: Financial time series, Stochastic volatility, Gibbs Sampler, Dynamic linear models.

1. INTRODUCTION

Understanding the behavior of asset prices is essential for capital allocation decisions between the available investment options. Such a decision depends on what one thinks about risks and returns associated with these investment options. The most accepted theory is that the returns on high volatility assets follow a random walk with some outliers points, that usually occur during abnormal volatility increases, such as in financial and political crisis events. The future returns would be unpredictable, but the volatility can be estimated and monitored in order to detect such events and anticipate their movements. Under the Bayesian perspective, the inferential procedure of the stochastic volatility models commonly used are mostly based on intensive computational methods, e.g., Markov Chain Monte Carlo (MCMC) methods using Metropolis-Hastings algorithms, which raises questions about the usage of more automatic and simpler computational implementation methods that can be used to bring fast and reliable results. Dealing with financial time series brings three main challenges, which include: finding a model that fits well to the data and accommodates the heavy tails that exist in non-Gaussian returns; is fast enough to bring results on time to be used by market agents; and is flexible to include a new source of data and accommodate outliers and skewness, that improve the model fit.

Many models have been developed for risk measuring purposes. For example, [5] adopt the Stochastic Volatility model (SV) and study the influence of inserting jumps to improve the model. They suggest including jumps on returns and volatility in order to improve the model dynamics in case of spot changes in volatility, as in financial crisis moments. [13] include the leverage effect on SV models, which refers to the increase in volatility following a previous drop in stock returns, and model it as the negative correlation coefficient between error terms of stock returns. [12] extend its application to Stochastic Volatility with Jumps (SVJ) models with heavy-tailed distribution, obtained by a scale mixture of a generalized gamma distributed mixture component together with a normally distributed error in order to generate generalized Skew-t distributed innovations, and discuss the fit gains on including such feature. [21] investigate sequential, or online, Bayesian estimation for inference of stochastic volatility with variance-gamma jumps in returns and compare its performance to the model that uses offline Markov Chain Monte Carlo. [7] use a stochastic volatility model with a two-step estimation method. In the first step, they nonparametrically estimate the instantaneous volatility process, and, in the second step, standard estimation methods for fully observed diffusion processes are employed, but with the filtered or estimated volatility process replacing the latent process. [2] also provide nonparametric methods for stochastic volatility modeling, allowing the joint evaluation of return and volatility dynamics with nonlinear drift and diffusion functions, nonlinear leverage effects and jumps in returns and volatility with state-dependent jump intensities. However, those models are either nonparametric or the parameter estimation is somewhat complex, since there is no closed-form to the full conditional posterior distributions, being necessary the use of MCMC methods with Metropolis-Hastings (MH) steps.

The main objective of this work is to find an alternative model that accommodates speculative financial asset returns data, allows for the innovations to assume heavy-tailed distributions, includes jumps on returns in order to

arXiv: 1809.01501
*Corresponding author.

get the impact of uncommon events on financial markets and has a simpler and more automatic inferential procedure, like Gibbs Sampling together with a block sampling structure, for estimating the model parameters while mitigating convergence issues. We introduce a non-Gaussian stochastic volatility model with jumps, as well as an application to real S&P-500 index and Brent Crude future returns time series.

The proposed model is innovative on bringing beneficial aspects and characteristics from two different classes of models: the Stochastic Volatility (SV), used by [5], and the Dynamic Linear Model, used by [6]. The result is a model that allows the inclusion of jumps, a multiplicative evolution equation, and a mixture component to produce heavy-tailed distributions. The full conditional distribution of model parameters has a closed-form, including the volatility. Thus, it is easier to sample these distributions, making the MCMC algorithm simpler by the use of Gibbs Sampler steps. Furthermore, an exact sample from the full conditional distribution of the volatility can be drawn using a smoothing procedure feature. The MCMC algorithm presented in this paper consists of Gibbs Sampler steps only, which is not seen on other volatility models with jumps available on literature. The proposed model structure allows an automatic sampling procedure and implementation simplicity, avoiding the use of Metropolis steps and nonlinear filters, which require additional implementing effort.

2. THE NON-GAUSSIAN STOCHASTIC VOLATILITY MODEL WITH JUMPS ON RETURNS

The non-Gaussian stochastic volatility model with jumps on returns (NGSVJ) for time series $\{y_t\}_{t=1}^{n}$ is given by:

$$y_t = \mu + J_t^y + \upsilon_t, \qquad \upsilon_t|\gamma_t \sim N(0, \gamma_t^{-1}\lambda_t^{-1}), \tag{1}$$

$$\lambda_t = \omega^{-1}\lambda_{t-1}\zeta_t, \qquad \zeta_t|Y_{t-1}, \varphi \sim Beta(\omega a_{t-1}, (1-\omega)a_{t-1}), \tag{2}$$

where $J_t^y = \xi_{t+1}^y N_{t+1}^y$, $\xi^y \sim N(\mu_y, \sigma_y^2)$, and $Pr(N_{t+1}^y = 1) = \rho_y$.

In this model, y_t represents the log-return in percentage, defined as $y_t = 100 \times (log(S_t) - log(S_{t-1}))$, where S_t is the asset price on time t. J_t^y is the jump, composed by the jump indicator N_{t+1}^y and magnitude $\xi_{t+1}^y \sim N(\mu_y, \sigma_y^2)$, in the same way proposed by [5]. μ represents the equilibrium log-return of y_t.

γ_t is the variance mixture component [6]. Using $\gamma_t \sim G(\frac{\nu}{2}, \frac{\nu}{2})$, the unconditional distribution of errors assume a $t_\nu(0,1)$ distribution. This structure allows the avoidance of nonlinear filters, since conditioned on the mixture component the model is Gaussian, but unconditionally it follows a Student-T distribution.

λ_t^{-1} is the volatility of returns, and the main interests lay on estimating its value over time since it is the main variable on risk and stock options pricing. ω is a discount factor and is specified, in order to avoid an MH step for its estimation, keeping the estimation procedure more automatic via Gibbs Sampling. a_{t-1} is the shape parameter of the filtering distribution of λ_t, which is described on details in [6].

The model provides the needed flexibility by using the State Space Model (SSM) form, with mixtures on variance in order to achieve non-Gaussian distribution for innovations. It has also a formulation that allows the full conditional posterior distributions to be available, so that Gibbs Sampler can be used to sample from the conditional posterior distributions, bringing implementation simplicity to the model.

In this case, there are no dimensionality issues with the parametric space, since all full conditional posterior distributions are obtained through the model properties and can be sampled via Gibbs Sampler. Using proper priors to parameters it is possible to obtain the full conditional posterior distribution, where the priors are chosen in order to obtain a conjugate posterior distribution. Another advantage lays on sampling mean μ and volatility $\lambda = \{\lambda_t\}_{t=1}^{n}$ in blocks, speeding up the sampling process. A more detailed description of the model procedures can be seen in [6].

The inclusion of jumps, adapted from [5], gets the advantage of models properties that guarantee a simpler computational implementation sampling method. This work only explores the addition of jumps in returns as the inclusion of jumps on volatility requires a more complex structure to preserve model properties and ensure fast inference procedures. The inclusion of jumps on volatility will be addressed to on future papers.

