

Stock return volatility, heavy tails, skewness and trading volume: A Bayesian approach

C. A. Abanto-Valle^{†1}, Dipak K. Dey[§] and V. H. Lachos[‡]

[†]Department of Statistics, Federal University of Rio de Janeiro, Brazil

[§] Department of Statistics, University of Connecticut, USA

[‡] Department of Statistics, Campinas State University, Brazil

Abstract

This article considers the relationship between stock return volatility and trading volume by using the modified mixture model (MMM) framework proposed by [Andersen \(1996\)](#) and [Mahieu and Bauer \(1998\)](#). We assume that the return shock has a skew-Student-t distribution with unknown degrees of freedom. This allows a parsimonious, flexible treatment of skewness and heavy tails in the conditional distribution of returns. We propose to construct an algorithm based on Markov chain Monte Carlo (MCMC) simulation methods to estimate all the parameters in the model using a Bayesian approach. A clear advantage of MCMC methods is that estimates of volatility are readily available for use in, for example, dynamic portfolio allocation and option pricing applications. The series of returns and trading volume of four common stocks of the New York stock exchange (NYSE) are analyzed.

Keywords: Markov chain Monte Carlo, nonlinear and non-Gaussian state space models, skew-Student-t, stochastic volatility, trading volume.

1 Introduction

The relationship between returns and trading volume has interested financial economists and analysts for a number of years. [Clark \(1973\)](#) started the discussion by presenting the simplest and intuitively version of the mixture of distributions hypothesis (MDH), which assumes a joint dependence of volatility and volume on the underlying information flow variable, i.e., price movements

¹Correspondence to: Carlos A. Abanto-Valle, Department of Statistics, Federal University of Rio de Janeiro, Caixa Postal 68530, CEP: 21945-970, RJ, Brazil. E-mail: cabantovalle@im.ufrj.br

and the trading volume changes are caused primarily by the arrival of new information and the volatility process that incorporates this information into market prices. In an extensive review of the literature, [Karpoff \(1987\)](#) cites several reasons why the price-volume relationship is important and observes that much of the previous research has been about the contemporaneous relationship using correlations. [Gallant et al. \(1992\)](#) also point out that previous empirical work on the price-volume relationship has focused primarily on the contemporaneous relationship between price changes and volume. Although much of the empirical research documents a positive correlation between trading volume and return volatility, the evidence on whether the observed relation can be reconciled with the predictions of market microstructure theory is mixed (see, for example [Tauchen and Pitts, 1983](#); [Richardson and Smith, 1994](#); [Foster and Viswanathan, 1995](#)).

The literature on MDH can be classified in two groups. The first one, under the assumption of MDH, focuses on estimation of the model parameters and latent variables to evaluate the goodness of fit with respect to real data (see, for example [Clark, 1973](#); [Epps and Epps, 1976](#); [Tauchen and Pitts, 1983](#); [Harris, 1987](#); [Andersen, 1996](#); [Liesenfeld, 1998](#)). The second one concentrates on the properties of the observed series, relying on an observable (realized) measure of volatility ([Bollerslev and Jubinski, 1999](#); [Luu and Martens, 2003](#)).

A first approach to merge the insights of the MDH with those of the market microstructure theory is the empirical model of daily return-volume relationship developed by [Andersen \(1996\)](#). He combines several important features of these models - for instance an asymmetric information structure and the presence of liquidity or noise traders - with the MDH and the related concept of stochastic volatility. The resulting model, called the modified mixture model (MMM), is estimated with a dynamic first order autoregressive stochastic volatility process for the log of the latent rate of information arrival, by using the generalized method of moments. As in [Liesenfeld \(1998\)](#), the estimated measure of volatility persistence drops significantly compared with the univariate specifications for the return volatility. [Mahieu and Bauer \(1998\)](#) and [Watanabe \(2000\)](#) implemented the MMM from a Bayesian viewpoint using simulation techniques based on MCMC methods to estimate the parameters and the latent process. They find that consistent with the bivariate model's hypothesis volatility does not decrease but remains high in the bivariate case. Their results suggest

that the choice of the estimation technique could be important in testing the validity of the MMM.

A large literature in financial econometrics has documented stylized facts which are frequently found in stock and foreign exchange returns: skewness, heavy-tailedness and volatility clustering. These properties are crucial not only for describing the return distributions but also for asset allocation, option pricing, forecasting and risk management.

Stochastic volatility (SV) models were introduced in the financial literature for describing time varying volatilities (Taylor, 1982, 1986). Various extensions of the simple SV model with normal errors have been discussed in the literature. For instance, many empirical studies have shown strong evidence of heavy-tailed conditional mean errors in financial time series (see for example Chib et al., 2002; Jacquier et al., 2004). In this context, recently Abanto-Valle et al. (2010) extended the basic SV model by assuming the flexible class of scale mixtures of normal distributions. The empirical evidence on the presence of asymmetry in the distribution of financial returns is not as clear-cut even though asymmetry plays a non-trivial role in shaping economic decisions. Corrado and Su (1997) suggests that fat tails and asymmetry jointly determine the so-called “volatility smile” in option pricing using the Black-Scholes approach and that explicit account of them improve accuracy in option pricing. Peiro (1999) provides further evidence of asymmetry in returns, both from stock market indices and from individual assets. Further, Mittnik and Paoletta (2000) argue that skewness and heavy tails should be taken into account explicitly in Value-at-Risk forecasts. Cappuccio et al. (2006) found empirical evidence on asymmetry in financial returns using a simple stochastic volatility modeling both skewness and heavy tails assuming that the conditional distribution of returns is a skew-generalized error distribution.

In this article we propose to expand the conditional distribution of the returns used in Mahieu and Bauer (1998) by introducing the skew-student-t distribution (Branco and Dey, 2001; Azzalini and Capitanio, 2003) which allows taking into consideration simultaneously skewness and heavy-tailedness. Inference in the MMM with skew-Student-t errors is performed under a Bayesian paradigm via MCMC methods, which permits to obtain the posterior distribution of parameters by simulation starting from reasonable prior assumptions on the parameters. We simulate the log-volatilities, the shape and skewness parameters by using the block sampling algorithm (Shephard

and Pitt, 1997; Watanabe and Omori, 2004; Abanto-Valle et al., 2010, 2011) and the Metropolis-Hastings sampling, respectively.

The rest of the article is organized as follows: Section 2 shows a brief review about skew-normal (Azzalini, 1986) and skew-t distributions (Branco and Dey, 2001) and their properties. Section 3 presents the relation between stock return volatility and trading volume with the extended specification for the conditional distribution of the returns. Section 4 shows the Bayesian estimation procedure using MCMC methods. Section 5 presents an empirical application on the return and trading volume series for four common stocks of the NYSE. Finally, section 6 concludes with suggested possible extensions.

2 The univariate skew-normal and skew-t distributions

We start by giving an important notation that will be used throughout the paper and present a review of the univariate skew normal (SN) and skew-t (ST) distributions and a study of some related properties of those distributions.

A univariate random variable X is said to follow a skew-normal distribution, $X \sim \mathcal{SN}(\zeta, \omega^2, \lambda)$, with location, scale and asymmetry parameters given by ζ , ω^2 and λ , respectively, if the density of this distribution has the form

$$p(x | \zeta, \omega^2, \lambda) = \frac{2}{\omega} \phi\left(\frac{x - \zeta}{\omega}\right) \Phi\left(\frac{\lambda}{\omega}(x - \zeta)\right), \quad (1)$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are, respectively, the probability density function (pdf) and the cumulative distribution function (cdf) of the standard normal distribution. When $\lambda = 0$, the density in equation (1) becomes $\mathcal{N}(\zeta, \sigma^2)$ (see Azzalini, 2005, for a comprehensive review). In the next sections, we use the following stochastic representation of the SN distribution (Azzalini, 1986; Henze, 1986). Let $W \sim \mathcal{N}_{[0, \infty)}(0, 1)$ and $\varepsilon \sim \mathcal{N}(0, 1)$, independently, and let $\delta \in (-1, 1)$, where $\mathcal{N}_{[0, \infty)}(\cdot, \cdot)$ and $\mathcal{N}(\cdot, \cdot)$ indicate the truncated normal and normal distribution, respectively. The random variable X , defined by

$$X = \zeta + \omega \delta W + \omega \sqrt{1 - \delta^2} \varepsilon, \quad (2)$$

