

# Bayesian modeling of financial returns: A relationship between volatility and trading volume

Carlos A. Abanto-Valle<sup>1</sup>, Helio S. Migon<sup>1</sup> and Hedibert F. Lopes<sup>2</sup>

<sup>1</sup> *Instituto de Matemática, Universidade Federal do Rio de Janeiro, Brazil.*

*e-mails: cabantovalle@im.ufrj.br, migon@im.ufrj.br*

<sup>2</sup> *Graduate School of Business, University of Chicago, USA.*

*e-mail: hlopes@chicagogsb.edu*

## Abstract

The modified mixture model with Markov switching volatility specification is introduced to analyze the relationship between stock return volatility and trading volume. We propose to construct an algorithm based on Markov Chain Monte Carlo (MCMC) simulation methods to estimate all the parameters in the model using the Bayesian approach. The series of returns and trading volume of the British Petroleum stock will be analyzed.

**Keywords:** Stochastic volatility, Non linear and non Gaussian state space models, Markov process of first order, Markov Chain Monte Carlo

## 1 Introduction

The dynamics of the relationship between stock return volatility and trading volume has a long history in the finance literature. Karpoff (1987) provides a good survey of this literature, discussing the return-volume relation in various financial markets. In considering this problem, Clark (1973) started the discussion by presenting the intuitively appealing mixture of distributions hypothesis (MDH). According to the MDH, return and trading volume are driven by the same underlying latent information flow variable, i.e. price movements 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. 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 & Pitts, 1983; Richardson & Smith, 1994; and Foster & Wiswanathan, 1995).

There are several variants of the MDH in the literature. These include the models of Clark (1973), Tauchen & Pitts (1983), Harris (1987), and Andersen (1996). 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 AR(1) stochastic volatility process for the latent rate of information arrival, as proposed by Andersen (1994), by using the generalized method of moments (GMM). Mahieu & 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.

In this article we propose to expand the log volatility specification used in Mahieu & Bauer (1998) by introducing the Markov switching volatility specification proposed by So, Lam & Li (1998), which allows taking into consideration different volatility regimes.

The rest of the article is organized as follows: Section 2 presents the relation between stock return volatility and trading volume with the extended specifica-

tion for the log-volatility. Section 3 shows the Bayesian estimation procedure using MCMC methods. Section 4 shows an application using an artificial data set. Section 5 presents an empirical application on the return and trading volume series for the IBM stock. Finally, section 6 concludes.

## 2 The Model

Andersen (1996) develops an empirical return volatility-trading volume model using the theoretical framework of Glosten & 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 & Bauer (1998), leads to the following specification:

$$y_t | h_t \sim \mathcal{N}(0, e^{h_t}) \quad (1)$$

$$v_t | h_t \sim \mathcal{P}(m_0 + m_1 e^{h_t}), \quad m_0, m_1 > 0 \quad (2)$$

$$h_t = \alpha + \phi h_{t-1} + \eta_t \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2), \quad (3)$$

where  $y_t$ ,  $v_t$  and  $h_t$  are respectively the compounded return, the trading volume and the log volatility on day  $t$ . In equation (2),  $m_0$  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 (1)-(3) will be denoted as SV-VOL. Note that the univariate stochastic volatility model (SV) used extensively in financial literature (see Jacquier, Polson & Rossi, 1994; Kim, Shepard & Chib, 1998; Mahieu & Schotman, 1998 and Abanto-Valle, Lopes & Migon, 2007 among others) is specified by equations (1) and (3).

In this article we modify the SV-VOL model by allowing the log-volatility specification to incorporate regime-switching properties; that is, the parameter determining the level of the log-volatility is allowed occasional discrete shifts:

$$h_t = \alpha_{S_t} + \phi h_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \quad (4)$$

In (4), the switching dynamic is governed by a  $k$ -state first-order Markov process with transition probabilities  $p_{ij} = Pr(S_t = j \mid S_{t-1} = i)$ , where  $i, j = 1, \dots, k$  and  $\sum_{j=1}^k p_{ij} = 1$ . In this new specification  $S_t$  is the indicator variable showing the mean level of the state at time  $t$ , that is, the state indicator  $S_t$  defines a particular regime for the parameter values. In this article, we consider only two regimes,  $S_t = 0$  and  $S_t = 1$ , which indicate high and low volatility regimes respectively. In order to avoid identifiability problems, we assume that  $\alpha_0 = \gamma_0$  and  $\alpha_1 = \gamma_0 + \gamma_1 S_t$  with the restriction  $\gamma_1 < 0$  enforced for identification of each regime. The resulting MMM with Markov switching defined by equations (1),(2) and (4) will be denoted as MSSV-VOL.

### 3 MSSV-VOL model estimation using MCMC

A Bayesian approach to parameter estimation in the MSSV-VOL model defined by equations (1),(2) and (4) relies on MCMC techniques. We propose to construct an algorithm based on MCMC simulation methods to make the Bayesian analysis of all the parameters feasible.

The MSSV model, defined by equations (1) and (4), has been studied from a Bayesian viewpoint using MCMC methods. Multi-move samplers have been used to update the log-volatilities. For example, So et al. (1998) use the mixture sampler Gibbs sampling (Kim et al., 1998), and Shibata & Watanabe (2005) the block sampler Gibbs sampling (Shephard & Pitt, 1997 and Watanabe & Omori, 2004).

Let  $\boldsymbol{\theta} = (\gamma_0, \gamma_1, \phi, \sigma_\eta^2, m_0, m_1, p_{00}, p_{11}, \pi_0)'$  be the vector of parameters of the MSSV-VOL model,  $\mathbf{h}_{0:T} = (h_0, h_1, \dots, h_T)'$  be the vector of the log volatilities and  $\mathbf{S}_{0:T} = (S_0, S_1, \dots, S_T)'$  the states of the first order Markov process. Let  $\mathbf{y}_{1:T} = (y_1, \dots, y_T)'$  and  $\mathbf{v}_{1:T} = (v_1, \dots, v_T)'$  represent the information available up to time  $T$ .