2.1 Bayesian inference

For the mean parameter of the log-returns, μ, a prior $N(m_0, C_0)$ is specified and the samples of its posterior distribution can be obtained through a Forward Filtering Backward Sampling (FFBS) algorithm, available in [15].

For ease of notation, let $\Phi = (\mu, \underset{\sim}{J^y}, \underset{\sim}{\gamma}, \underset{\sim}{\lambda}, \mu_y, \sigma_y^2, \underset{\sim}{\xi}, \rho_y, \underset{\sim}{N^y})$, excluding the parameter being evaluated, i.e. $\Phi_{[-\lambda]} = (\mu, \underset{\sim}{J^y}, \underset{\sim}{\gamma}, \mu_y, \sigma_y^2, \underset{\sim}{\xi}, \rho_y, \underset{\sim}{N^y})$.

With the prior distribution for λ_t, which is given by $\lambda_t|Y_{t-1} \sim G(\omega a_{t-1}, \omega b_{t-1})$, and following the method proposed by [6], the updating distribution is:

$$p(\lambda_t|Y_t, \Phi_{[-\lambda]}) \sim G\left(\omega a_{t-1} + \frac{1}{2}, \omega b_{t-1} + \gamma_t \frac{(y_t - \mu - J_t)^2}{2}\right). \tag{3}$$

The procedure for sampling from $(\lambda_t|Y_n, \Phi_{[-\lambda]})$ can be seen on Appendix A [6]. ω is a fixed discount factor.

For the mixture component γ_t, a prior $G(\frac{\nu}{2}, \frac{\nu}{2})$ is defined, and, when mixed as γ_t^{-1} (resulting in Inverse-Gamma) leads to a Student-t with ν degrees of freedom to the innovations. The full conditional posterior distribution is:

$$p(\gamma_t|Y_t, \Phi_{[-\gamma]}) \sim G\left(\frac{\nu}{2} + \frac{1}{2}, \frac{\nu}{2} + \lambda_t \frac{(y_t - \mu - J_t)^2}{2}\right). \tag{4}$$

The parameter ν will be specified, since its posterior distribution does not have a closed form, leading to a Metropolis step. A grid of values is made for it, comparing deviance information criterion (DIC), for different values of ν. Recalling that the main objective of the NGSVJ model is to keep an automatic and simple procedure for the model estimation and, once it is determined for a specific asset time series, ν does not need to be changed for future observations.

The jump sizes ξ^y_{t+1} follow a $N(\mu_y, \sigma^2_y)$. For the mean μ_y a non-informative prior $N(m, v)$ is set, resulting in a full conditional posterior:

$$p(\mu_y|Y_n, \Phi_{[-\mu_y]}) \sim \mathrm{N}\left(\frac{m\sigma^2_y + vn\bar{\xi^y}}{\sigma^2_y + nv}, \frac{v\sigma^2_y}{\sigma^2_y + nv}\right). \tag{5}$$

For the variance σ^2_y a prior $IG(\alpha, \beta)$ is assumed, resulting in the full conditional posterior:

$$p(\sigma^2_y|Y_n, \Phi_{[-\sigma^2_y]}) \sim IG\left(\alpha + \frac{n}{2}, \beta + \frac{\sum_{\substack{i=1\\ J_i \neq 0}}^{t} (\xi^y_{i+1} - \mu_y)^2}{2}\right). \tag{6}$$

In both cases, n is the number of times that the jump is observed, and $\bar{\xi^y}$ the mean of jump sizes ξ^y_{t+1}. As the prior of jump sizes is assumed to be Normal, the full conditional posterior is also Normal, given by:

$$p(\xi^y_{t+1}|Y_t, \Phi_{[-\xi]}) \sim N\left(m^*_\xi, v^*_\xi\right), \tag{7}$$

where:

$$m^*_\xi = \frac{\mu_y \gamma_t^{-1}\lambda_t^{-1} + y_t\sigma^2_y - \mu\sigma^2_y}{\sigma^2_y + \gamma_t^{-1}\lambda_t^{-1}}, \tag{8}$$

$$v^*_\xi = \frac{\sigma^2_y \gamma_t^{-1}\lambda_t^{-1}}{\sigma^2_y + \gamma_t^{-1}\lambda_t^{-1}}. \tag{9}$$

For jump probabilities ρ_y, a prior $Beta(\alpha, \beta)$ is set. The full conditional posterior is given by:

$$p(\rho_y|Y_n, \Phi_{[-\rho_y]}) \sim Beta\left(\alpha + \sum_{t=1}^{n} N^y_t, \beta + n - \sum_{t=1}^{n} N^y_t\right). \tag{10}$$

Since the jump indicator N^y can assume only two values, 0 or 1, the probability of observation at $t+1$ be a jump is given by:

$$P(N^y_{t+1} = 1|Y_{t+1}, \Phi_{[-N]}) \propto \rho_y \times P(Y_{t+1}|N^y_{t+1} = 1, \Phi_{[-N]}), \tag{11}$$

which is easy to calculate, since $P(Y_{t+1}|N^y_{t+1} = 1, \Phi_{[-N]})$ is a Normal distribution. Using the concept proposed by [3], if $P(N^y_{t+1} = 1|Y_{t+1}, \Phi_{[-N]})$ is greater than a threshold α, then $N^y_{t+1} = 1$. The threshold α is chosen such that the number of jumps identified corresponds to the estimate of the jump intensity ρ_y.

2.2 Gibbs Sampler

Here we present the Gibbs Sampler algorithm to sample from the NGSVJ model's parameters.

Let $Y_n = \{y_t\}_{t=1}^n$, $\underset{\sim}{J} = \{J^y_t\}_{t=1}^n = \{\xi^y_{t+1}N^y_{t+1}\}_{t=1}^n$, $\underset{\sim}{\gamma} = \{\gamma_t\}_{t=1}^n$, $\underset{\sim}{\lambda} = \{\lambda_t\}_{t=1}^n$, $\underset{\sim}{\xi} = \{\xi^y_{t+1}\}_{t=1}^n$, $\underset{\sim}{N} = \{N^y_{t+1}\}_{t=1}^n$ and prior probability densities $\pi(\gamma), \pi(\mu_y), \pi(\sigma^2_y), \pi(\xi), \pi(\rho_y)$ are set for $\underset{\sim}{\gamma}, \mu_y, \sigma^2_y, \underset{\sim}{\xi}, \rho_y$. Then, a sample of size M from the joint posterior distribution $\pi(\mu, \underset{\sim}{\lambda}, \underset{\sim}{\gamma}, \mu_y, \sigma^2_y, \underset{\sim}{J}, \rho_y|Y_n)$ is drawn via Gibbs Sampler, as follows:

i Initialize $\mu^{(0)}, \underset{\sim}{\lambda}^{(0)}, \underset{\sim}{\gamma}^{(0)}, \mu^{(0)}_y, (\sigma^2_y)^{(0)}, \underset{\sim}{\xi}^{(0)}, \underset{\sim}{N}^{(0)}$ and $\rho^{(0)}_y$.
ii Set $j = 1$.
iii Sample $\mu^{(j)}|Y_n, \underset{\sim}{J}^{(j-1)}, \underset{\sim}{\lambda}^{(j-1)}, \underset{\sim}{\gamma}^{(j-1)}$ using FFBS algorithm.
iv Block sample $\underset{\sim}{\lambda}^{(j)}|Y_n, \mu^{(j)}, \underset{\sim}{J}^{(j-1)}, \underset{\sim}{\gamma}^{(j-1)}$ using algorithm described on Appendix A.
v Block sample $\underset{\sim}{\gamma}^{(j)}|Y_n, \mu^{(j)}, \underset{\sim}{J}^{(j-1)}, \underset{\sim}{\lambda}^{(j)}$ as in Eq. (4).
vi Sample $\mu^{(j)}_y|\underset{\sim}{\xi}^{(j-1)}, (\sigma^2_y)^{(j-1)}$ as in Eq. (5).
vii Sample $(\sigma^2_y)^{(j)}|\underset{\sim}{\xi}^{(j-1)}, \mu^{(j)}_y$ as in Eq. (6).
viii Block sample $\underset{\sim}{J}^{(j)}|Y_n, \mu^{(j)}, \underset{\sim}{\lambda}^{(j)}, \underset{\sim}{\gamma}^{(j)}, \mu^{(j)}_y, (\sigma^2_y)^{(j)}$ by
  a Block sample $\underset{\sim}{\xi}^{(j)}|Y_n, \mu^{(j)}, \underset{\sim}{\lambda}^{(j)}, \underset{\sim}{\gamma}^{(j)}, \mu^{(j)}_y, (\sigma^2_y)^{(j)}$ as in Eq. (7).
  b Block sample $\underset{\sim}{N}^{(j)}|Y_n, \mu^{(j)}, \underset{\sim}{\lambda}^{(j)}, \underset{\sim}{\gamma}^{(j)}, \underset{\sim}{\xi}^{(j)}$ as in Eq. (11).
ix Sample $\rho^{(j)}_y|\underset{\sim}{J}^{(j)}$ as in Eq. (10).
x Set $j = j + 1$.
xi If $j \leq M$, go to iii, otherwise stop.

Since all full conditional posterior distribution have closed-form, only Gibbs Sampler steps are used.

2.3 Model diagnostics

The approach to compare different specifications for model parameters is the BIC and DIC criteria, defined by:

$$BIC = -2\,ln(\hat{L}) + k \times ln(n), \tag{12}$$
$$DIC = D(\underset{\sim}{y}, \bar{\Phi}) + 2pD, \tag{13}$$
$$pD = \bar{D}(\underset{\sim}{y}, \Phi) - D(\underset{\sim}{y}, \bar{\Phi}). \tag{14}$$

Where, for BIC, $\hat{L}$ is the maximized value of the likelihood function for the model, k is the number of parameters evaluated and n is the sample size, and, for DIC, statistical deviance $D(\underset{\sim}{y}, \Phi)$ is defined as:

$$D(\underset{\sim}{y}, \bar{\Phi}) = -2\,ln\left(p(\underset{\sim}{y}|\bar{\Phi})\right) \tag{15}$$

for data y, model parameters Φ and its posterior mean $\bar{\Phi}$. The posterior mean deviance is given by:

$$\bar{D}(\underset{\sim}{y}, \Phi) = E\left[D(\underset{\sim}{y}, \Phi)|\underset{\sim}{y}\right]. \tag{16}$$

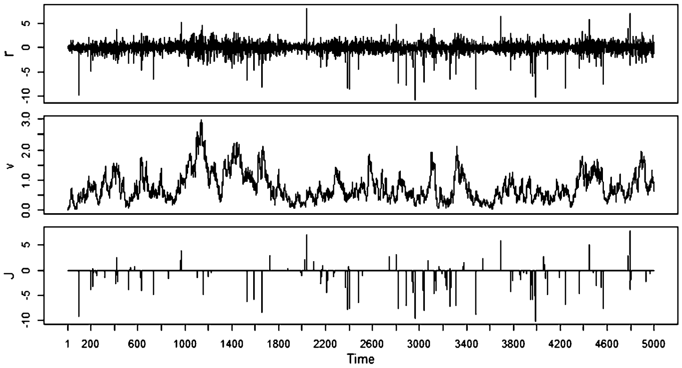

Figure 1. Simulation Study: Simulated realization ($n = 5{,}000$).

3. SIMULATION

To illustrate the performance of the NGSVJ, we apply the method to synthetic data from the model proposed by [21], and compare it to the performance of non-Gaussian State Space Model (NGSSM) in order to highlight the effect of including jumps on absorbing the abnormal impacts on observations and preventing them to propagate to volatility. To generate the volatility, we use:

$$
\begin{aligned}
v_t = v_{t-1} + \kappa(\theta - v_{t-1})\Delta + \rho\sigma_v\sqrt{v_{t-1}\Delta}\epsilon_{1,t} \\
+\sigma_v\sqrt{(1-\rho^2)v_{t-1}\Delta}\epsilon_{2,t},
\end{aligned}
\tag{17}
$$

where $\epsilon_{1,t}$ and $\epsilon_{2,t} \sim N(0,1)$. Synthetic data for returns is then generated from:

$$
r_t = N(\mu + \mathcal{J}_t, \gamma_t^{-1}v_t), \tag{18}
$$

$$
\mathcal{J}_t = N_t\xi_t, \tag{19}
$$

where the jump times, N_t are generated from a $Bernoulli(\rho_y)$, jump sizes ξ_t from $N(\mu_y, \sigma_y^2)$, and γ_t from $G(\frac{\nu}{2}, \frac{\nu}{2})$. Setup of parameters was: log-returns mean $\mu = 0.05$; jump probability $\rho_y = 0.015$; jump magnitude mean $\mu_y = -2.5$ and standard deviation $\sigma_y = 4$; variance mixture degrees of freedom $\nu = 30$; volatility components $\Delta = 1$, $\theta = 0.8$, $\kappa = 0.015$, $\sigma_v = 0.1$, $\rho = 0.4$, same used by [21].

The simulated time series consists of approximately 20 years of daily data ($n = 5{,}000$). Figure 1 shows one realization generated from the model. All codes for the model estimation were written in R software [18], using package Rcpp, available at The Comprehensive R Archive Network (CRAN). Machine specifications are Intel Core i7-8700K 3.70 GHz processor, 16 GB RAM, using a 64-bit Windows 10 Operating System. MCMC retained 20,000 samples after 300,000 iterations, burn-in of 60,000 and lag of 11 iterations.

For the NGSVJ model, Table 1 shows a grid with different values of ν that was made in order to choose the value that gives better adjustment according to DIC criteria. Smallest

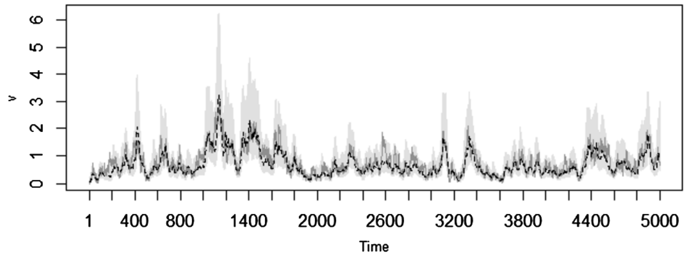

Figure 2. Simulation Study: Posterior estimates of instantaneous volatility v_t for simulated data. True volatility series shown in solid gray; posterior mean estimates λ_t^{-1} in dashed black; and 95% credibility interval is the light gray area.