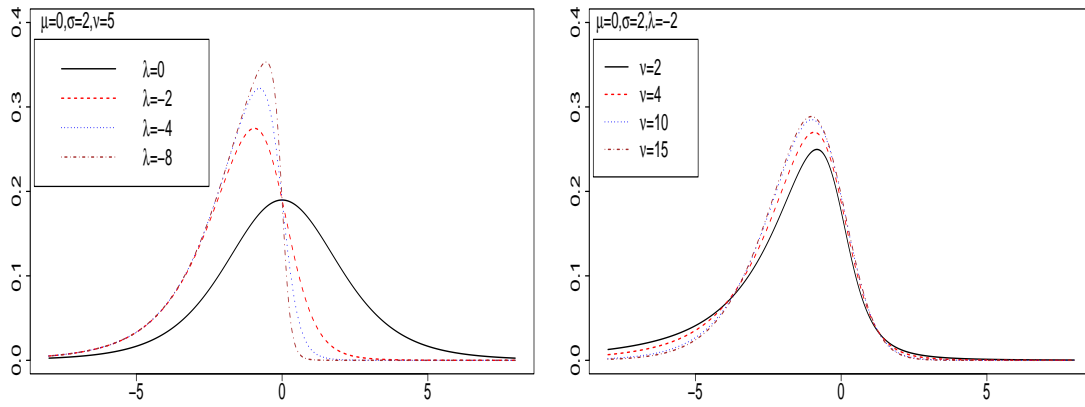


Figure 1: The skew-t distribution. Left: $\zeta = 0, \omega = 2, \nu = 5$ (fixed), $\lambda = 0, -2, -4, -8$. Right: $\zeta = 0, \omega = 2, \lambda = -2$ (fixed), $\nu = 2, 4, 10$ and 15 .

follows a univariate skew-normal distributions, that is, $X \sim \mathcal{SN}(\zeta, \omega^2, \lambda)$, where $\lambda = \delta/\sqrt{1 - \delta^2}$.

The kurtosis coefficient of a skew-normal distribution is restricted to the interval [3, 3.8692]. To achieve a higher degree of excess kurtosis, the skew-t distribution has been introduced by Branco and Dey (2001) and latter studied by Azzalini and Capitanio (2003). A univariate random variable X follows the scalar skew-t distribution, $X \sim \mathcal{ST}(\zeta, \omega^2, \lambda, \nu)$, if it has the following stochastic representation

$$X = \zeta + U^{-1/2} \omega \delta W + U^{-\frac{1}{2}} \omega (1 - \delta^2)^{\frac{1}{2}} \varepsilon, \quad (3)$$

where $W \sim \mathcal{N}_{[0, \infty)}(0, 1)$, $\varepsilon \sim \mathcal{N}(0, 1)$ and $U \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$ are independently distributed. The Gamma distribution $\mathcal{G}(a, b)$ is defined with density $p(u | a, b) = b^a u^{a-1} e^{-bu} / \Gamma(a)$. The pdf of X is then given by

$$f(X | \zeta, \omega^2, \lambda, \nu) = \frac{2}{\omega} t_\nu \left(\frac{x - \zeta}{\omega} \right) T_{\nu+1} \left(\lambda \omega^{-1} (x - \zeta) \sqrt{\frac{\nu + 1}{\nu + \omega^{-2} (x - \zeta)^2}} \right), \quad (4)$$

where $t_\nu(\cdot)$ and $T_\nu(\cdot)$ denote the pdf and cdf of a standard Student-t distribution with ν degrees of freedom. From (3), we have that

$$E(X) = \zeta + \sqrt{\frac{2}{\pi}} k_1 \omega \delta, \quad (5)$$

$$V(X) = \omega^2 k_2 - \frac{2}{\pi} k_1^2 \omega^2 \delta^2, \quad (6)$$

where $\delta = \lambda / \sqrt{1 + \lambda^2}$ and $k_m = E(U^{-m/2})$. $E(\cdot)$ and $V(\cdot)$ denote the expected value and variance, respectively. The skew-t nests the traditional symmetric Student's t distribution as a special case when $\lambda = 0$, and the conditional normal distribution as $\nu \rightarrow \infty$, and can capture left-tailed or negative skewness when $\lambda < 0$, and positive skewness when $\lambda > 0$.

To interpret the parameters (λ, ν) in relation to the skewness and heavy-tailedness, skew-t densities are plotted using several combinations of the parameter values in Figure 1 with ζ and ω fixed at 0 and 2, respectively. In Figure 1, left, the densities are drawn using $\lambda = 0, -2, -4, -8$ with ν fixed at 5. As mentioned, $\lambda = 0$ corresponds to a symmetric Student's t-density. A lower value of λ implies a more negative skewness or left-skewness as well as heavier tails. Figure 1, right, shows the densities for ν at 2,4,10 and 15 with λ fixed equal to -2. As ν becomes larger, the

density becomes less skewed and has lighter tails. Hence the skewness and heavy-tailedness are determined jointly by the combination of the parameter values of λ and ν .

3 The Model

Andersen (1996) develops an empirical return volatility-trading volume model using the theoretical framework of Glosten and Milgrom (1985). In his specification, the trading volume has two components which are directly related to informed and uninformed traders. The uninformed component is governed by a time invariant Poisson process with constant intensity m_0 , while the informed volume has a Poisson distribution with parameter which is a function of the information flow, that is $m_1 e^{h_t}$. An empirical version of the MMM of Andersen (1996), which was formulated by Mahieu and Bauer (1998), leads to the following specification:

$$y_t = e^{\frac{h_t}{2}} \varepsilon_t, \quad (7)$$

$$v_t | h_t \sim \mathcal{P}(m_0 + m_1 e^{h_t}), \quad m_0, m_1 > 0 \quad (8)$$

$$h_{t+1} = \mu + \varphi(h_t - \mu) + \sigma_\eta^2 \eta_t, \quad (9)$$

where y_t , v_t and h_t are respectively the compounded return, the trading volume and the log volatility on day t . $\mathcal{P}(\cdot)$ indicates the Poisson distribution. We assume that $|\varphi| < 1$, i.e., the log-volatility process is stationary and that the initial value $h_1 \sim \mathcal{N}(\mu, \frac{\sigma_\eta^2}{1-\varphi^2})$, ε_t and η_t are uncorrelated with normal distribution with zero mean and unit variance. In equation (8), m_0 reflects the uninformed component of trading volume and is related to liquidity traders. The remaining part of trading volume that is induced by new information is represented by $m_1 e^{h_t}$. The MMM defined by equations (7)-(9) will be denoted as SV-N-VOL. Note that the univariate stochastic volatility model (SV) used extensively in financial literature see (see Jacquier et al., 1994; Kim et al., 1998; Mahieu and Bauer, 1998; Abanto-Valle et al., 2010, among others) is specified by equations (7) and (9).

We modify the normality specification of the returns in (7) in order to capture heavy-tailedness and skewness features in the marginal distribution of random errors using the stochastic represen-

tation of the skew-Student-t in (3), as

$$y_t = (\zeta + \omega \delta W_t U_t^{-\frac{1}{2}}) e^{\frac{h_t}{2}} + e^{\frac{h_t}{2}} U_t^{-\frac{1}{2}} \omega (1 - \delta^2)^{\frac{1}{2}} \varepsilon_t, \quad (10a)$$

$$W_t \sim \mathcal{N}_{[0, \infty)}(0, 1), \quad (10b)$$

$$U_t | \mathbf{v} \sim \mathcal{G}\left(\frac{\mathbf{v}}{2}, \frac{\mathbf{v}}{2}\right), \quad (10c)$$

where ε_t and η_t are mutually independent and normally distributed with zero mean and unit variance, $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$, $\mathcal{G}(\cdot, \cdot)$ denotes the Gamma distribution. We set ζ and ω in such a way that $E(y_t | h_t) = 0$ and $V(y_t | h_t) = e^{h_t}$. Model defined by equations (8),(9),(10a)-(10c) will be denoted as SV-ST-VOL. In this setup, equations (8),(9),(10a) and (10c) with $\lambda = 0$ (equivalently $\delta = 0$) and $\forall t = 1, \dots, T$ define the SV-T-VOL model. Finally, equations (8),(9), (10a) and (10b) with $U_t = 1, \forall t = 1, \dots, T$, results the SV-SN-VOL model.