The Bayesian approach for estimating the MSSV-VOL model uses the data augmentation principle, which considers  $\mathbf{h}_{0:T}$  and  $\mathbf{S}_{0:T}$  as latent parameters. By the use of Bayes' theorem, the joint posterior density of parameter and latent variables has the following decomposition

$$\begin{aligned} p(\mathbf{h}_{0:T}, \mathbf{S}_{0:T}, \boldsymbol{\theta} \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}) &\propto p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T})p(\mathbf{v}_{1:T} \mid \mathbf{h}_{0:T}, \boldsymbol{\theta}) \\ &\times p(\mathbf{h}_{0:T} \mid \mathbf{S}_{0:T}, \boldsymbol{\theta})p(\mathbf{S}_{0:T} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}) \end{aligned} \quad (5)$$

where

$$p(\mathbf{y}_{1:T} \mid \mathbf{h}_{0:T}) \propto \prod_{t=1}^T e^{-\frac{h_t + y_t^2}{2} e^{-h_t}} \quad (6)$$

$$p(\mathbf{v}_{1:T} \mid \mathbf{h}_{0:T}, \boldsymbol{\theta}) \propto \prod_{t=1}^T [m_0 + m_1 e^{h_t}]^{v_t} [e^{-m_0 - m_1 e^{h_t}}] \quad (7)$$

$$p(\mathbf{h}_{0:T} \mid \mathbf{S}_{0:T}, \boldsymbol{\theta}) \propto e^{-\frac{1-\phi^2}{2\sigma_\eta^2} (h_0 - \frac{\alpha S_0}{1-\phi})^2} \prod_{t=1}^T e^{-\frac{1}{2\sigma_\eta^2} (h_t - \alpha S_t - \phi h_{t-1})^2} \quad (8)$$

$$p(\mathbf{S}_{0:T} \mid \boldsymbol{\theta}) \propto \pi_{S_0} \prod_{t=1}^T p_{S_{t-1}S_t} \quad (9)$$

$$\pi_i = P(S_0 = i) \quad i = 0, 1. \quad (10)$$

For the unknown parameters in the MSSV-VOL model, the prior distributions are set as:  $\phi \sim \mathcal{N}_{[-1,1]}(\bar{\phi}, \sigma_\phi^2)$ ,  $\sigma_\eta^2 \sim \mathcal{IG}(\frac{T_0}{2}, \frac{M_0}{2})$ ,  $p_{ii} \sim \mathcal{Be}(\lambda_{i0}, \lambda_{i1}), i = 0, 1$ ,  $\pi_0 \sim \mathcal{Be}(l_0, l_1)$ ,  $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)' \sim N_{2(\gamma_1 < 0)}(\bar{\boldsymbol{\gamma}}, \bar{\mathbf{B}})$ ,  $m_0 \sim \mathcal{G}(a_0, b_0)$ ,  $m_1 \sim \mathcal{G}(a_1, b_1)$ , where as usual  $\mathcal{N}_{[.,.]}(\cdot, \cdot)$ ,  $\mathcal{IG}(\cdot, \cdot)$ ,  $\mathcal{Be}(\cdot, \cdot)$ ,  $\mathcal{G}(\cdot, \cdot)$ ,  $\mathcal{N}_{K(\cdot)}(\cdot, \cdot)$  represent the truncated normal, inverse gamma, beta, gamma and the K-multivariate truncated normal distribution respectively.

Since the posterior density  $p(\mathbf{h}_{0:T}, \mathbf{S}_{0:T}, \boldsymbol{\theta} \mid \mathbf{y}_{0:T}, \mathbf{v}_{0:T})$  does not have closed form, first we sample the parameters  $\boldsymbol{\theta}$  and next the latent variables  $\mathbf{S}_{0:T}$  and  $\mathbf{h}_{0:T}$  using Gibbs sampling. The sampling scheme is described by Algorithm 3.1.

**Algorithm 3.1**

1. Set  $i = 0$  and get starting values for the parameters  $\boldsymbol{\theta}^{(i)}$ , the states  $\mathbf{S}_{0:T}^{(i)}$  and  $\mathbf{h}_{0:T}^{(i)}$
2. Draw  $\boldsymbol{\theta}^{(i+1)} \sim p(\boldsymbol{\theta} \mid \mathbf{h}_{0:T}^{(i)}, \mathbf{S}_{0:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
3. Draw  $\mathbf{S}_{0:T}^{(i+1)} \sim p(\mathbf{S}_{0:T} \mid \boldsymbol{\theta}^{(i+1)}, \mathbf{h}_{0:T}^{(i)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
4. Draw  $\mathbf{h}_{0:T}^{(i+1)} \sim p(\mathbf{h}_{0:T} \mid \boldsymbol{\theta}^{(i+1)}, \mathbf{S}_{0:T}^{(i+1)}, \mathbf{y}_{1:T}, \mathbf{v}_{1:T})$
5. Set  $i = i + 1$  and return to 2 until achieving convergence.

As described by algorithm 3.1, the Gibbs sampler requires sampling parameters and latent variables from their full conditionals. Sampling the log-volatilities  $\mathbf{h}_{0:T}$  in step 4 is the most difficult task due to the non linear setup in equations (1) and (2). In order to avoid the higher correlations due to the Markovian structure of the  $h_t$ 's, we develop a multi-move sampler (Shephard & Pitt, 1997; Watanabe & Omori, 2004 and Omori & Watanabe, 2008) to sample the  $\mathbf{h}_{0:T}$  by blocks. Details on the full conditionals of  $\boldsymbol{\theta}$  and the latent variable  $\mathbf{S}_{0:T}$  are given in the appendix.