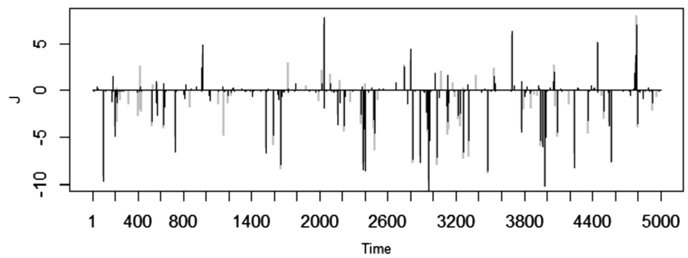

Figure 3. Simulation Study: Posterior estimates of instantaneous jumps $\mathcal{J}_t$ for simulated data. True values of the time series shown in gray; posterior mean estimates J_t in black.

DIC was obtained with $\nu = 30$, which matches the degrees of freedom used to generate simulated data.

Table 1. Grid analysis for ν

ν	5	15	30	60
DIC	13,572	12,197	**12,139**	12,145

Figure 2 shows the time series of true instantaneous volatility v_t, together with the estimated volatility λ_t^{-1}. The NGSVJ is able to closely track the latent state. Almost every point from the true volatility is inside the 95% credibility interval, even when the estimate mean slightly deviates from the true value. Figure 3 shows the time series of true instantaneous jumps $\mathcal{J}_t$, together with the estimated jumps J_t. Note that the jumps represent moments of punctual abnormal returns, caused by the market's speculative movements. The NGSVJ is able to catch most of the simulated jumps, together with their magnitudes. Jump points that are not captured by the jump component of the model are propagated to volatility (λ^{-1}) or heavy tail (γ^{-1}) components.

Table 2 presents the model estimates for each static parameter. The NGSVJ model is able to get estimates very close to true parameters. Figure 4 compares posterior estimates of instantaneous volatility on NGSVJ and NGSSM

Table 2. Posterior estimate of NGSVJ and NGSSM (without jumps) static parameters for simulated daily returns ($n = 5{,}000$)

		NGSVJ			NGSSM		
	True	Mean	SD	RMSE	Mean	SD	RMSE
μ	0.05	0.0529	0.0024	0.0037	0.0187	0.0012	0.0313
ρ_y	0.015	0.01544	0.0022	0.0024	-	-	-
μ_y	-2.5	-2.2799	0.5648	0.6245	-	-	-
σ_y	4	4.4445	0.4016	0.7475	-	-	-
$\log \mathcal{L}$	-		$-5{,}783$			$-6{,}322$	
BIC	-		**11,634**			12,669	
DIC	-		**12,139**			13,938	

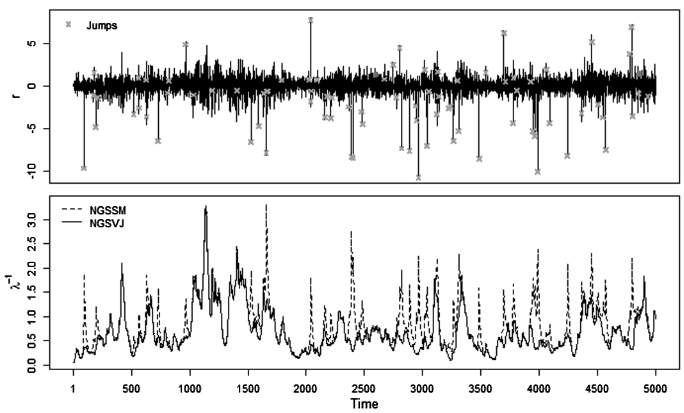

Figure 4. Simulation Study: Top graph shows posterior estimates of main instantaneous jumps $\mathcal{J}_t$ for simulated data together with simulated returns. Bottom graph shows posterior estimates for instantaneous volatility λ_t^{-1}. NGSVJ mean estimates in solid line and NGSSM (without the jump structure) mean estimates in dashed line.

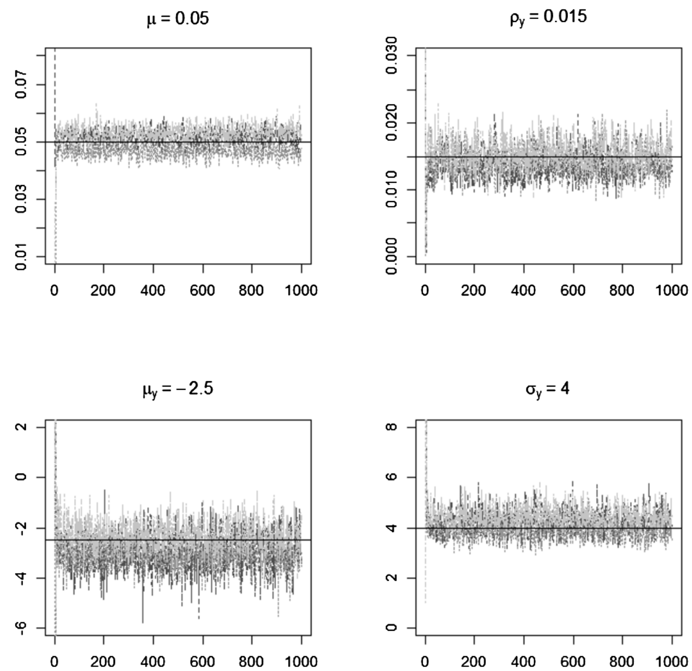

Figure 5. MCMC chain convergence: trace plot for the static parameters of the model for simulated time series using three different initial values. Each line represents a chain.

(without jumps) models. It can be seen that the jump component absorbs the impact of abnormal returns so that they do not propagate to volatility measure. Also, BIC and DIC values are smaller for the NGSVJ model than NGSSM (without the jump component), indicating a better fit, despite having more parameters to estimate.

Another advantage of the model, as can be seen in Figure 5, is that it rapidly achieves the converges due to its automatic and simple sampling structure. The Geweke test was also performed to check convergence of MCMC chains, to confirm that there is no significative change on chain's mean from the first 10% and the last 50% samples, and results are in Table 3. This allows the model to be used in on-line day-trade operations to estimate market volatility and orient day-trade arbitrage strategies.

Similar results are obtained across alternative simulation scenarios, using different initial values for the parameters. As the sample size n grows very large, the dimensionality of the model raises exponentially, since there are four dynamic parameters to be estimated: volatility, λ^{-1}; mixture component, γ^{-1}; jump times, N; and jump size, ξ. Thus, its usage is limited by concerns such as computational time and available computer memory.

One way to deal with larger data is to reduce the sample size and the number of iterations made by the MCMC algorithm, taking advantage of the model's inferential procedure that enables fast convergence due to its automated characteristic via Gibbs sampler. As seen in Figure 5, convergence was achieved on the first 200 iterations, regardless of the initial values for the parameters.