4 Parameter estimation via MCMC

Let $\theta = (\mu, \varphi, \sigma_\eta^2, \nu, \lambda, m_0, m_1)'$ be the full parameter vector of the SV-ST-VOL model, $\mathbf{h}_{1:T} = (h_1, \dots, h_T)'$ be the vector of the log volatilities, $\mathbf{U}_{1:T} = (U_1, \dots, U_T)'$ and $\mathbf{W}_{1:T} = (W_1, \dots, W_T)'$ be the mixing variables, $\mathbf{y}_{1:T} = (y_1, \dots, y_T)'$ and $\mathbf{v}_{1:T} = (v_1, \dots, v_T)'$ be the information available up to time T , while \mathbf{v} is the degrees of freedom parameter vector associated with the mixture distribution and λ the skewness parameter. The Bayesian approach to estimate the parameters in the SV-ST-VOL model uses the data augmentation principle, which considers $\mathbf{h}_{1:T}$, $\mathbf{W}_{1:T}$ and $\mathbf{U}_{1:T}$ as latent variables. The joint posterior density of parameters and latent unobservable variables can be written as

$$\begin{aligned} p(\theta, \mathbf{W}_{1:T}, \mathbf{U}_{1:T}, \mathbf{h}_{1:T} | \mathbf{y}_{1:T}, \mathbf{v}_{1:T}) &\propto p(\mathbf{y}_{1:T} | \theta, \mathbf{W}_{1:T}, \mathbf{U}_{1:T}, \mathbf{h}_{1:T}) p(\mathbf{v}_{1:T} | \theta, \mathbf{h}_{1:T}) \\ &\times p(\mathbf{h}_{1:T} | \theta) p(\mathbf{W}_{1:T}) p(\mathbf{U}_{1:T} | \theta) p(\theta), \end{aligned} \quad (11)$$

where $p(\theta)$ is the prior distribution. Since the posterior density $p(\theta, \mathbf{W}_{1:T}, \mathbf{U}_{1:T}, \mathbf{h}_{1:T} | \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$ does not have closed form, we first sample the parameters θ , followed by the latent variables $\mathbf{W}_{1:T}$, $\mathbf{U}_{1:T}$ and $\mathbf{h}_{1:T}$ using Gibbs sampling. The sampling scheme is described by Algorithm 1. Sampling

the log-volatilities $\mathbf{h}_{1:T}$ in step 5 of Algorithm 1 is the most difficult task due to the nonlinear setup in the observational equation in equations (10a) and (8). In order to avoid the higher correlations due to the Markovian structure of the h_t 's, in the next subsection we develop a multi-move block sampler to sample $\mathbf{h}_{1:T}$ by blocks (Shephard and Pitt 1997; Watanabe and Omori 2004; Abanto-Valle et al. 2010, 2011). Details on the full conditionals of θ and the latent variables $\mathbf{U}_{1:T}$ and $\mathbf{W}_{1:T}$ are given in Appendix.

Algorithm 1

1. Set $i = 0$ and set starting values for the parameters $\theta^{(i)}$ and the latent quantities $\mathbf{W}_{1:T}^{(i)}$, $\mathbf{U}_{1:T}^{(i)}$ and $\mathbf{h}_{1:T}^{(i)}$.
2. Generate $\theta^{(i+1)}$ in turn from its full conditional distribution, given $\mathbf{U}_{1:T}^{(i)}$, $\mathbf{W}_{1:T}^{(i)}$, $\mathbf{h}_{1:T}^{(i)}$, $\mathbf{y}_{1:T}$ and $\mathbf{v}_{1:T}$.
3. Draw $\mathbf{W}_{1:T}^{(i+1)} \sim p(\mathbf{W}_{1:T} | \theta^{(i)}, \mathbf{U}_{1:T}^{(i)}, \mathbf{h}_{1:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$.
4. Draw $\mathbf{U}_{1:T}^{(i+1)} \sim p(\mathbf{U}_{1:T} | \theta^{(i+1)}, \mathbf{W}_{1:T}^{(i+1)}, \mathbf{h}_{1:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$.
5. Generate $\mathbf{h}_{1:T}^{(i+1)}$ by blocks as:
 - i) For $l = 1, \dots, K$, the knot positions are generated as k_l , the floor of $[T \times \{(l + u_l)/(K + 2)\}]$, where the u_l 's are independent realizations of the uniform random variable on the interval $(0,1)$.
 - ii) For $l = 1, \dots, K$, generate $h_{k_{l-1}+1:k_l-1}$ jointly conditional on $\theta^{(i+1)}$, $\mathbf{W}_{k_{l-1}+1:k_l-1}^{(i+1)}$, $\mathbf{U}_{k_{l-1}+1:k_l-1}^{(i+1)}$, $h_{k_{l-1}}^{(i)}$, $h_{k_l}^{(i)}$, $\mathbf{y}_{k_{l-1}:k_l-1}$, $\mathbf{v}_{k_{l-1}:k_l-1}$.
 - iii) For $l = 1, \dots, K$, draw $h_{k_l}^{(i+1)}$ conditional on $\mathbf{y}_{1:T}$, $\theta^{(i)}$, $\mathbf{W}_{k_l}^{(i+1)}$, $U_{k_l}^{(i+1)}$, $h_{k_{l-1}}^{(i+1)}$ and $h_{k_{l+1}}^{(i+1)}$.
6. Set $i = i + 1$ and return to 2 until convergence is achieved.

In the SV-ST-VOL model considered so far, an important modelling assumption is the regularization penalty $p(v)$ on the tail thickness. A default Jeffreys' prior was developed by Fonseca

et al. (2008), with a number of desirable properties particularly when learning a fat-tail from a finite dataset. The default Jeffreys's prior for ν takes the form

$$p(\nu) \propto \left(\frac{\nu}{\nu+3} \right)^{\frac{1}{2}} \left\{ \psi' \left(\frac{\nu}{2} \right) - \psi' \left(\frac{\nu+1}{2} \right) - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}}, \quad (12)$$

where $\psi'(a) = \frac{d\{\psi(a)\}}{da}$ and $\psi(a) = \frac{d\{\log\Gamma(a)\}}{da}$ are the trigamma and digamma functions, respectively. The interesting feature of this prior is its behavior as ν goes to infinity and it has polynomial tails of the form $p(\nu) \propto \nu^{-4}$. In this case, the tail of the prior decays rather fast for large values of ν and assessing the degree of tail thickness can require prohibitively large samples. To the skewness parameter, we assume that $\lambda \sim t_{0.5}(0.0, \frac{\pi^2}{4})$, a Jeffreys' prior suggested by Bayes and Branco (2007), where $t_a(c, d)$ denotes the Student-t distribution with location a , scale b and a degrees-of-freedom.

4.1 Block sampler

In order to simulate $\mathbf{h}_{1:T} = (h_1, \dots, h_T)'$ in the SV-ST-VOL model, we consider a two-step process: first, we simulate h_1 conditional on $\mathbf{h}_{2:T}$, next $\mathbf{h}_{2:T}$ conditional on h_1 . To sample the vector $\mathbf{h}_{2:T}$, we develop a multi-move block algorithm. In our block sampler, we divide it into $K+1$ blocks, $\mathbf{h}_{k_{l-1}+1:k_l-1} = (h_{k_{l-1}+1}, \dots, h_{k_l-1})'$ for $l = 1, \dots, K+1$, with $k_0 = 1$ and $k_{K+1} = T$, where $k_l - 1 - k_{l-1} \geq 2$ is the size of the l -th block. We sample the block of disturbances $\boldsymbol{\eta}_{k_{l-1}:k_l-2} = (\eta_{k_{l-1}}, \dots, \eta_{k_l-2})'$ given the end conditions $h_{k_{l-1}}$ and h_{k_l} instead of $\mathbf{h}_{k_{l-1}+1:k_l-1}$. In order to facilitate the exposition, we omit the dependence on $\boldsymbol{\theta}$, $\mathbf{W}_{t+1:t+k}$, $\mathbf{U}_{t+1:t+k}$, $\mathbf{y}_{t+1:t+k}$ and $\mathbf{v}_{t+1:t+k}$, and suppose that $k_{l-1} = t$ and $k_l = t+k+1$ for the l -th block, such that $t+k < T$. Then $\boldsymbol{\eta}_{t:t+k-1} = (\eta_t, \dots, \eta_{t+k-1})'$ are sampled at once from their full conditional distribution $f(\boldsymbol{\eta}_{t:t+k-1} | h_t, h_{t+k+1}, \mathbf{y}_{t:t+k})$, which without the constant terms is expressed in log scale as

$$\begin{aligned} \log f(\boldsymbol{\eta}_{t:t+k-1} | h_t, h_{t+k+1}) &= \text{const} - \frac{1}{2} \sum_{r=t}^{t+k-1} \eta_r^2 + \sum_{r=t+1}^{t+k} l(h_r) \\ &\quad - \frac{1}{2\sigma_\eta^2} [h_{t+k+1} - \mu - \boldsymbol{\varphi}(h_{t+k} - \mu)]^2 \mathbb{I}(t+k < T), \end{aligned} \quad (13)$$

where $\mathbb{I}(\cdot)$ is an indicator function. We denote the first and second derivatives of $l(h_r)$ with respect to h_r by l' and l'' , where $l(h_r) = \log p(v_r | m_0, m_1, h_r) + \log p(y_r | \nu, \lambda, W_r, U_r, h_r)$ is obtained

from equation (8) and (10a). As (13) does not have closed form, we use the Metropolis-Hastings acceptance-rejection algorithm (Tierney, 1994; Chib and Greenberg, 1995) to sample from. We propose to use the following artificial Gaussian state space model as a proposed density to simulate the block $\eta_{t+1:t+k}$

$$\hat{y}_r = h_r + \xi_r, \quad \xi_r \sim \mathcal{N}(0, d_r), \quad r = t+1, \dots, t+k, \quad (14)$$

$$h_{r+1} = \mu + \varphi(h_r - \mu) + \sigma_\eta \eta_r, \quad \eta_r \sim \mathcal{N}(0, 1), \quad r = t, t+1, \dots, t+k-1, \quad (15)$$

where the auxiliary variables d_r and \hat{y}_r for $r = t+1, \dots, t+k-1$ and $t+k = T$ are defined as follows:

$$\begin{aligned} d_r &= -\frac{1}{l_F''(\hat{h}_r)}, \\ \hat{y}_r &= \hat{h}_r + d_r l'(\hat{h}_r). \end{aligned} \quad (16)$$