Let us consider the MSSV-VOL model defined by equations (1), (2) and (4). In order to simulate  $\mathbf{h}_{0:T}$ , we break the problem into two steps: first, we simulate  $h_0$  conditional on  $\mathbf{h}_{1:T}$  and next  $\mathbf{h}_{1:T}$  conditional on  $h_0$ . In our block sampler, we divide  $\mathbf{h}_{1:T}$  into  $K + 1$  blocks,  $\mathbf{h}_{k_{i-1}+1:k_i-1} = (h_{k_{i-1}+1}, \dots, h_{k_i-1})'$  for  $i = 1, \dots, K + 1$ , with  $k_0 = 0$  and  $k_{K+1} = T$ , where  $k_i - k_{i-1} \geq 2$  is the size of the  $i$ -th block. Following Shephard & Pitt (1997) and Omori & Watanabe (2008),

the  $K$  knots  $(k_1, \dots, k_K)$  are generated randomly using

$$k_i = \text{int}[T \times \{(i + u_i)/(K + 2)\}], \quad i = 1, \dots, K, \quad (11)$$

where the  $u_i$ 's are independent realizations of the uniform random variable on the interval  $(0,1)$  and  $\text{int}[x]$  represents the floor of  $x$ . We sample the block of disturbances  $\boldsymbol{\eta}_{k_{i-1}+1:k_i-1} = (\eta_{k_{i-1}+1}, \dots, \eta_{k_i-1})$  instead of  $\mathbf{h}_{k_{i-1}+1:k_i-1} = (h_{k_{i-1}+1}, \dots, h_{k_i-1})$ , exploiting the fact that the innovations  $\eta_t$  are *i.i.d.* with  $\mathcal{N}(0, \sigma_\eta^2)$ .

Suppose that  $k_{i-1} = t$  and  $k_i = t+k+1$  for the  $i$ -th block, such that  $t+k < T$ . Then  $\boldsymbol{\eta}_{t+1:t+k} = (\eta_{t+1}, \dots, \eta_{t+k})$  are sampled at once from their full conditional distribution  $f(\boldsymbol{\eta}_{t+1:t+k} | h_t, h_{t+k+1}, \mathbf{y}_{t+1:t+k}, \mathbf{S}_{t+1:t+k+1})$ , which is expressed in the log scale as

$$\begin{aligned} & \log f(\boldsymbol{\eta}_{t+1:t+k} | h_t, h_{t+k+1}, \mathbf{y}_{t+1:t+k}, \mathbf{v}_{t+1:t+k}, \mathbf{S}_{t+1:t+k+1}) \\ &= \text{const} - \frac{1}{2\sigma_\eta^2} \sum_{r=t+1}^{t+k} \eta_r^2 + \sum_{r=t+1}^{t+k} l(h_r) - \frac{1}{2\sigma_\eta^2} (h_{t+k+1} - \alpha_{S_{t+k+1}} - \phi h_{t+k})^2, \end{aligned} \quad (12)$$

where  $l(h_r)$  is the log of  $f(y_r, v_r | h_r)$  given by

$$l(h_r) = \text{const} - \frac{h_r}{2} - \frac{1}{2} y_r^2 e^{-h_r} - (m_0 + m_1 e^{h_r}) + v_r \log(m_0 + m_1 e^{h_r}).$$

Note that when  $t+k = T$ , the last term in (12) is omitted and denotes the first and second derivatives of  $l(h_r)$  with respect to  $h_r$  by  $l'$  and  $l''$ .

Applying a Taylor's series expansion to  $\sum_{r=t+1}^{t+k} l(h_r)$  in equation (12) around some preliminary estimate of  $\boldsymbol{\eta}_{t:t+k}$ , denoted by  $\hat{\boldsymbol{\eta}}_{t:t+k}$ , we have

$$\begin{aligned} & \log f(\boldsymbol{\eta}_{t+1:t+k} | h_t, h_{t+k+1}, \mathbf{y}_{t+1:t+k}, \mathbf{v}_{t+1:t+k}, \mathbf{S}_{t+1:t+k+1}) \\ & \approx \text{const} - \frac{1}{2\sigma_\eta^2} \sum_{r=t+1}^{t+k} \eta_r^2 - \frac{1}{2\sigma_\eta^2} (h_{t+k+1} - \alpha_{S_{t+k+1}} - \phi h_{t+k})^2 \\ & \quad + \sum_{r=t+1}^{t+k} \left\{ l(\hat{h}_r) + (h_r - \hat{h}_r) l'(\hat{h}_r) + \frac{1}{2} (h_r - \hat{h}_r)^2 l''_F(\hat{h}_r) \right\} \end{aligned} \quad (13)$$

As  $l(h_r)$  is not concave, we propose to use  $l''_F(h_r)$  in place of  $l''(h_r)$ , which can be positive for some values of  $h_r$ . To ensure that  $l''_F(h_r)$  is everywhere strictly negative<sup>1</sup>, it is defined as

$$l''_F(h_r) = \mathbb{E}[l''(h_r)] = -\frac{1}{2} - \frac{m_1^2 e^{2h_r}}{m_0 + m_1 e^{h_r}}. \quad (14)$$

The expectation in (14) is taken with respect to the joint density of  $y_r$  and  $v_r$  conditional on  $h_r$ .

After some simple but tedious algebra in (13), we have the approximating normal density  $g$  as follows

$$\begin{aligned} \log f(\boldsymbol{\eta}_{t+1:t+k} | h_t, h_{t+k+1}, \mathbf{y}_{t+1:t+k}, \mathbf{v}_{t+1:t+k}, \mathbf{S}_{t+1:t+k+1}) \\ = \text{const} - \frac{1}{2\sigma_\eta^2} \sum_{r=t+1}^{t+k} \eta_r^2 + \frac{1}{2} \sum_{r=t+1}^{t+k-1} l''_F(\hat{h}_r) \left( \hat{h}_r - \frac{l'(\hat{h}_r)}{l''_F(\hat{h}_r)} - h_r \right)^2 \\ - \frac{\phi^2 - l''_F(\hat{h}_{t+k})\sigma_\eta^2}{2\sigma_\eta^2} \left\{ \frac{\sigma_\eta^2}{\phi^2 - l''_F(\hat{h}_{t+k})} \left( l'(\hat{h}_{t+k}) - l''_F(\hat{h}_{t+k})\hat{h}_{t+k} + \frac{\phi - \alpha_{S_{t+k+1}}}{\sigma_\eta^2} h_{t+k+1} \right) - h_{t+k} \right\}^2 \\ = \log g, \end{aligned} \quad (15)$$

where  $\hat{\mathbf{h}}_{t+1:t+k}$  is the estimate of  $\mathbf{h}_{t+1:t+k}$  corresponding to  $\hat{\boldsymbol{\eta}}_{t+1:t+k}$ .