4. MODEL APPLICATIONS

In this section we show two applications of the NGSVJ model and compare it to the NGSSM proposed by [6] and SV model proposed by [8], implemented on R's stochvol package available on CRAN, in order to attest its efficiency. Since we

Table 3. MCMC chain convergence: Geweke test statistic for the static parameters of the model

	μ	ρ_y	μ_y	σ_y
Z-score	−0.6212	0.6731	−0.425	−0.0173
p-value	0.2672	0.2504	0.3354	0.4930

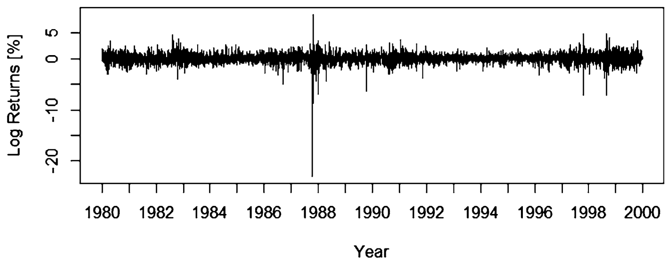

Figure 6. Log-returns of the S&P 500 Index from January 2, 1980, to December 31, 1999 ($n = 5{,}055$).

cannot guarantee full reproduction of models in [5] and [21], which would be ideally natural competitors to NGSVJ, we chose to use models that have codes already available and widely known on literature. The first application deals with the estimation of S&P500 volatility, as done by [5, 21]. The second is an application to intra-day Brent Crude future returns, to estimate volatility before the next minute return information arrives, in order to present the model as a tool for strategy making in a real decision problem, while the user is operating at the market.

4.1 S&P500 index data

The NGSVJ is applied to stock market index data and results will be compared to the NGSSM proposed by [6] and SV model proposed by [8] for stock returns data, which do not include jumps on returns.

The dataset contains S&P 500 stock index returns from January 2, 1980, to December 31, 1999. Excluding weekends and holidays, there are 5,055 daily observations for the S&P500. Table 4 provides descriptive statistic for the log-returns, scaled by 100.

Table 4. S&P500 log-returns [×100%] descriptive statistics

Descriptive Statistic	Value
Sample size	5,055
Mean	0.05205
Variance	0.9978
Skewness	−2.6357
Kurtosis	63.0710
Min	−22.8997
Max	8.7089

4.1.1 Parameters specification

When applying on stock market index data, a grid analysis was made using a sample of the last 500 observations in order to choose ν. Since the model is relatively robust for this parameter, the same value was used in all applications. Based on grid analysis conclusion that suggests $\nu = 30$, a G(15,15) prior distribution was specified for γ_t in order to obtain the Student's t_{30}-errors for the observation and system disturbances. Recall that we avoid appealing to Metropolis steps in order to estimate ν, in order to keep an automatic procedure of sampling via Gibbs sampler. The threshold was fixed at $\alpha = 0.7$ and the discount factor ω fixed at 0.9. For the mean components, μ and μ_y, N(0,100) priors were specified and for σ_y^2 a IG(0.1,0.1) prior. Also, $a_0 = 0.1$ and $b_0 = 0.1$, as suggested in [22], cited by [6]. A Beta(2,40) prior distribution was specified to ρ_y, as in [5]. For the NGSSM model, the same specifications were made, except that this model does not include the jump components.

The results were obtained with a 300,000 iteration chain, a burn-in of 60,000 observations, with a lag of 11 observations, resulting in 20,000 samples. MCMC chains convergence was verified through graphs and the Geweke diagnostic test, using R package *coda* [14]. All programming was done in the R software [18], using the Rcpp package.

4.1.2 Results

Table 5 shows model estimates for each of the static parameters. For NGSVJ, as observed by [5], it is possible to see that jumps in returns are infrequent, since jump probability ρ_y is small, but have a large magnitude, as can be seen by the magnitude of jump size mean, μ_y. BIC and DIC criteria favor NGSVJ over NGSSM and SV models, which indicates the former has a better fit to the data. Computational time for NGSVJ is close to SV model, which makes it competitive since it includes the jump structure that benefits from the automatic inference procedure to boost its speed. The inclusion of jumps makes NGSVJ 55.6% slower than the NGSSM model, but it has a better fit and a more precise estimate for volatility since abnormal returns are captured by the jump component.

Figure 7 shows the posterior mean estimates of instantaneous square root volatility $\lambda^{-1/2}$ for NGSVJ, NGSSM and SV models. As expected, the NGSVJ model estimates a lower magnitude volatility measure, since part of the log-returns variation is absorbed by the jump component, so they do not propagate to volatility as an increase on risk.

Figure 8 shows the posterior mean estimates of instantaneous volatility λ^{-1} for NGSVJ with the 95% credibility interval. Also, a zoom into two specific moments known as market crisis: the Black Monday (1987) and the Asian/Russian financial crisis (1997, 1998). The solid line

Table 5. Posterior inference of static parameters for NGSVJ, NGSSM and SV models for S&P500 daily returns. Computational time (CT) is given in seconds

	NGSVJ		NGSSM		SV	
	Mean	SD	Mean	SD	Mean	SD
μ	0.0616	0.0020	0.0522	0.0012	0	-
ρ_y	0.0042	0.0012	-	-	-	-
μ_y	-2.4598	1.4302	-	-	-	-
σ_y	5.2793	1.2598	-	-	-	-
$\log \mathcal{L}$	$\mathbf{-5{,}972}$		$-6{,}086$		$-8{,}468$	
BIC	**12,012**		12,197		12,206	
DIC	**12,338**		12,763		21,719	
CT	738		**474**		753	

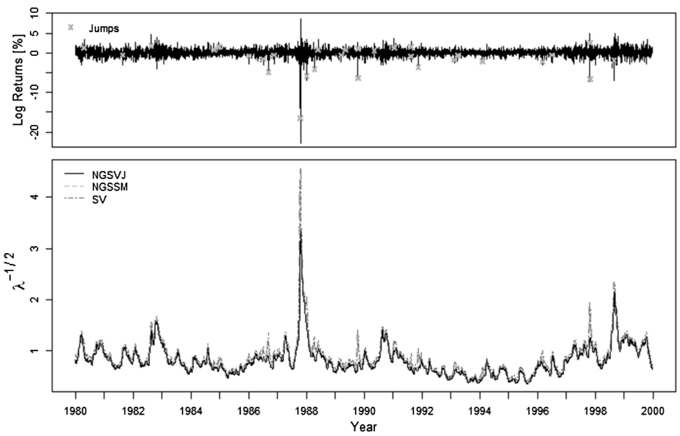

Figure 7. Top graph shows posterior estimates of main instantaneous jumps J_t for S&P500 together with log-returns [$\times$100%]. Bottom graph shows posterior estimates of instantaneous square-root of volatility $\lambda_t^{-1/2}$ for S&P 500 Index data. NGSVJ mean estimates in solid line; NGSSM mean estimates in dashed line; and SV mean estimates in dotdashed line.

represents the posterior mean and the gray area indicates the 95% percentile credibility interval for spot volatility. Results are consistent with [5] findings, with volatility peaks occurring at the same time.

Figure 9 provides jump sizes and probabilities for each observation. Before moments of higher volatility, it is possible to observe an increase in jump sizes and probabilities, when compared to periods with lower volatility, thus, evidencing that such moments are preceded by speculative movements, captured in the model as jumps.