For $r = t+k < T$, it follows that

$$\begin{aligned} d_r &= \frac{\sigma_\eta^2}{\varphi^2 - \sigma_\eta^2 l_F''(\hat{h}_{t+k})}, \\ \hat{y}_r &= d_r \left[l'(\hat{h}_r) - l_F''(\hat{h}_r) \hat{h}_r + \frac{\varphi}{\sigma_\eta^2} [h_{r+1} - \mu(1 - \varphi)] \right]. \end{aligned} \quad (17)$$

We obtain the measurement equation (14) by a second-order expansion of l_r around some preliminary estimate of η_r , denoted by $\hat{\eta}_r$, where \hat{h}_r is the estimate of h_r equivalent to $\hat{\eta}_r$, and

$$l_F''(h_r) = E[l''(h_r)] = -\frac{1}{2} - \frac{(\zeta + \omega \delta W_t U_t^{-\frac{1}{2}})^2}{4\omega^2(1 - \delta^2)} U_r - \frac{m_1^2 e^{2h_r}}{m_0 + m_1 e^{h_r}}, \quad (18)$$

which is everywhere strictly negative. The expectation in (18) is taken with respect to y_r and v_r conditional on h_r, W_r, U_r, θ . Since (14)-(15) define a Gaussian state space model, we can apply de Jong and Shephard's simulation smoother (de Jong and Shephard, 1995) to perform the sampling. We denote this density by g . Since f is not bounded by g , we use the Metropolis-Hastings acceptance-rejection algorithm to sample from f , as recommended by Chib and Greenberg (1995). In the SV-SN-VOL case, we use the same procedure with $U_t = 1$ for $t = 1, \dots, T$.

The procedure to select the expansion block $\hat{\mathbf{h}}_{t+1:t+k}$ is described in the Algorithm 2.

Algorithm 2

1. Initialize $\hat{\mathbf{h}}_{t+1:t+k}$.
2. Evaluate recursively $l'(\hat{h}_r)$ and $l''_F(\hat{h}_r)$ for $r = t + 1, \dots, t + k$.
3. Conditional on the current values of the vector of parameters θ , $\mathbf{U}_{t+1:t+k}$, $\mathbf{W}_{t+1:t+k}$, h_t and h_{t+k+1} , define the auxiliary variables \hat{y}_r and d_r using equations (16) or (17) for $r = t + 1, \dots, t + k$.
4. Consider the linear Gaussian state-space model in (14) and (15). Apply the Kalman filter and a disturbance smoother (Koopman, 1993) and obtain the posterior mean of $\eta_{t:t+k}(\mathbf{h}_{t:t+k})$ and set $\hat{\eta}_{t:t+k}(\hat{\mathbf{h}}_{t:t+k})$ to this value.
5. Return to step 2 and repeat the procedure until achieving convergence.

Finally, we describe the updating procedure for h_1 and the knot conditions h_{k_l} , for $l = 1, \dots, K$. First, we simulate h_1 from $p(h_1 | h_2, \theta, \mathbf{y}_{1:T})$ by using the Metropolis-Hasting (MH) algorithm with the normal density, $\mathcal{N}(\mu + \varphi[h_2 - \mu], \sigma_\eta^2)$, as a proposal. Let h_1^p and $h_1^{(i-1)}$ denote the proposal and the previous iteration values, the acceptance probability is given by $\alpha_{MH} = \min\{1, \frac{Q(h_1^p)}{Q(h_1^{(i-1)})}\}$, where $Q(h_1)$ is the conditional density of $y_1, v_1 | \theta, W_1, U_1, h_1$. As the density $p(h_{k_l} | h_{k_l-1}, h_{k_l+1})$ does not have a closed form, we use the MH algorithm with $\mathcal{N}(\frac{\mu(1-\varphi)^2 + \varphi(h_{k_l-1} + h_{k_l+1})}{1+\varphi^2}, \frac{\sigma_\eta^2}{1+\varphi^2})$, as the proposal distribution. Let $h_{k_l}^p$ and $h_{k_l}^{(i-1)}$ denote the proposal value and the previous iteration value. Thus, the acceptance probability is given by $\alpha_{MH} = \min\{1, \frac{Q(h_{k_l}^p)}{Q(h_{k_l}^{(i-1)})}\}$, where $Q(h_{k_l})$ is the conditional density of $y_{k_l}, v_{k_l} | \theta, W_{k_l}, U_{k_l}, h_{k_l}$.

5 Empirical Application

This section analyzes the daily closing prices and daily number of traded shares corrected by dividends and stock splits for the common stocks of the Chevron-Texaco (CVX), International Business Machine (IBM), The Coca-Cola Company (KO) and Petroleo Brasileiro (PBR) listed on the New York stock exchange (NYSE). The data set was obtained from the Yahoo finance

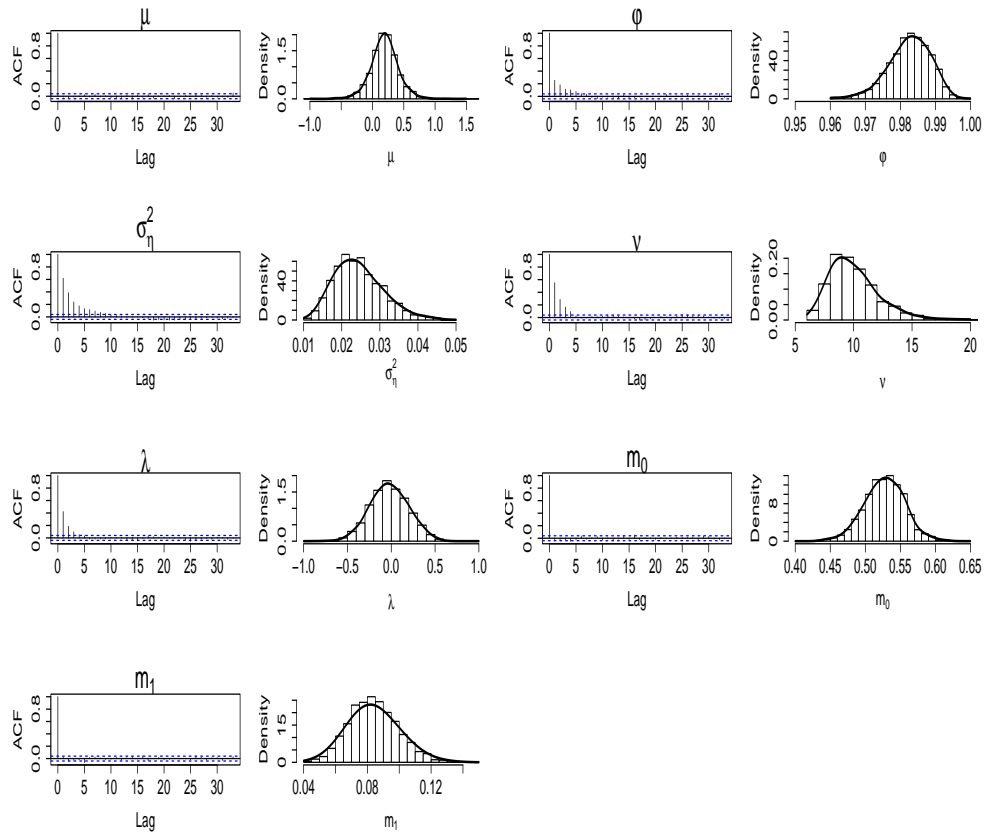


Figure 2: Estimation results for IBM data set. SV-ST-VOL model. We plot the autocorrelation function (ACF) and the histograms of the parameters μ , ϕ , σ_{η}^2 , v , λ , m_0 and m_1 .

web site at <http://finance.yahoo.com>. The analyzed period starts January 2, 2004 and ends April 05, 2012, yielding 2081 observations. Throughout we work with compounded returns: $y_t = 100 \times \{\log(P_t) - \log(P_{t-1})\}$ where P_t is the closing price on day t . To make the volume series stationary, the volume data are adjusted by regressing the log of the trading volume on a constant and on time $t = 1, 2, \dots, T$. The exponential function of the residuals of this regression is then linearly transformed so that the raw data and the detrended data have the same mean and variance. For the following results, the detrended series is multiplied by 10^{-7} .

Summary statistics of the return and detrended volume series are given in Table 1. As one can see, the returns and trading volume are clearly not normally distributed. The kurtosis of each series exceeds the value of three which would be expected for a normally distributed variable and the Jarque-Bera statistics (JB) overwhelmingly reject normality. Finally, there is significant positive contemporaneous correlation coefficient between return and volume, as confirmed by the correlation coefficient between volume and absolute returns $\rho_{|y|,v}$. All these findings are consistent with the MMM.