From (15), we define auxiliary variables  $d_r$  and  $\hat{y}_r$  for  $r = t+1, \dots, t+k-1$  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  $r = t+k < T$ ,

$$\begin{aligned} d_r &= \frac{\sigma_\eta^2}{\phi - \sigma_\eta^2 l''_F(\hat{h}_{t+k})} \quad \text{and} \\ \hat{y}_r &= d_r \left[ l'(\hat{h}_r) - l''_F(\hat{h}_r)\hat{h}_r + \frac{(\phi - \alpha_{S_{r+1}})}{\sigma_\eta^2} h_{r+1} \right]. \end{aligned} \quad (17)$$

---

<sup>1</sup>In the context of the SV-VOL model, Watanabe (2000) uses an alternative expression,  $\min\{l''(h_r), -0.001\}$ , to ensure the strictly negative condition on  $l''_F(h_r)$



When  $r = t + k = T$ , we use (16) to define the auxiliary variables.

The resulting normalized density in (15),  $g$ , is a  $k$ -dimensional normal density, which is the exact density of  $\boldsymbol{\eta}_{t+1:t+k}$  conditional on  $\hat{\boldsymbol{y}}_{t+1:t+k}$  in the linear Gaussian state space model:

$$\hat{y}_r = h_r + \epsilon_r, \quad \epsilon_r \sim N(0, d_r), \quad (18)$$

$$h_r = \alpha_{s_r} + \phi h_{r-1} + \eta_r, \quad \eta_r \sim N(0, \sigma_\eta^2). \quad (19)$$

Applying the de Jong & Shepard (1995) simulation smoother to this model with the artificial  $\hat{\boldsymbol{y}}_{t+1:t+k}$  enables us to sample  $\boldsymbol{\eta}_{t+1:t+k}$  from the density  $g$ . Since  $g$  does not bound  $f$ , we use the Metropolis-Hastings acceptance-rejection algorithm to sample from  $f$ .

We select the expansion block  $\hat{\mathbf{h}}_{t+1:t+k}$  as follows. Once an initial expansion block  $\hat{\mathbf{h}}_{t+1:t+k}$  is selected, we can calculate the artificial  $\hat{\boldsymbol{y}}_{t+1:t+k}$ . Then, applying the Kalman filter and a disturbance smoother to the linear Gaussian state space model consisting of equations (18) and (19) with the artificial  $\hat{\mathbf{h}}_{t+1:t+k}$  yields the mean of  $\hat{\mathbf{h}}_{t+1:t+k}$  conditional on  $\hat{\boldsymbol{y}}_{t+1:t+k}$  in the linear Gaussian state space model, which is used as the next  $\hat{\mathbf{h}}_{t+1:t+k}$ . We use five iterations of this procedure to obtain a good sequence of  $\hat{\mathbf{h}}_{t+1:t+k}$  to use as the expansion block. The procedure is summarized in algorithm 3.2.

### Algorithm 3.2

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. Define the auxiliary variables  $y_r$  and  $d_r$  using equations (16) or (17) for  $r = t + 1, \dots, t + k$ .
3. Consider the linear Gaussian state-space model in (18) and (19). Apply the

Kalman Filter and a disturbance smoother (Koopman, 1993) and obtain the posterior mean of  $\boldsymbol{\eta}_{t:t+k}$  ( $\mathbf{h}_{t:t+k}$ ) and set  $\hat{\boldsymbol{\eta}}_{t:t+k}$  ( $\hat{\mathbf{h}}_{t:t+k}$ ) to this value.

4. Return to step 2 and repeat the procedure five times.

## 4 Numerical illustration with artificial data set

In order to assess the performance of the MCMC algorithms described in the previous section, we present results based on a simulated data set. All the calculations were performed running our own code implemented in C++ using the Scythe statistical library<sup>2</sup>(Pemstein, Quinn & Martin, 2007), on an Intel Pentium 4 +2.8 GHz with 1 GB of RAM. We simulated a data set of 1500 observations using  $\gamma_0 = -0.5$ ,  $\gamma_1 = -0.25$ ,  $\phi = 0.7$ ,  $\sigma_\eta^2 = 0.2$ ,  $m_0 = 0.85$ ,  $m_1 = 0.15$ ,  $p_{00} = 0.98$  and  $p_{11} = 0.98$ .

We set the prior distributions as:  $\phi \sim \mathcal{N}_{[-1,1]}(0.95, 10)$ ,  $\sigma_\eta^2 \sim \mathcal{IG}(\frac{5}{2}, \frac{0.1}{2})$ ,  $p_{00} \sim \mathcal{Be}(50, 1.5)$ ,  $p_{11} \sim \mathcal{Be}(1.5, 50)$ ,  $\pi_0 \sim \mathcal{Be}(0.5, 0.5)$ ,  $\boldsymbol{\gamma} \sim N_{2(\gamma_1 < 0)}(\bar{\boldsymbol{\gamma}}, \bar{\mathbf{B}})$ ,  $m_0 \sim \mathcal{G}(0.08, 0.1)$ ,  $m_1 \sim \mathcal{G}(1, 10)$ , where  $\bar{\boldsymbol{\gamma}} = (-0.5, -0.25)'$  and  $\bar{\mathbf{B}} = \text{diag}(4.0, 1.0)$ .

The number of blocks, K, in the block sampler was set equal to 60, so that each block contained 25  $h_t$ 's on average. We conducted the MCMC simulation for 40000 iterations. The first 10000 draws were discarded as a burn-in period, and then the next 30000 were recorded. Table 1 reports the posterior means, the Monte Carlo (MC) error of the posterior means, the 95% intervals and the convergence diagnostic (CD) statistics proposed by Geweke (1992) for all the parameters. The 95% credibility intervals are estimated using the 2.5th and the 97.5th percentiles of the posterior samples.

---

<sup>2</sup>The Scythe statistical library is available for free at the website <http://scythe.wustl.edu>. It is an open source C++ library for statistical computation. It includes a suite of matrix manipulation functions, a suite of random number generators and a suite of numerical optimizers.