4.2 Intra-day returns data

The NGSVJ model is applied to intra-day returns of Brent Crude Futures so that a market agent can use this information to choose strategies on day-trade operations. In order to be used during market operations, the model must be able to deliver reliable and fast results. The proposed

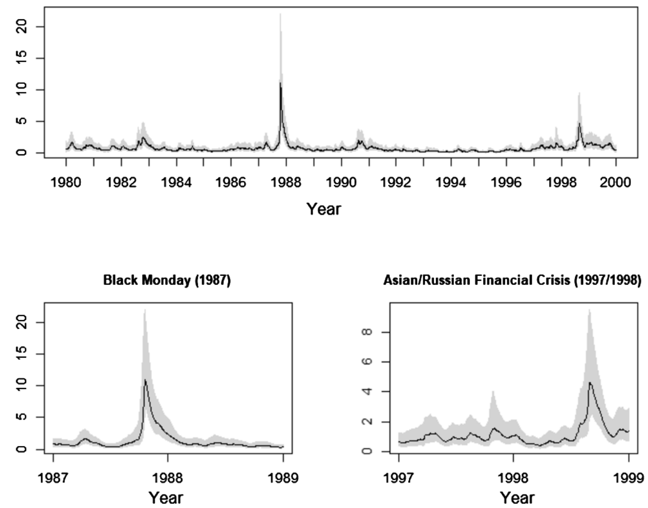

Figure 8. Posterior estimates of instantaneous volatility λ_t^{-1} for S&P 500 Index data. NGSVJ mean estimates in solid line; and 95% credibility interval is the gray area.

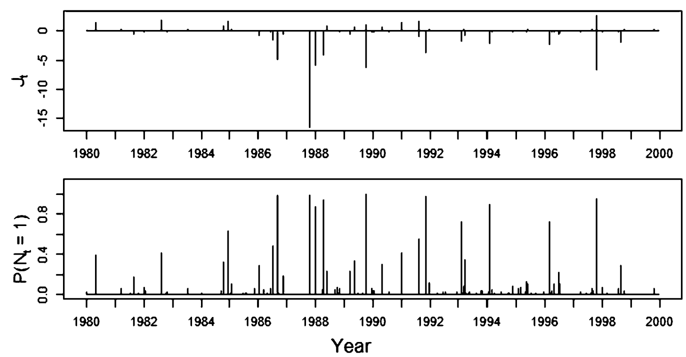

Figure 9. Posterior estimates of jump times, N, and jump sizes, J, for S&P 500 Index data from NGSVJ model.

model performance will be compared to NGSSM and SV models.

The dataset contains Brent Crude futures, ICE:BRN, log returns from August 13, 2018, to August 17, 2018, in a total of 6,241 minute observations. Table 6 provides summary statistics for the log returns scaled by 100.

Table 6. Descriptive statistics of ICE:BRN log-returns [$\times$100%]

Descriptive Statistic	Value
Sample size	6,241
Mean	-0.00026
Variance	0.00204
Skewness	-1.6751
Kurtosis	42.3833
Min	-0.95748
Max	0.39991

4.2.1 Parameters specifications

The NGSVJ setup considers the same specifications defined on Section 4.1.1. The results were obtained with a 10,000 iteration chain, burn-in of 6,000 observations with a lag of 2 samples, resulting in 1,667 samples. MCMC chains convergence was verified through graphs and the Geweke diagnostic test, using R package *coda* [14]. This specification takes advantage of the fast convergence which enables the model to give results in about 30 seconds so that a strategy decision can be taken before the next minute observation arrives.

4.2.2 Results

Table 7 shows model estimates for each of the static parameters. BIC and DIC criteria strongly favor NGSVJ against NGSSM and SV models, since chain convergence is not achieved for all parameters on the latter model. Both NGSVJ and NGSSM models are able to deliver reliable results in less than one minute, before the next information arrives, but the former has the advantage of including jumps in returns, thus, providing a better fit to the data and a more precise volatility estimate.

Table 7. Posterior inference of static parameters for NGSVJ, NGSSM and SV models for ICE:BRN minute returns. Computational time (CT) is given in seconds

	NGSVJ		NGSSM		SV	
	Mean	SD	Mean	SD	Mean	SD
μ	0.0001	0.0001	-0.0001	0.0000	0	-
ρ_y	0.0087	0.0015	-	-	-	-
μ_y	-0.02305	0.0334	-	-	-	-
σ_y	0.0498	0.0119	-	-	-	-
$\log \mathcal{L}$	**12737**		12407		4235	
BIC	**-25405**		-24787		-8435	
DIC	**-24919**		-23747		-8409	
CT	31.83		**20.60**		28.64	

Figure 10 shows trace plot of the final sample and Table 8 presents the results of the Geweke test for the static parameters of NGSVJ model. As seen on the simulation study, the model is able to rapidly achieve convergence due to its automatic and simple sampling structure. This property makes NGSVJ both fast and reliable enough to provide results in such a time frame that enables the user to make strategic decisions.

Figure 11 shows a trace plot for static parameters for ICE:BRN time series for the SV model, provided by package stochvol, for the final sample. Despite being fast enough to bring results on the required time frame, the SV model was not able to achieve convergence for all parameters. It can be seen that ρ and σ_y parameters have still not converged after six thousand iterations, hence such results are not reliable to be used for taking strategic decisions.

Figure 12 shows the posterior mean estimates of instantaneous square root volatility, $\lambda^{-1/2}$, for NGSVJ, NGSSM,

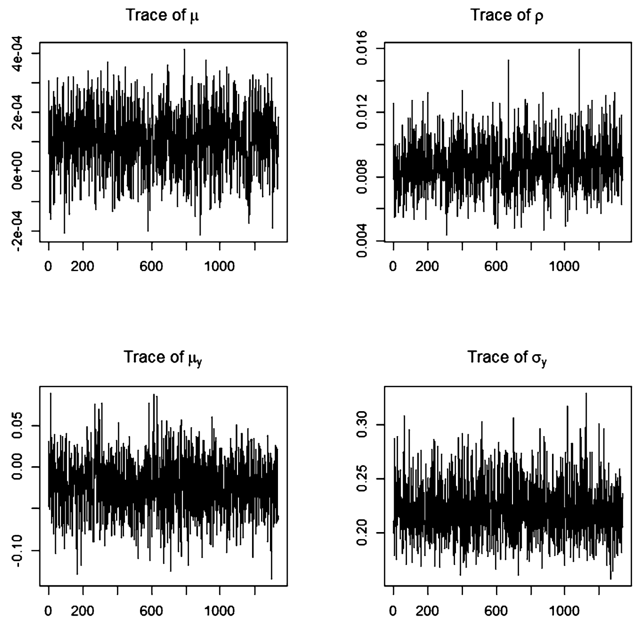

Figure 10. MCMC chain convergence: trace plot for static parameters for ICE:BRN time series for NGSVJ model. Convergence is achieved for all static parameters.

Table 8. MCMC chain convergence: Geweke test for static parameters for ICE:BRN time series for NGSVJ model. Convergence is achieved for all static parameters

	μ	ρ_y	μ_y	σ_y
Z-score	-1.248	1.278	-0.4858	-0.7208
p-value	0.1060	0.1006	0.3136	0.2355

and SV models. As observed before, the NGSVJ model estimates a lower magnitude volatility measure, since part of the log-returns variation is absorbed by the jump component, so they do not propagate to volatility as an increase on risk.

Figure 13 shows the posterior mean estimates of instantaneous volatility λ^{-1} for NGSVJ with the 95% credibility interval, jump sizes and probabilities for each observation. The solid line represents the posterior mean and the gray area indicates the 95% percentile credibility interval for spot volatility.