Now, we fitted the SV, SV-T, SV-SN, SV-ST, SV-N-VOL, SV-T-VOL, SV-SN-VOL and SV-

Table 1: Summary statistics of the returns and the detrended volumes series

	Returns							
	Mean	St. Dev.	Min	Max	Skewness	Kurtosis	JB	
CVX	0.06	1.82	-13.34	18.94	0.09	16.18	15065.6	
IBM	0.04	1.40	-8.67	10.90	0.03	8.67	2786.8	
KO	0.03	1.21	-9.06	12.99	0.42	16.63	16179.9	
PBR	0.07	3.12	-26.24	26.45	-0.16	13.09	8837.2	
	Trading Volume							
	Mean	St. Dev.	Min	Max	Skewness	Kurtosis	JB	$\rho_{ y ,v}$
CVX	0.99	0.44	0.20	4.87	1.73	9.07	4240.9	0.56
IBM	0.68	0.32	0.01	3.11	2.19	10.88	7055.1	0.61
KO	0.89	0.43	0.89	5.92	2.69	18.94	24574.3	0.55
PBR	1.36	0.83	0.03	8.87	2.21	12.32	9238.0	0.54

ST-VOL models. In all cases, we simulated the h_t 's in a multi-move fashion with stochastic knots

based on the method described in Section 4. The number of blocks, K is set at 60, in a such way that each block contained 35 h'_i s on an average. We set the prior distribution of the parameters as: $\varphi \sim \mathcal{N}_{(-1,1)}(0.95, 100)$, $\sigma^2 \sim \mathcal{IG}(2.5, 0.025)$, $\mu \sim \mathcal{N}(0, 100)$, $m_0 \sim \mathcal{G}(0.08, 0.01)$ and $m_1 \sim \mathcal{N}_{(0,\infty)}(0.10, 100)$. We assume that $\lambda \sim t_{0.5}(0.0, \frac{\pi^2}{4})$, a Jeffreys' prior suggested by Bayes and Branco (2007). Finally, for v , we assume that the prior given by equation (12). For the parameter φ the priors' mean and variance are 0.0032 and 0.3328, respectively. This prior setup is equivalent to the uniform distribution on interval $(-1, 1)$, which gives zero mean and variance of 0.3333. All the calculations were performed running stand-alone code developed by us using an open source C++ library for statistical computation, the Scythe statistical library (Pemstein et al., 2007), which is available for free download at <http://scythe.wustl.edu>.

For all models, we conducted the MCMC simulation for 130000 iterations. In all cases, the first 30000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, only every 40th values of the chain were stored. With the resulting 2500 values, we calculated the posterior means and the 95% credibility intervals. All the sequences passed the convergence diagnostic (CD) statistics (Geweke, 1992) but the results are not reported. Table 2 summarizes the results for the respective SV models for each one of the analyzed stocks and Tables 3 and 4 summarizes the results for the SV-VOL class. Figure 2 shows a rapid decay of autocorrelations for all the parameters in the SV-ST-VOL model for the IBM stock. It is an indicator that our MCMC algorithm mixed well.

From Table 2, consistent with the existing evidence of great persistence in the log-volatility process, we found for all the stocks, that the posterior means of φ and the 95% posterior credibility intervals very close to unity. The values of posterior mean for the SV-T (SV-ST) being greater than that of the SV-N (SV-ST), respectively. As expected, we found the persistence parameter more persistent for stocks with higher kurtosis in the returns. The posterior mean of σ_η^2 is smaller in the SV-T (SV-ST) than that of the SV-N (SV-ST), indicating that the log-volatility process of the heavy tailed version of the SV models are less variable than those of the SV-N and SV-SN. From Tables 3 and 4, we found that the values of the posterior means of the persistence parameter φ in the SV-VOL models are slightly lower than the equivalent SV models, but very close to the

unity, indicating a high persistence. This is in sharp contrast with the results in [Andersen \(1996\)](#) and [Liesenfeld \(1998\)](#), who found substantially lower persistence in volatility in the case of the modified mixture model. The opposite occurs with σ_{η}^2 , so the posterior means are slightly greater than the respective SV model.

We now turn to discuss our results for ν and λ the tail-fatness and skewness parameters respectively. Our estimates on the SV and the SV-VOL models are consistently with the well established stylized fact of presence of heavy tails. For example in the SV-T and SV-ST of the KO stock, the posterior mean and credibility intervals of ν are 8.0742 (5.9779,11.1720) and 8.3837 (6.1100,12.0790) respectively. The results for this stock considering the SV-T-VOL and SV-ST-VOL models are 8.6196 (6.2750,12.1830) and 9.1227 (6.4790,13.4790), respectively. In both cases, our results indicate a strong evidence of heavy-tails. Similar results are obtained for the IBM stock. In the PBR stock we have the posterior mean and credibility intervals for ν in the SV-T and SV-ST-VOL are 22.7346 (11.7300,38.3400) and 22.4125 (10.9900,38.0500) respectively. We found that the skewness parameter λ in the SV-ST model is significant for the CVX and PTB stocks (the posterior credibility interval does not contain zero). In fact, we have the posterior mean and credibility interval are -0.7642 (-1.2200,-0.0521) and -0.7045 (-1.1207,-0.09421), respectively. The same occurs for the SV-ST-VOL, where the posterior mean and credibility intervals are -0.5534 (-0.7601,-0.0254) and -0.7434 (-1.1688,-0.1395), respectively. For both stocks, the posterior means of the ν parameter are greater than the IBM and KO stocks. On the other hand for the IBM and KO stocks we found that the posterior mean is different from zero, but the posterior credibility interval of λ parameter contains zero and there is a strong evidence of heavy-tails, as mentioned earlier. The results are consistent with the SV model equivalent for each one of the stocks.

As mentioned earlier, m_0 reflects the noisy component of trading volume generated by liquidity traders. The remaining part of trading volume that is induced by new information is represented by $m_1 e^{h_t}$. We find that the distributions of m_0 and m_1 have posterior means close to 0.75, 0.53, 0.71 and 0.87 (m_0) and 0.08, 0.08, 0.07 and 0.06 (m_1) for the CVX, IBM, KO and PBR stocks for all the SV-VOL models considered here. It is important to note that the posterior mean of the

skewness parameter is modified with the inclusion of the trading volume. For example, for the PBR stock. In the SV-ST model, the posterior mean and 95% credibility interval for λ are -0.7045 (-1.1207,-0.09421) and considering the SV-ST-VOL model, -0.7434 (-1.1688,-0.1395). Similar findings are found for the other stocks. We conclude that considering the trading volume can also explain asymmetric volatility.

To assess the goodness of the estimated models, we calculate the Bayesian predictive information criteria, BPIC (Ando, 2006, 2007). The BPIC criterion is defined as

$$BPIC = -2E_{\theta|\mathbf{y}_{1:T}}[\log\{p(\mathbf{y}_{1:T} | \theta)\}] + 2T\hat{b}, \quad (19)$$

where \hat{b} is given by

$$\hat{b} \approx \frac{1}{T} \left\{ E_{\theta|\mathbf{y}_{1:T}}[\log\{p(\mathbf{y}_{1:T} | \theta)p(\theta)\}] - \log[p(\mathbf{y}_{1:T} | \hat{\theta})p(\hat{\theta})] + \text{tr}\{J_T^{-1}(\hat{\theta})I_T(\hat{\theta})\} + 0.5q \right\}. \quad (20)$$

Here q is the dimension of θ , $E_{\theta|\mathbf{y}_{1:T}}[\cdot]$ denotes the expectation with respect to the posterior distribution, $\hat{\theta}$ is the posterior mode, and

$$I_T(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^T \left(\frac{\partial \eta_T(y_t, \theta)}{\partial \theta} \frac{\partial \eta_T(y_t, \theta)}{\partial \theta'} \right) \Big|_{\theta=\hat{\theta}},$$

$$J_T(\hat{\theta}) = \frac{1}{T} \sum_{t=1}^T \left(\frac{\partial^2 \eta_T(y_t, \theta)}{\partial \theta \partial \theta'} \right) \Big|_{\theta=\hat{\theta}},$$

with $\eta_T(y_t, \theta) = \log p(y_t | \mathbf{y}_{1:t-1}, \theta) + \log p(\theta)/T$. In all the applications here, the log-likelihood function, $\log p(\mathbf{y}_{1:T} | \theta)$, is estimated using the auxiliary particle filter (see, e.g., Pitt and Shephard, 1999) with 10000 particles. The best model has the smallest BPIC.