Parameter	True	Mean	MC error	95% interval	CD
$\gamma_0$	-0.50	-0.5367	0.0105	(-0.9160,-0.2868)	-0.16
$\gamma_1$	-0.25	-0.2401	0.0069	(-0.5492,-0.0201)	-0.16
$\phi$	0.70	0.6861	0.0055	(0.4450, 0.8202)	-0.72
$\sigma_\eta^2$	0.20	0.2677	0.0033	(0.1728, 0.3965)	-0.22
$p_{00}$	0.98	0.9698	0.0004	(0.9175, 0.9966)	-0.03
$p_{11}$	0.98	0.9717	0.0007	(0.9224, 0.9954)	-1.32
$\pi_0$	0.50	0.4925	0.0024	(0.0588,0.9344)	-0.15
$m_0$	0.85	0.8332	0.0003	(0.7779, 0.8855)	-0.54
$m_1$	0.15	0.0880	0.0017	(0.0023, 0.3105)	0.98

Table 1: Simulated data set: posterior mean, standard error of the posterior mean, 95% interval and convergence diagnostic(CD) for the MSSV-VOL.

The proposed algorithm is evaluated in terms of how well it estimates the true parameter values. It can be seen that the estimated results for the parameters appear quite reasonable. In Figure 1 all the 95% credibility intervals include true values. All parameters passed the convergence diagnostic test of Geweke (1992) and also Heidelberger & Welch (1983), although the last one is not reported.

In Figure 2, the smoothed mean of  $e^{\frac{h_t}{2}}$  calculated from MCMC output is showed (dotted line). It is compared with the true volatilities (solid line), showing that the estimated values follow the behavior of the true volatilities.

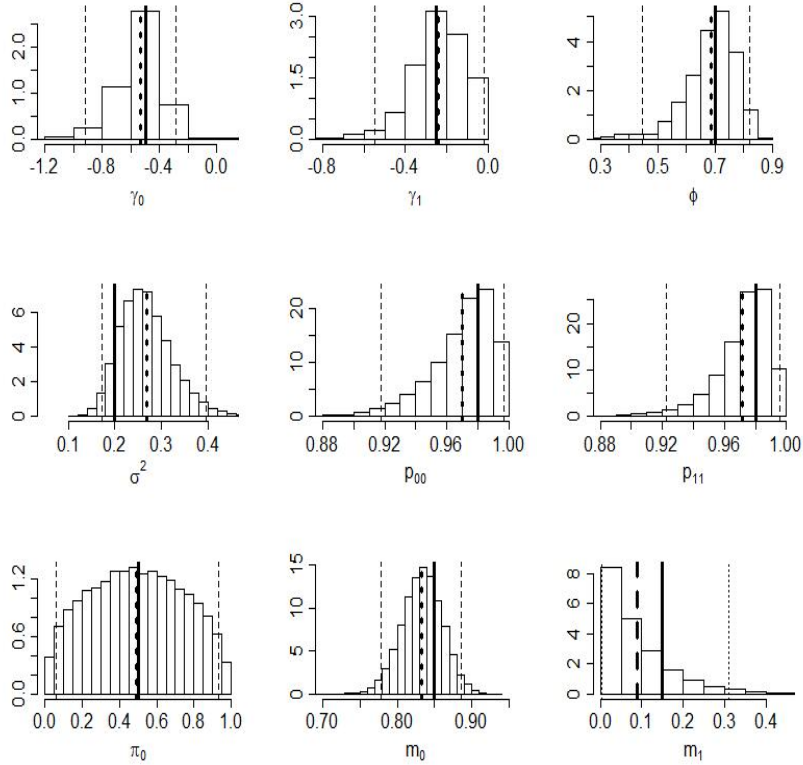


Figure 1: Simulated data set: Histograms of the parameters from MCMC output for the MSSV-VOL model. The tiny dotted and the dotted line indicate the 2.5% percentile, the posterior mean and the 97.5% percentile respectively. The solid line indicates the true value

Figure 3 shows the probability that  $S_t$  are in the high volatility period (top), and the true values of state indicator variable  $S_t$  versus the estimated values using the MSSV-VOL model (bottom). In the majority of cases the state indicators are well estimated.<sup>3</sup>

---

<sup>3</sup>The state indicators and  $P(S_t = 0)$  obtained with the MSSV model give similar results, but they are not reported.

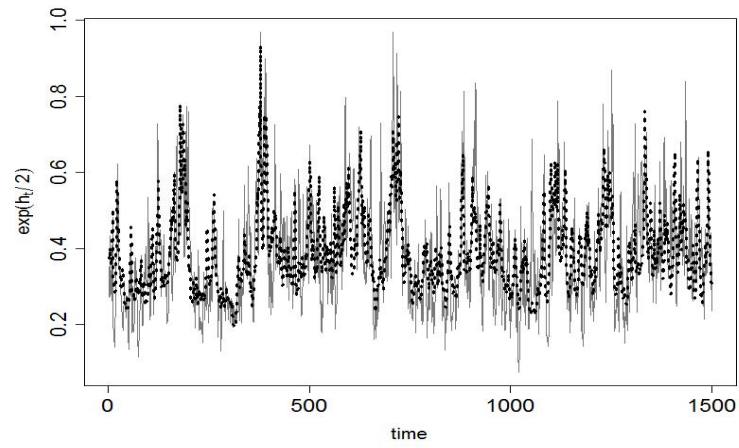


Figure 2: Simulated data set. True volatilities (solid line) vs estimated smoothed mean of  $e^{\frac{h_t}{2}}$  by the MSSV-VOL model (dotted line).