5. DISCUSSION

The NGSVJ model was able to capture speculative movements in the market through the jump components and detect periods with increased market risk through the volatility component. Results obtained by applying the model to the S&P 500 return series are consistent with [5] findings since volatility peaks occurred at the same time. Historical events of known volatility effects on financial markets, such as Black Monday and Asian/Russian Finan-

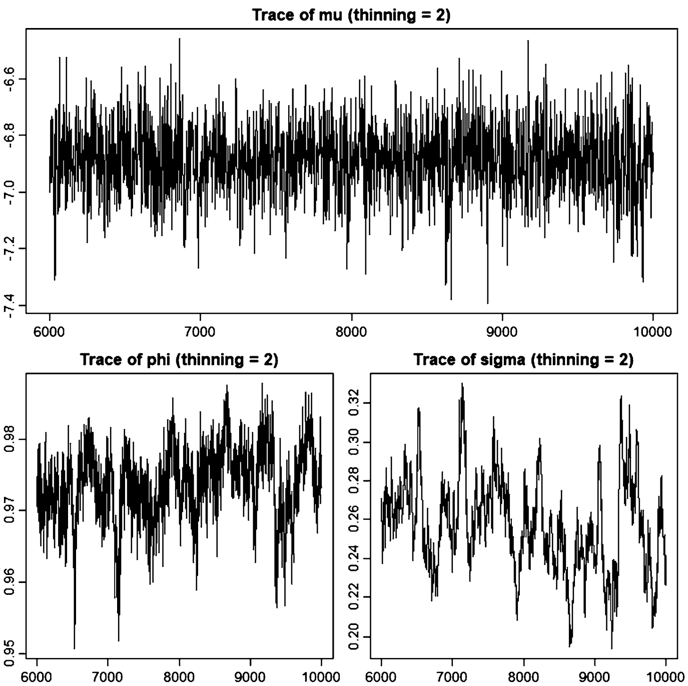

Figure 11. MCMC chain convergence: trace plot for static parameters for ICE:BRN time series for SV model, provided by package stochvol, after burn-in. Convergence is not achieved for all static parameters.

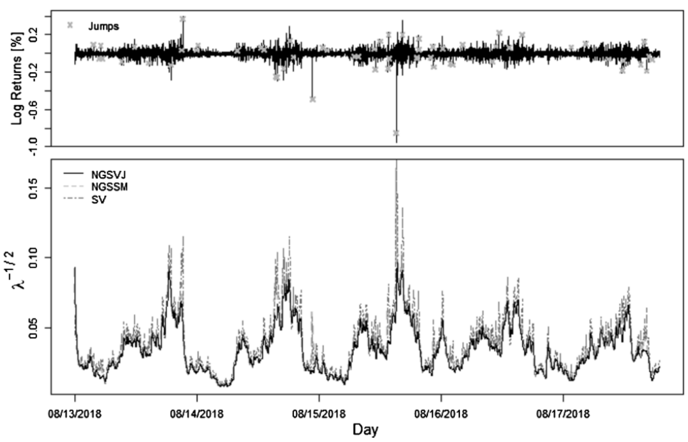

Figure 12. Top graph shows posterior estimates of main instantaneous jumps J_t for ICE:BRN together with log-returns [$\times$100%]. Bottom graph shows posterior estimates of instantaneous square-root of volatility $\lambda_t^{-1/2}$ for ICE:BRN data. NGSVJ mean estimates in solid line; NGSSM mean estimates in dashed line; and SV mean estimates in dotdashed line.

cial Crisis, are detected by the model precisely. Also, the inclusion of jumps improves the model fit when compared to NGSSM and SV models.

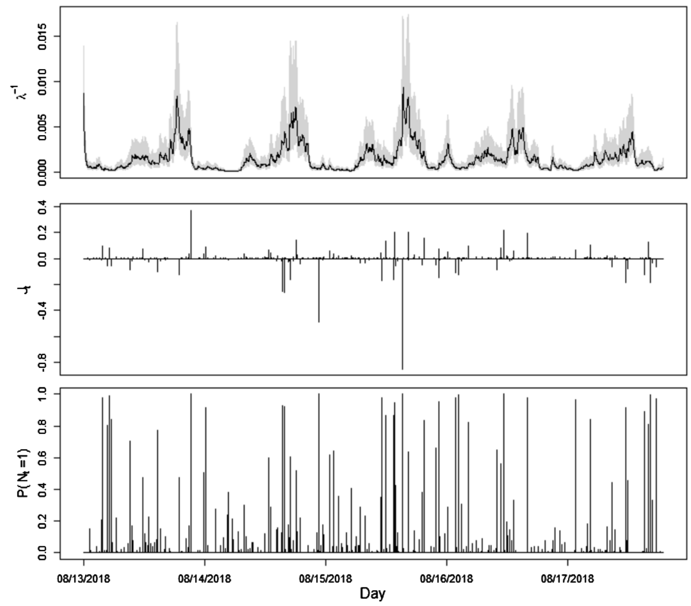

Figure 13. Top: Posterior mean estimates of instantaneous volatility λ_t^{-1} for ICE:BRN data. NGSVJ mean estimates in solid line, and 95% credibility interval is the gray area; Middle: Posterior mean estimates for jump sizes, J_t; Bottom: Posterior estimates for jump times N_t.

The most notable advantage of using NGSVJ is its computational simplicity. Other benefits include using the NGSSM structure and automatic sampling process for parameters which allow sampling the volatility in block via Gibbs sampler, mitigating convergence issues. This structure allows model parameters to achieve convergence with less MCMC iterations, providing fast and reliable estimates and allowing the model to be used in practical situations, such as taking decisions on intra-day operations arbitrage strategies.

Using Gibbs Sampler to draw a sample from conditional posterior distributions is computationally cheaper than recurring to Metropolis based algorithms. Since Metropolis is an accept-reject algorithm, it can take several steps until a full representative sample is obtained in order to make suitable statistical inferences, unless a very good proposal distribution is given to the algorithm. On using Gibbs Sampler, a representative sample can potentially be obtained in fewer steps. This is especially relevant when dealing with small time frames, as in intra-day operations.

As the model is built over the dynamic linear model with scale mixtures proposed by [6] an exact sample of volatility parameter λ_t^{-1} is drawn. The three main advantages that come from this process are: there is no need to make approximations in order to estimate volatility; the volatility is sampled in blocks from the proposed model; and there is no need to appeal to Metropolis based algorithms.

Another advantage is the model flexibility. It can be adapted to include jumps, covariates, heavy tails and different distributions can be adopted based on the mixture component used. In this work, Student-t distribution was obtained through a Gamma mixed component, but other distributions can also be used to give satisfactory results. Also, it was shown that the inclusion of jumps raises the performance of the model when compared to the NGSSM model.

The inclusion of jumps in the model reduces substantially the volatility estimate. This would have several impacts on risk analysis, since less volatility indicates less risk, in other words, knowing that some event was a jump, or tail event, and not a recurrent market event would mean that such asset is still a safe bet. Usually, an increase of jump frequency is observed near the occurrences of market anomalies, such as Black Monday and the Asian and Russian financial crisis, which may indicate that it can also be used for predicting near future increases on market risk.

For day-to-day operations, the automatic model is effective and can be used in order to estimate market volatility, which can be used for options pricing, VaR calculations, measuring market regimes, etc.