As the quantity of information available is different between models with and without trading volume, we compare first the SV-N, SV-T, SV-SN and SV-ST models. From Table 2, according to the BPIC, the SV-SN and SV-ST give the worst and the best fit for the returns data set for CVX, IBM, KO and PB stocks, respectively. So, it gives evidence that for all the stocks it is important to take account simultaneously flat-tailness as well as skewness, not skewness only. Now,

we compare the SV-N-VOL, SV-T-VOL, SV-SN-VOL and SV-ST-VOL. As in the corresponding SV model, the best fit is given by the SV-ST-VOL model and the worst for the SV-SN-VOL. It confirms our previous comment, about the importance to model jointly skewness and heavy-tails.

In Figure 3 , we plot the smoothed mean of $e^{\frac{h_t}{2}}$ obtained from the MCMC output for the SV-N-VOL (solid line) and the SV-ST-VOL (dotted line) for all the series of returns, that is the CVX, IBM, KO and PTB. From a practical point of view, we are mainly interested in whether we find a significant difference between the two series. Therefore, in the right of Figure 3, we plot the smoothed mean and 95% credibility intervals of the differences of $e^{\frac{h_t}{2}}$ from the SV-N-VOL and SV-ST-VOL models. This graph shows us that these series do not differ very much in most periods, but in some periods of high volatility, we observe differences in percentages of more than 2%. This can have a substantial impact, for instance, in the valuation of derivative instruments and several strategic or tactical asset allocation topics.

Table 3: Estimation results for the SV-N-VOL and SV-T-VOL models. We report the posterior mean and the 95% posterior credibility interval between parenthesis. The Bayesian predictive information criteria (BPIC) is reported.

	CVX	IBM	SV-N-VOL	KO	PBR
μ	0.6299 (0.1014,1.1280)	0.1370 (-0.1970,0.4821)	-0.2341 (-0.5948,0.1179)	1.6762 (1.2360,2.0910)	
φ	0.9873 (0.9778,0.9956)	0.9688 (0.9513,0.9834)	0.9716 (0.9543,0.9852)	0.9812 (0.9695,0.9917)	
σ_η^2	0.0153 (0.0100,0.0226)	0.0477 (0.0309,0.0698)	0.0453 (0.0279,0.0695)	0.0271 (0.0182,0.0380)	
m_0	0.7562 (0.6847,0.8265)	0.5257 (0.4714,0.5789)	0.7119 (0.6494,0.7728)	0.8751 (0.7806,0.9666)	
m_1	0.0819 (0.0585,0.1092)	0.0864 (0.0600,0.1170)	0.1365 (0.0927,0.1865)	0.0571 (0.0451,0.0708)	
BPIC	14019.7	10297.6	10260.8	24567.1	
			SV-T-VOL		
μ	0.6385 (0.1312,1.1981)	0.2015 (-0.2361,0.6434)	-0.2005 (-0.9458,0.4381)	1.7768 (1.2780,2.2680)	
φ	0.9878 (0.9788,0.9961)	0.9826 (0.9694,0.9934)	0.9865 (0.9756,0.9951)	0.9861 (0.9765,0.9948)	
σ_η^2	0.0144 (0.0094,0.0215)	0.0243 (0.0135,0.0396)	0.0198 (0.0114,0.0321)	0.0176 (0.0112,0.0263)	
v	31.6031 (19.2800,39.6300)	9.6450 (6.7210,14.8970)	8.6196 (6.2750,12.1830)	22.7346 (11.7300,38.3400)	
m_0	0.7470 (0.6713,0.8178)	0.5267 (0.4703,0.5813)	0.7089 (0.6466,0.7729)	0.8330 (0.7238,0.9370)	
m_1	0.0850 (0.0589,0.1144)	0.0840 (0.0571,0.1174)	0.1362 (0.0899,0.1909)	0.0571 (0.0451,0.0708)	
BPIC	11572.5	9351.4	9601.9	14777.4	

Table 4: Estimation results for the SV-SN-VOL and SV-ST-VOL models. We report the posterior mean and the 95% posterior credibility interval between parenthesis. The Bayesian predictive information criteria (BPIC) is reported.

	CVX	IBM	SV-SN-VOL	KO	PBR
μ	0.6331 (0.0807,1.1221)	0.1321 (-0.2003,0.4508)	-0.2407 (-0.6086,0.1264)	1.6775 (1.2180,2.1110)	
φ	0.9876 (0.9785,0.9958)	0.9691 (0.9518,0.98307)	0.9738 (0.9584,0.9873)	0.9813 (0.9691,0.9917)	
σ_{η}^2	0.0147 (0.0095,0.0218)	0.0471 (0.0299,0.0688)	0.0417 (0.0260,0.0629)	0.0269 (0.0181,0.0375)	
λ	-0.7524 (-1.2974,0.3265)	0.2718 (-0.7086,1.0340)	0.7032 (-0.3659,1.1950)	-0.3663 (-1.0940,0.6186)	
m_0	0.7554 (0.6818,0.8289)	0.5244 (0.4698,0.5776)	0.7087 (0.6468,0.7690)	0.8716 (0.7749,0.9625)	
m_1	0.0821 (0.0566,0.1111)	0.0872 (0.0602,0.1171)	0.1402 (0.0972,0.1910)	0.0576 (0.0451,0.0718)	
BPIC	14169.4	10517.3	10470.3	24306.7	
			SV-ST-VOL		
μ	0.6630 (0.1073,1.1557)	0.1949 (-0.2454,0.6341)	-0.2091 (-0.7409,0.3409)	1.7796 (1.2540,2.2680)	
φ	0.9880 (0.9789,0.9964)	0.9826 (0.9693,0.9927)	0.9862 (0.9749,0.9955)	0.9867 (0.9765,0.9952)	
σ_{η}^2	0.0140 (0.0090,0.0210)	0.0244 (0.0143,0.0390)	0.0204 (0.0121,0.0330)	0.0169 (0.0110,0.0253)	
ν	31.3496 (19.0100,39.6500)	10.0585 (6.8560,15.4120)	9.1227 (6.4790,13.4790)	22.4125 (10.9900,38.0500)	
λ	-0.5534 (-0.7601,-0.0254)	-0.0206 (-0.4416,0.4065)	0.1654 (-0.2379,0.5782)	-0.7434 (-1.1688,-0.1395)	
m_0	0.7463 (0.6697,0.8199)	0.5272 (0.4696,0.5840)	0.7080 (0.6435,0.7704)	0.8339 (0.7215,0.9414)	
m_1	0.0852 (0.0598,0.1148)	0.0840 (0.0562,0.1165)	0.1386 (0.0943,0.1926)	0.0622 (0.0476,0.0780)	
BPIC	11317.2	9339.6	9598.9	14490.3	

6 Conclusions

This article studies the joint distribution of daily returns and trading volume based on the modified mixture model. We extend this specification to take account for skewness and heavier-tailed for the returns using the skew-Student-t distribution. We have constructed an algorithm based on Markov Chain Monte Carlo (MCMC) simulation methods to estimate all the parameters and latent quantities in the model using the Bayesian approach. As a by product of the MCMC algorithm, we were able to produce an estimate of the latent information process which can be used in financial modeling. The SV-ST-VOL model is supported by the data set of the common stocks of the CVX, IBM, KO and PBR.

This article makes certain contributions, but several extensions are still possible. First, we specify the log of volatility as a simple AR(1) process, but more elaborate models, such as long memory models, may be required to specify volatility. Second, we can include a dynamic pattern in the parameters m_0 and m_1 , considering them as time varying parameters, and finally we can extend the model to include many assets.

Acknowledgements

The research of Carlos A. Abanto-Valle was supported by CNPq. V. H. Lachos acknowledges financial support from FAPESP and CNPq.

Appendix A: The full conditionals

In this appendix, we describe the full conditional distributions for the parameters and the mixing latent variables $\mathbf{U}_{1:T}$ and $\mathbf{W}_{1:T}$ for the SV-ST model.