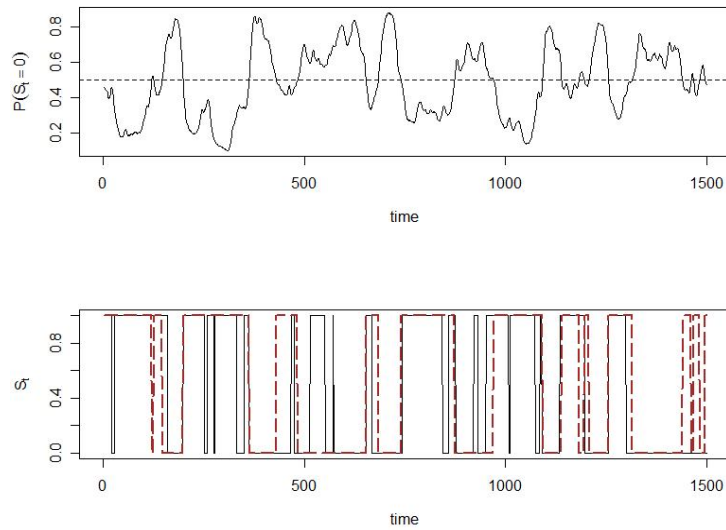


Figure 3: Simulated data set: Top:  $P(S_t = 0)$ . Bottom: True states  $S_t$ . (solid line) vs estimated  $S_t$  by the MSSV-VOL model (dotted line).

## 5 Empirical Application

This section analyzes the daily closing prices and trading volume corrected by dividends and stock splits for the British Petroleum Company stock series (BP) listed on the London Stock Exchange (LSE) <sup>4</sup>. The analyzed period starts January 5, 1999 and ends July 17, 2008, yielding 2398 observations. Throughout we work with the mean corrected returns computed as

$$y_t = 100 \left\{ (\log P_t - \log P_{t-1}) - \frac{1}{T} \sum_{j=1}^T (\log P_j - \log P_{j-1}) \right\}$$

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  $10^{-6}$ .

	Mean	S.D.	Max	Min	Skewness	Kurtosis
Returns	0.00	1.57	8.33	-7.92	-0.12	4.98
Volume	3.33	1.43	17.41	0.63	2.17	12.77

Table 2: Summary statistics for BP stock series

Table 2 summarize descriptive statistics for the corrected compounded return and the detrended trading volume; the time series plots are shown in Figure 4. For the return series, the basic statistics, the mean, standard deviation, skewness and kurtosis are given as 0.00, 1.57, -0.12 and 4.98, respectively. Note that the

<sup>4</sup>The data set was obtained from the Yahoo finance web site at <http://finance.yahoo.com>.

kurtosis of the returns is above three, so that daily BP stock returns not likely to follow a normal distribution.

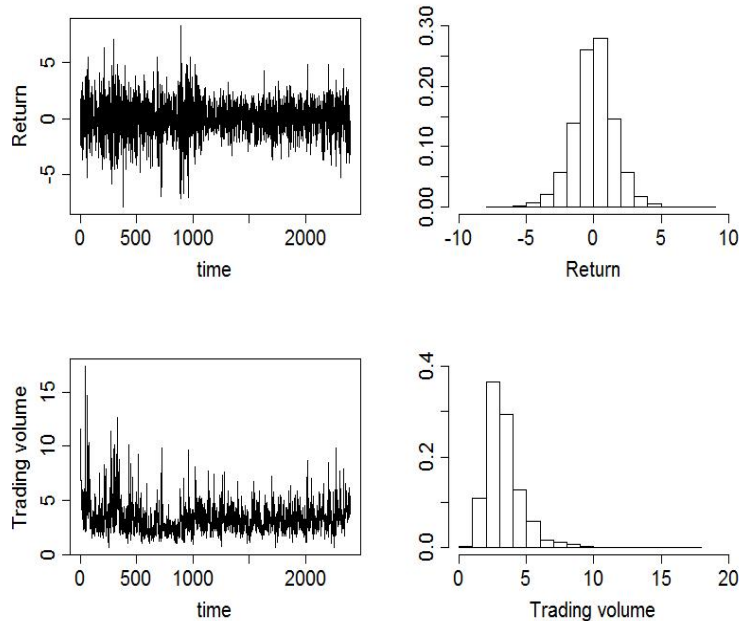


Figure 4: BP data set with sample period from January 5, 1999 to July 17, 2007. Top: raw series and histogram of corrected returns. Bottom: raw series and histogram of trading volume.

We fitted the SV, MSSV, SV-VOL and the MSSV-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 3. For the SV and SV-VOL models we set the prior distributions as:  $\alpha \sim \mathcal{N}(0, 10)$ ,  $\phi \sim \mathcal{N}_{[-1,1]}(0.95, 10)$ ,  $\sigma_\eta^2 \sim \mathcal{IG}(\frac{5}{2}, \frac{0.05}{2})$ ,  $m_0 \sim \mathcal{G}(0.08, 0.1)$  and  $m_1 \sim \mathcal{G}(1, 10)$ . We set  $K$ , the number of blocks, as 60, in a such way that each block contained 40  $h_t$ 's on average. We conducted the MCMC simulation for 60000 iterations. The first 10000 draws were discarded as a burn-in period, and then the next 50000 were recorded. Using these 50000

draws we calculated the posterior means, the Monte Carlo (MC) error of the posterior means, the 95% intervals and the convergence diagnostic (CD). Tables 3 and 4 summarize the results. According to the CD values, the null hypothesis that the sequence of 50000 draws is stationary is accepted at the 5% level for all the parameters in all the models considered here.

SV model				
Parameter	Mean	MC error	95% interval	CD
$\alpha$	0.0429	0.0005	(0.0233, 0.0674)	0.011
$\phi$	0.9372	0.0007	(0.9044,0.9629)	0.060
$\sigma_{\eta}^2$	0.0588	0.0009	(0.0353,0.0940)	-0.336
MSSV model				
Parameter	Mean	MC error	95% interval	CD
$\gamma_0$	0.3026	0.0047	(0.1848,0.4972)	0.160
$\gamma_1$	-0.2456	0.0040	(-0.4094, -0.1414)	-0.083
$\phi$	0.7506	0.0035	( 0.6133, 0.8349)	-0.147
$\sigma_{\eta}^2$	0.1456	0.0014	(0.1010, 0.2080)	0.522
$p_{00}$	0.9928	0.0001	(0.9819,0.9987)	-0.074
$p_{11}$	0.9874	0.0001	( 0.9875,0.9987)	0.571
$\pi_0$	0.6013	0.0013	(0.1331,0.9588)	-1.970

Table 3: PB data set: posterior mean, MC error, 95% interval and convergence diagnostic(CD) for the SV (top) and the MSSV (bottom) models.