For future works it is intended to extend the model to the multivariate case, where, instead of analyzing one asset individually, an asset portfolio risk is analyzed as a whole. Another possible extension is the inclusion of jumps in volatility, as suggested by [5], without having to appeal to Metropolis based algorithms, in order to keep the inferential procedure of the model fast and accessible. Some other possibilities include working with a skew heavy-tailed distribution, as in [12] but still maintaining the block sampling so that it allows the capture of the leverage effect and correlations between mean and volatility.

APPENDIX A. APPENDIX SECTION

This appendix shows how the posterior sample from λ is drawn [6].

The joint distribution of $p(\lambda|Y_n, \Phi_{[-\lambda]})$ has density

$$p(\lambda_n|\Phi_{[-\lambda]}, Y_n) \prod_{t=1}^{n-1} p(\lambda_t|\lambda_{t+1}, \Phi_{[-\lambda]}, Y_t) p(\Phi_{[-\lambda]}|Y_n) \tag{20}$$

where the distribution of $(\lambda_t|\lambda_{t+1}, \Phi_{[-\lambda]}, Y_t)$ is given by

$$\lambda_t - \omega\lambda_{t+1}|\lambda_{t+1}, \Phi_{[-\lambda]}, Y_t \sim G\left((1-\omega)a_t, b_t\right), \forall t \geq 0, \tag{21}$$

where a_t and b_t are the filtering parameters. The on-line or updated distribution $\lambda_t|\Phi_{[-\lambda]}, Y_t$ is given by:

$$p(\lambda_t|\Phi_{[-\lambda]}, Y_{t-1}) \sim G(\omega a_{t-1}, \omega b_{t-1}) \tag{22}$$

$$p(\lambda_t|\Phi_{[-\lambda]}, Y_t) \propto p(\lambda_t|\Phi_{[-\lambda]}, Y_{t-1}) p(Y_t|\lambda_t, \Phi_{[-\lambda]}) \tag{23}$$

$$p(\lambda_t|\Phi_{[-\lambda]}, Y_t) \sim G\left(\omega a_{t-1} + \frac{1}{2}, \omega b_{t-1} + \gamma_t \frac{(Y_t - \mu - J_t)^2}{2}\right) \tag{24}$$

Based on Theorem 2 of [6] and with a sample of $p(\Phi_{[-\lambda]}|Y_n)$, an exact sample of the joint distribution $(\lambda|Y_n, \Phi_{[-\lambda]})$ can be obtained following the algorithm:

1. set $t = n$ and sample $p(\lambda_n|\Phi_{[-\lambda]}, Y_n)$;
2. set $t = t - 1$ and sample $p(\lambda_t|\lambda_{t+1}, \Phi_{[-\lambda]}, Y_t)$;
3. if t > 1, go back to step 2; otherwise, the sample of $(\lambda_1, ..., \lambda_n|\Phi_{[-\lambda]}, Y_n)$ is complete.

This procedure allows the implementation of the algorithm, described in Section 2.2, step *iv* and enables us to obtain an exact sample from the smoothed distribution of the states conditioned on other parameters.

Received 5 October 2018

REFERENCES

[1] Abanto-Valle, C. A., Lachos, V. H., and Dey, D. K. (2015). Bayesian estimation of a skew-student-t stochastic volatility model. Methodology and Computing in Applied Probability, 17(3), 721–738. MR3377857
[2] Bandi, F. M., and Renò, R. (2018). Nonparametric stochastic volatility. Econometric Theory, 1–49. MR3873848
[3] Brooks, C., and Prokopczuk, M. (2013). The dynamics of commodity prices. Quantitative Finance, 13(4), 527–542. MR3048016
[4] Chib, S., Nardari, F., and Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. Journal of Econometrics, 108(2), 281–316. MR1894758
[5] Eraker, B., Johannes, M., and Polson, N. (2003). The impact of jumps in volatility and returns. The Journal of Finance, 58(3), 1269–1300
[6] Gamerman, D., dos Santos, T. R., and Franco, G. C. (2013). A non-Gaussian family of state-space models with exact marginal likelihood. Journal of Time Series Analysis, 34(6), 625–645. MR3127211
[7] Kanaya, S., and Kristensen, D. (2016). Estimation of stochastic volatility models by nonparametric filtering. Econometric Theory, 32(4), 861–916. MR3530455
[8] Kastner, G. (2016). Dealing with stochastic volatility in time series using the R package stochvol. Journal of Statistical Software, 69(5), 1–30
[9] Lee, S. S., and Mykland, P. A. (2007). Jumps in financial markets: A new nonparametric test and jump dynamics. The Review of Financial Studies, 21(6), 2535–2563.
[10] Li, H., Wells, M. T., and Yu, C. L. (2006). A Bayesian analysis of return dynamics with Lévy jumps. The Review of Financial Studies, 21(5), 2345–2378.
[11] Merener, N. (2016). Concentrated production and conditional heavy tails in commodity returns. Journal of Futures Markets, 36(1), 46–65.
[12] Nakajima, J., and Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. Computational Statistics & Data Analysis, 53(6), 2335–2353. MR2665832
[13] Omori, Y., Chib, S., Shephard, N., and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. Journal of Econometrics, 140(2), 425–449. MR2408914
[14] Plummer, M., Best, N., Cowles, K. and Vines, K. (2006). CODA: convergence diagnosis and output analysis for MCMC. R News, 6(1), 7–11.
[15] Prado, R., and West, M. (2010). Time series: modeling, computation, and inference. CRC Press. MR2655202
[16] Raftery, A. E., and Lewis, S. (1991). How many iterations in the Gibbs sampler? Washington Univ Seattle Dept of Statistics.
[17] Schwartz, E., and Smith, J. E. (2000). Short-term variations and long-term dynamics in commodity prices. Management Science, 46(7), 893–911.

[18] R x64 3.4.3 [Computer software]. (2018). Retrieved from http://www.r-project.org.

[19] Todorov, V. (2009). Variance risk-premium dynamics: The role of jumps. The Review of Financial Studies, 23(1), 345–383.

[20] Triantafyllopoulos, K. (2008). Multivariate stochastic volatility with bayesian dynamic linear models. Journal of Statistical Planning and Inference, 138(4), 1021–1037. MR2384504

[21] Warty, S. P., Lopes, H. F., and Polson, N. G. (2018). Sequential Bayesian learning for stochastic volatility with variance-gamma jumps in returns. Applied Stochastic Models in Business and Industry, 34(4), 460–479. MR3845895

[22] West, M., and Harrison, J. (2006). Bayesian forecasting and dynamic models. Springer Science & Business Media. MR1020301

[23] Wit, E., Heuvel, E. V. D., and Romeijn, J. W. (2012). 'All models are wrong...': an introduction to model uncertainty. Statistica Neerlandica, 66(3), 217–236. MR2955417

Arthur T. Rego
Department of Statistics
Universidade Federal de Minas Gerais
Belo Horizonte, Minas Gerais, 31270-901
Brazil
E-mail address: arthur.tarso@outlook.com

Thiago R. dos Santos
Department of Statistics
Universidade Federal de Minas Gerais
Belo Horizonte, Minas Gerais, 31270-901
Brazil
E-mail address: thiagords@est.ufmg.br