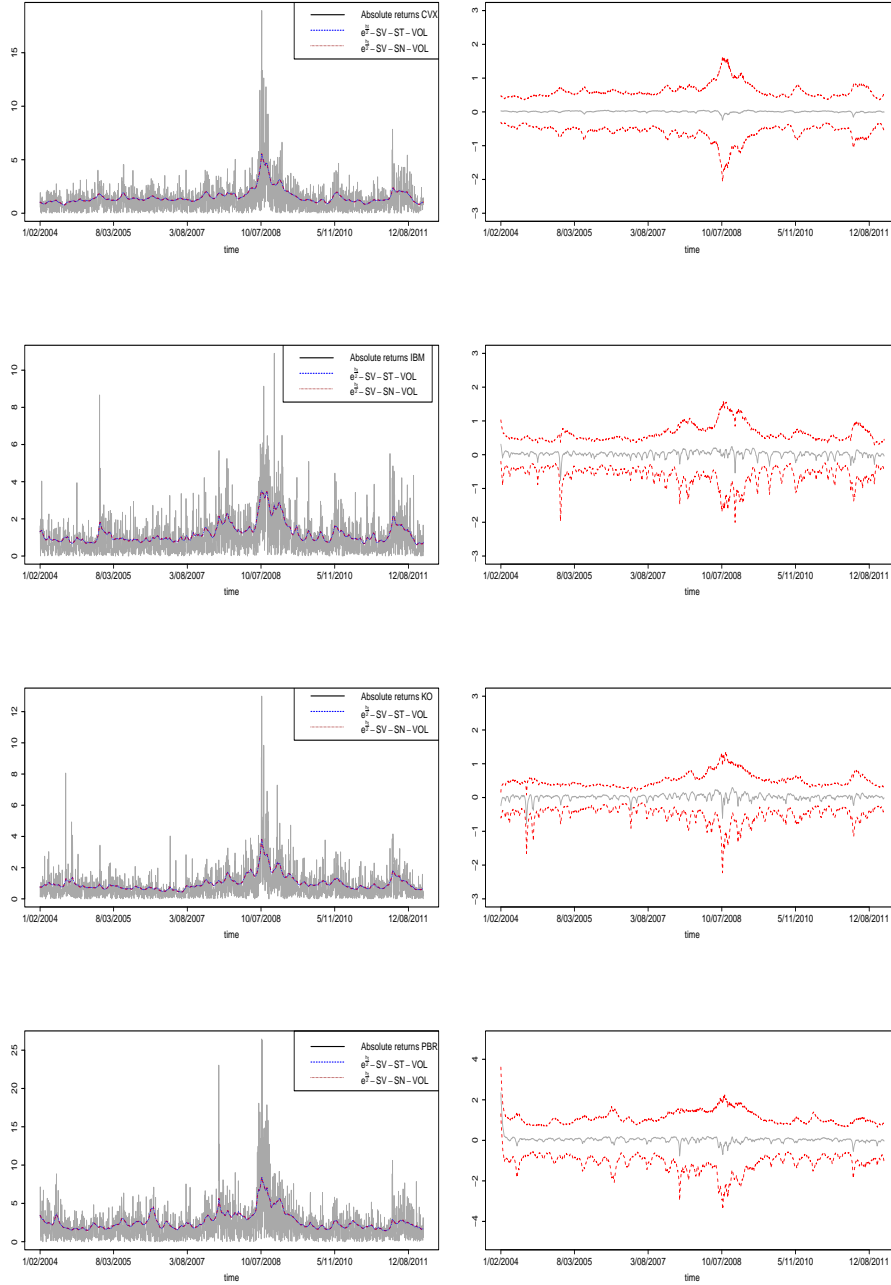


Figure 3: Left: Posterior smoothed mean of $e^{h_t/2}$ and the absolute returns. Right: Posterior smoothed mean of the difference of $e^{h_t/2}$ in the SV-ST-VOL and SV-N-VOL.

Full conditional distribution of μ , φ and σ_η^2

The prior distributions of the common parameters are set as: $\mu \sim N(\bar{\mu}, \sigma_\mu^2)$, $\varphi \sim \mathcal{N}_{(-1,1)}(\bar{\varphi}, \sigma_\varphi^2)$, $\sigma_\eta^2 \sim \mathcal{IG}(\frac{T_0}{2}, \frac{M_0}{2})$. We have the following full conditional for μ :

$$\mu \mid \mathbf{h}_{1:T}, \varphi, \sigma_\eta^2 \sim \mathcal{N}\left(\frac{b_\mu}{a_\mu}, \frac{1}{a_\mu}\right), \quad (\text{A.1})$$

where $a_\mu = \frac{1}{\sigma_\alpha^2} + \frac{(T-1)(1-\varphi)^2}{\sigma_\eta^2} + \frac{(1-\varphi)^2}{\sigma_\eta^2}$ and $b_\mu = \frac{\bar{\mu}}{\sigma_\mu^2} + \frac{(1-\varphi^2)}{\sigma_\eta^2} h_1 + \frac{\sum_{t=1}^{T-1} (h_{t+1} - \varphi h_t)(1-\varphi)}{\sigma_\eta^2}$. In a similar way, the conditional posterior distribution of φ is given by

$$p(\varphi \mid \mathbf{h}_{1:T}, \mu, \sigma_\eta^2) \propto Q(\varphi) \exp\left\{-\frac{a_\varphi}{2} (\psi - \frac{b_\varphi}{a_\varphi})^2\right\} \mathbb{I}_{|\varphi| < 1}, \quad (\text{A.2})$$

where $Q_\varphi = \sqrt{1-\varphi^2} \exp\left\{-\frac{1}{2\sigma_\eta^2} [(1-\varphi^2)(h_1 - \mu)^2]\right\}$, $a_\varphi = \frac{\sum_{t=1}^{T-1} (h_t - \mu)^2}{\sigma_\eta^2} + \frac{1}{\sigma_\psi^2}$, $b_\varphi = \frac{\sum_{t=1}^{T-1} (h_t - \mu)(h_{t+1} - \mu)}{\sigma_\eta^2} + \frac{\bar{\varphi}}{\sigma_\varphi^2}$ and $\mathbb{I}_{|\varphi| < 1}$ is an indicator variable. As $p(\varphi \mid \mathbf{h}_{0:T}, \alpha, \sigma_\eta^2)$ in (A.2) does not have closed form, we sample from it by using the Metropolis-Hastings algorithm with truncated $\mathcal{N}_{(-1,1)}(\frac{b_\varphi}{a_\varphi}, \frac{1}{a_\varphi})$ as the proposal density.

Finally, the full conditional of σ_η^2 is $\mathcal{IG}(\frac{T_1}{2}, \frac{M_1}{2})$, where $T_1 = T_0 + T$ and $M_1 = M_0 + [(1 - \psi^2)(h_1 - \mu)^2] + \sum_{t=1}^{T-1} [h_{t+1} - \mu - \psi(h_t - \mu)]^2$.

Full conditional of ν , λ , U_t and W_t

We, set ζ and ω in such a way that $E(y_t \mid h_t) = 0$ and $V(y_t \mid h_t) = e^{h_t}$. So, we have $\zeta = -\sqrt{\frac{2}{\pi}} k_1 \delta \omega$ and $\omega^2 = \left[k_2 - \frac{2}{\pi} k_1^2 \delta^2\right]^{-1}$, where $k_1 = \sqrt{\frac{\nu}{2} \frac{\Gamma(\frac{\nu-1}{2})}{\Gamma(\frac{\nu}{2})}}$, $k_2 = \frac{\nu}{\nu-2}$ and $\delta = \frac{\lambda}{\sqrt{1+\lambda^2}}$. Then the full conditionals of ν and λ are as follows:

$$\begin{aligned} p(\nu \mid \cdot) &\propto \left(\frac{\nu}{\nu+3}\right)^{\frac{1}{2}} \left\{ \psi' \left(\frac{\nu}{2}\right) - \psi' \left(\frac{\nu+1}{2}\right) - \frac{2(\nu+3)}{\nu(\nu+1)^2} \right\}^{\frac{1}{2}} \\ &\times \left(\frac{1}{\omega}\right)^T e^{-\frac{1}{2\omega^2(1-\delta^2)} \sum_{t=1}^T U_t e^{-h_t} (y_t - \zeta - \omega \delta W_t U_t^{-\frac{1}{2}} e^{\frac{h_t}{2}})^2}, \end{aligned} \quad (\text{A.3})$$

$$p(\lambda \mid \cdot) \propto \left(1 + \frac{2\lambda}{\pi^2}\right)^{-\frac{3}{4}} \left(\frac{1}{1-\delta^2}\right)^{\frac{T}{2}} e^{-\frac{1}{2\omega^2(1-\delta^2)} \sum_{t=1}^T U_t e^{-h_t} (y_t - \zeta - \omega \delta W_t U_t^{-\frac{1}{2}} e^{\frac{h_t}{2}})^2}. \quad (\text{A.4})$$

Since the above full conditional distributions are not in any known closed form, we must simulate ν and λ using the Metropolis-Hastings algorithm. The proposal density used are $\mathcal{N}_{(\nu>2)}(\mu_\nu, \tau_\nu^2)$

and $\mathcal{N}(\mu_\lambda, \tau_\lambda^2)$, with $\mu_\nu = x - \frac{q'(x)}{q''(x)}$ and $\tau_\nu^2 = \max\{0.001, (-q''(x))^{-1}\}$ for $\nu = \nu$ or λ , where x is the value of the previous iteration, $q(\cdot)$ is the logarithm of the conditional posterior density, and $q'(\cdot)$ and $q''(\cdot)$ are the first and second derivatives respectively.