The top of Table 3 reports the estimation results of SV model. The posterior mean and 95% interval of  $\phi$  are 0.9372 and (0.9044,0.9629) respectively, exhibiting high persistence in return volatility. The persistence parameter in the MSSV



model,  $\phi$ , drops from 0.9372 to 0.7506, which agrees with previous results in MSSV models (So et al., 1998; Shibata & Watanabe, 2005 and Carvalho & Lopes, 2007). On the other hand, the posterior mean and the 95% interval of  $\sigma_\eta^2$  are 0.1456 and (0.1010, 0.2080), which are higher than the 0.0588 and (0.0353,0.0940) in the SV model.

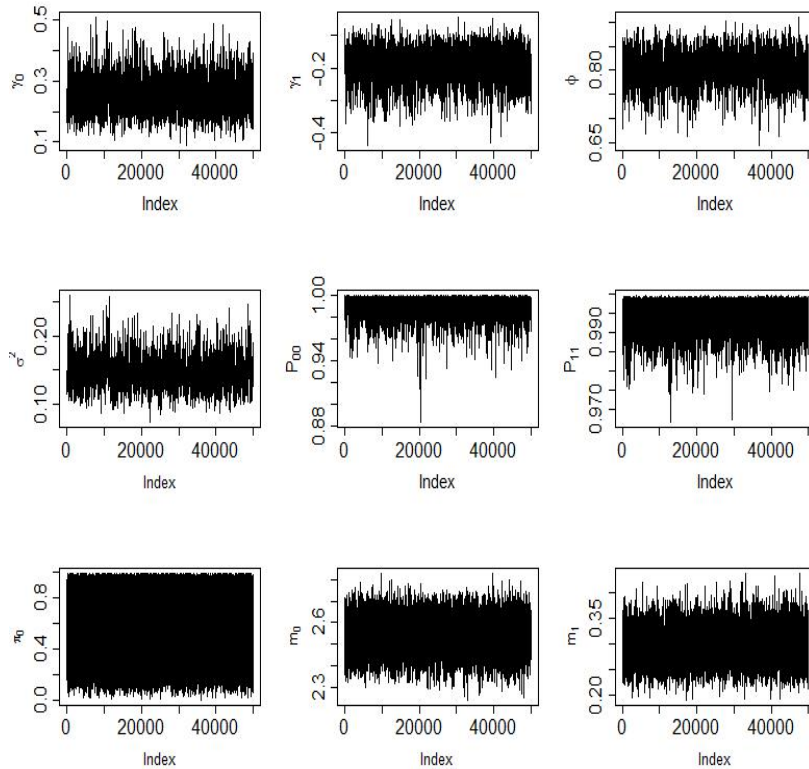


Figure 5: MSSV-VOL model, BP data set: sample paths for parameters obtained from MCMC output.

In the SV-VOL model, we find that posterior mean and 95% interval of  $\phi$  are 0.9281 and (0.9027,0.9503), respectively, both of which are slightly smaller than those in the SV model. 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 posterior mean of  $m_0$  is 2.5924 and the distribution of  $m_1$  has a posterior mean of 0.2655.

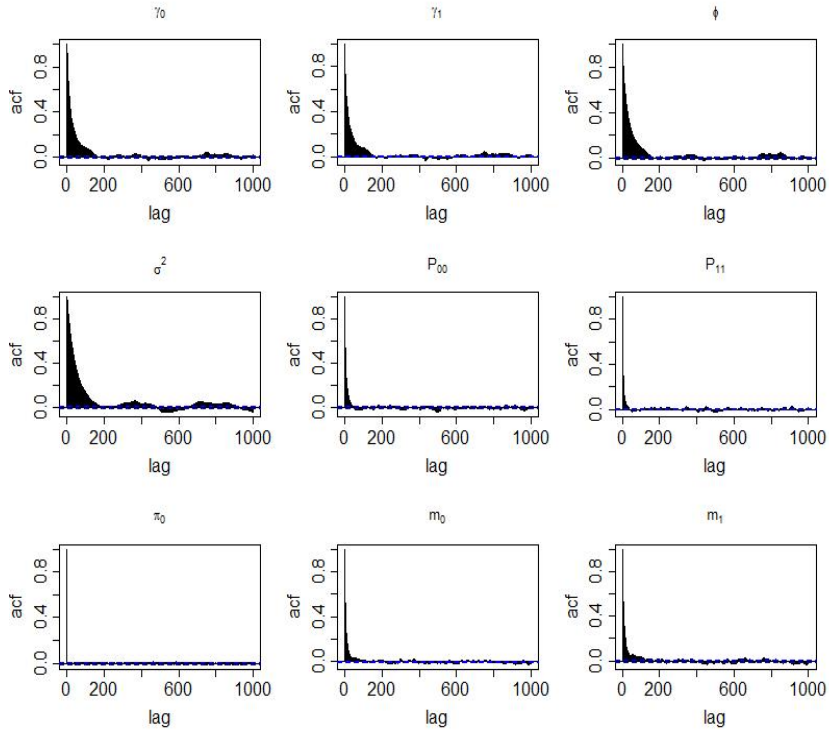


Figure 6: MSSV-VOL model, BP data set: autocorrelation functions for parameters obtained from MCMC output.