As $U_t \sim \mathcal{G}(\frac{\nu}{2}, \frac{\nu}{2})$, the conditional posterior of U_t is given by

$$p(U_t | h_t, W_t, \nu, \lambda) \propto Q(U_t) U_t^{\frac{\nu+1}{2}-1} e^{-\frac{U_t}{2} [\nu + \frac{e^{-h_t} (y_t - \zeta e^{\frac{h_t}{2}})^2}{\omega^2 (1-\delta^2)}]}, \quad (\text{A.5})$$

where $Q(U_t) = e^{-\frac{U_t^{\frac{1}{2}} \delta W_t e^{-\frac{h_t}{2} (y_t - \zeta e^{\frac{h_t}{2}})}}{\omega (1-\delta^2)}}$. As $p(U_t | h_t, W_t, \nu, \lambda)$ in (A.5) does not have closed form, we sample it by using the Metropolis-Hastings algorithm with $\mathcal{G}(\frac{\nu+1}{2}, \frac{1}{2} [\nu + \frac{e^{-h_t} (y_t - \zeta e^{\frac{h_t}{2}})^2}{\omega^2 (1-\delta^2)}])$ as the proposal density. Finally, from equations (10a) and (10b), we have the full conditional of W_t is the $\mathcal{N}_{(0,\infty)}(\frac{\delta U_t^{-\frac{1}{2}} e^{-\frac{h_t}{2} [y_t - \zeta e^{\frac{h_t}{2}}]}}{\omega}, \frac{1}{1-\delta^2})$.

Full conditionals of m_0 and m_1

We assume that prior distributions are, respectively, $m_0 \sim \mathcal{G}(a_0, b_0)$ and $m_1 \sim \mathcal{N}(a_1, b_1)$. Then the full the conditionals of m_0 and m_1 are as follows:

$$p(m_0 | \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, m_1) \propto e^{-(b_0+T)m_0} e^{\sum_{t=1}^T \log[m_0^{\frac{a_0-1}{T}} (m_0+m_1 e^{h_t})^{\nu_t}]}, \quad (\text{A.6})$$

$$p(m_1 | \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, m_0) \propto e^{[-\frac{1}{2b_1} (m_1 - a_1)^2 + \sum_{t=1}^T e^{h_t}]} e^{\sum_{t=1}^T \log[(m_0+m_1 e^{h_t})^{\nu_t}]}. \quad (\text{A.7})$$

Since the above full conditional distributions are not in any known closed form, we must simulate m_0 and m_1 using the Metropolis-Hastings algorithm. The proposal density used is $\mathcal{N}_{(m_i>0)}(\mu_{m_i}, \tau_{m_i}^2)$, with $\mu_{m_i} = x - \frac{q'(x)}{q''(x)}$ and $\tau_{m_i}^2 = (-q''(x))^{-1}$ for $i = 0, 1$, where x is the value of the previous iteration, $q(\cdot)$ is the logarithm of the conditional posterior density, and $q'(\cdot)$ and $q''(\cdot)$ are the first and second derivatives respectively.

References

Abanto-Valle CA, Bandyopadhyay D, Lachos VH, Enriquez I. 2010. Robust Bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions. *Computational Statistics & Data Analysis* **54**: 2883–2898.

- Abanto-Valle CA, Migon HS, Lachos VH. 2011. Stochastic volatility in mean with heavy-tailed distributions. *Brazilian Journal of Probability and Statistics* (**forthcoming**).
- Andersen T. 1996. Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance* **51**: 169–204.
- Ando T. 2006. Bayesian inference for nonlinear and non-Gaussian stochastic volatility model with leverage effect. *Journal of Japan Statistical Society* **36**: 173–197.
- Ando T. 2007. Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models. *Biometrika* **94**: 443–458.
- Azzalini A. 1986. Further results on a class of distributions which include the normal ones. *Statistica* **46**: 199–208.
- Azzalini A. 2005. The skew-normal distribution and related multivariate families. *Scandinavian Journal of Statistics* **32**: 159–188.
- Azzalini A, Capitanio A. 2003. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution. *Journal of Royal Statistical Society: Series B* **65**: 367–389.
- Bayes CL, Branco MD. 2007. Bayesian inference for the skewness parameter of the scalar skew-normal distribution. *Brazilian Journal of Probability and Statistics* **21**: 141–163.
- Bollerslev T, Jubinski D. 1999. Equity trading volume and volatility: Latent information arrivals and common long-run dependencies. *Journal of Business & Economic Statistics* **17**: 9–21.
- Branco MD, Dey DK. 2001. A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis* **79**: 99–113.
- Cappuccio N, Lubian D, Raggi D. 2006. Investigating asymmetry in U.S. stock market indexes: evidence from a stochastic volatility model. *Applied Financial Economics* **16**: 479–490.

- Chib S, Greenberg E. 1995. Understanding the Metropolis-Hastings algorithm. *The American Statistician* **49**: 327–335.
- Chib S, Nardari F, Shepard N. 2002. Markov Chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics* **108**: 281–316.
- Clark PK. 1973. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica* **41**: 135–155.
- Corrado CJ, Su T. 1997. Implied volatility skews and stock index skewness and kurtosis implied by S&P500 index option prices. Working paper, University of Missouri - Columbia.
- de Jong P, Shephard N. 1995. The simulation smoother for time series models. *Biometrika* **82**: 339–350.
- Epps T, Epps M. 1976. The stochastic dependence of stochastic price changes and transaction volume: Implications for the mixture of distribution hypothesis. *Econometrica* **44**: 305–321.
- Fonseca TCO, Ferreira MAR, Migon HS. 2008. Objective Bayesian analysis for the student-t regression model. *Biometrika* **95**: 325–333.
- Foster FD, Viswanathan S. 1995. Can speculative trading explain the trading-volume relation. *Journal of Business and Economic Statistics* **13**: 379–396.
- Gallant AR, Rossi PE, Tauchen G. 1992. Stock price and volume. *Review of Financial Studies* **5**: 199–242.
- Geweke J. 1992. Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In Bernardo JM, Berger JO, Dawid AP, Smith AFM (eds.) *Bayesian Statistics. Vol. 4*. Oxford, U.K.: Oxford University Press, 169–193.
- Glosten LR, Milgrom PR. 1985. Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* **14**: 71–100.

- Harris L. 1987. Transaction data test of the mixture distribution hypothesis. *The Journal of Financial and Quantitative Analysis* **22**: 127–141.
- Henze N. 1986. A probabilistic representation of the skew-normal distribution. *Scandinavian Journal of Statistics* **13**: 271–275.
- Jacquier E, Polson N, Rossi P. 1994. Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics* **12**: 371–418.
- Jacquier E, Polson N, Rossi P. 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. *Journal of Econometrics* **122**: 185–212.
- Karpoff JM. 1987. The relation between price changes and trading volume: A survey. *The Journal of Financial and Quantitative Analysis* **22**: 109–126.
- Kim S, Shepard N, Chib S. 1998. Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies* **65**: 361–393.
- Koopman S. 1993. Disturbance smoothers for State Space models. *Biometrika* **80**: 117–126.
- Liesenfeld R. 1998. Dynamic bivariate mixture models: modeling the behavior of prices and trading volume. *Journal Business & Economic Statistics* **12**: 101–109.
- Luu JC, Martens M. 2003. Testing the mixture-of-distributions hypothesis using “realized” volatility. *The Journal of Future Markets* **7**: 661–679.
- Mahieu R, Bauer R. 1998. A Bayesian analysis of stock return volatility and trading volume. *Applied Financial Economics* **8**: 671–687.
- Mittnik S, Paolella MS. 2000. Conditional density and value-at-risk prediction of asian currency exchange rates. *Journal of Forecasting* **19**: 313–333.
- Peiro A. 1999. Skewness in financial returns. *Journal of Banking and Finance* **23**: 847–862.

- Pemstein D, Quinn KV, Martin AD. 2007. The Scythe statistical library: An open source C++ library for statistical computation. *Journal of Statistical Software* **V**: 1–29.
- Pitt M, Shephard N. 1999. Filtering via simulation: Auxiliary particle filter. *Journal of the American Statistical Association* **94**: 590–599.
- Richardson M, Smith T. 1994. A direct test of the mixture distribution hypothesis: Measuring the daily flow of information. *The Journal of Financial and Quantitative Analysis* **15**: 183–202.
- Shephard N, Pitt M. 1997. Likelihood analysis of non-Gaussian measurements time series. *Biometrika* **84**: 653–667.
- Tauchen GE, Pitts M. 1983. The price variability-volume relationships in speculative markets. *Econometrica* **51**: 485–506.
- Taylor S. 1982. Financial returns modelled by the product of two stochastic processes—a study of the daily sugar prices 1961–75. In Anderson O (ed.) *Time Series Analysis: Theory and Practice, Vol 1*. Amsterdam: North-Holland, 203–226.
- Taylor S. 1986. *Modeling Financial Time Series*. Chichester: Wiley.
- Tierney L. 1994. Markov chains for exploring posterior distributions (with discussion). *Annals of Statistics* **21**: 1701–1762.
- Watanabe T. 2000. Bayesian analysis of dynamic bivariate mixture models: Can they explain the behavior of returns and trading volume? *Journal of Business and Economic Statistics* **18**: 199–210.
- Watanabe T, Omori Y. 2004. A multi-move sampler for estimate non-Gaussian time series model: Comments on Shepard and Pitt (1997). *Biometrika* **91**: 246–248.