We now consider the estimation based on the the MSSV-VOL model. In Figures 5, 6 and 7, we show the sample paths, the autocorrelation function and the histograms for parameters respectively, showing that the persistence parameter,  $\phi$ , drops from 0.9281 to 0.8052, when compared with the SV model. This value, 0.8052, is slightly greater than the posterior mean of the MSSV model (0.7506). The 95% posterior interval of  $\phi$  is different than that obtained from the MSSV model. The posterior means of  $m_0$  and  $m_1$  are 2.5422 and 0.2952. They are similar to the values in the SV-VOL model and also, the posterior 95% intervals (see Table 4). The posterior means of the transition probabilities  $p_{00}$  and  $p_{11}$

SV-VOL model				
Parameter	Mean	MC error	95% interval	CD
$\alpha$	0.0484	0.0002	(0.0295, 0.0696)	-0.6007
$\phi$	0.9281	0.0003	(0.9027,0.9503)	0.7282
$\sigma_\eta^2$	0.0934	0.0005	(0.0672, 0.1254)	-1.1420
$m_0$	2.5924	0.0011	(2.4442, 2.7320)	-1.1220
$m_1$	0.2655	0.0005	(0.2120,0.3255)	1.0370
MSSV-VOL model				
Parameter	Mean	MC error	95% interval	CD
$\gamma_0$	0.2576	0.0018	(0.1713,0.3660)	0.2178
$\gamma_1$	-0.1970	0.0015	(-0.2907,-0.1202)	-0.0577
$\phi$	0.8052	0.0013	(0.7386,0.8609)	-0.1518
$\sigma_\eta^2$	0.1465	0.0011	(0.1070, 0.1956)	0.7103
$p_{00}$	0.9898	0.0001	(0.9724,0.9977)	0.2312
$p_{11}$	0.9949	0.00004	(0.9877,0.9987)	0.3383
$\pi_0$	0.6241	0.0011	(0.1727,0.9615)	0.5141
$m_0$	2.5422	0.0013	(2.3992,2.6791)	0.3715
$m_1$	0.2952	0.0007	(0.2387,0.3592)	-0.8472

Table 4: PB data set: posterior mean, standard error of the posterior mean, 95% interval and convergence diagnostic(CD) for SV-VOL (top) and the MSSV-VOL (bottom).

are 0.9898 and 0.9949, respectively, which are slightly different than these the values 0.9928 and 0.9874 from the MSSV model, indicating that the probability of switching between high- and low-volatility states is quite low. In the MSSV-VOL model, a volatility shock lasts about 98 days in the high-volatility state compared to about 196 days in the low-volatility state. In the MSSV model, a volatility shock lasts about 138 and 79 days respectively. The duration of the shock in state  $i$  is obtained as  $(1 - p_{ii})^{-1}$ .

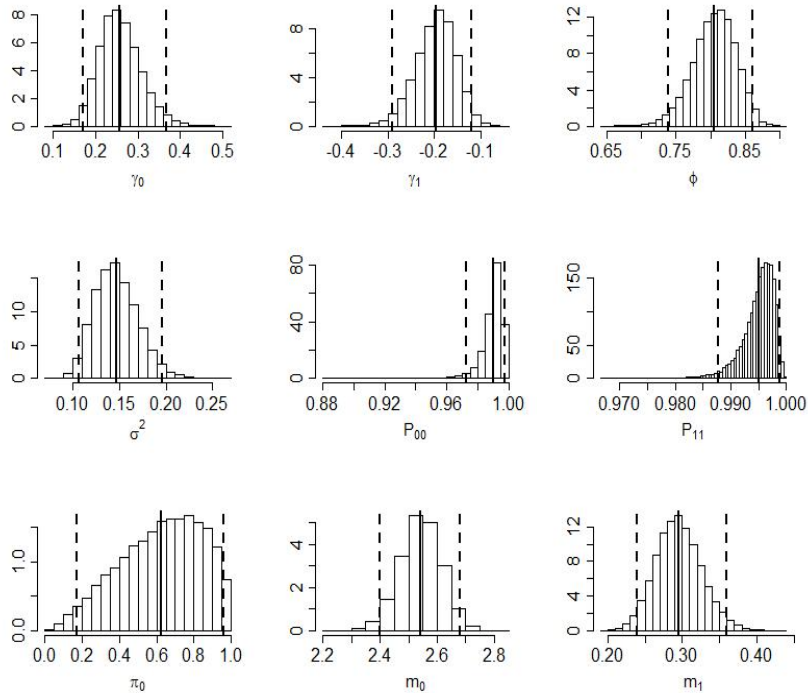


Figure 7: MSSV-VOL model, BP data set: histograms for parameters obtained from MCMC output. The dotted lines indicate the 2.5% and 97.5% percentiles, respectively, and the solid line the posterior mean.

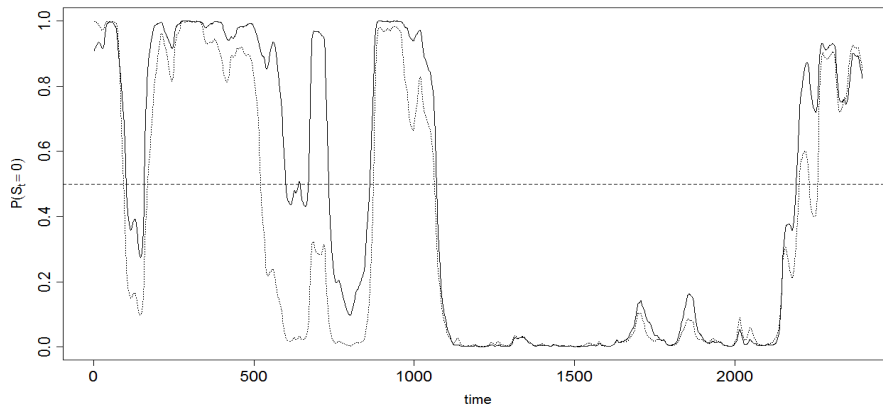


Figure 8: BP data set: Posterior probability of high-volatility state  $P(S_t = 0)$ . Solid line MSSV model, dotted line MSSV-VOL model.

Figure 8 depicts the posterior probabilities of the high-volatility state as inferred from the MSSV-VOL model (dotted line). Following Hamilton (1988), we consider an observation as belonging to a high-volatility state if the smoothed probability is higher than 0.5. Then, the high-volatilities periods are: 01/06/1999-05/17/1999, 09/03/1999-01/25/2001, 10/04/2007-11/16/2007 and 12/27/2007-07/17/2008. The last two regimes of high volatility are explained by a series of events that caused the oil price to exceed \$92/barrel by October 2007, and \$99.29/barrel for December futures in New York on November 21, 2007. Throughout the first half of 2008, oil regularly reached record high prices. On February 29, 2008, oil prices peaked at \$103.05 per barrel, and reached \$110.20 on March 12, 2008, the sixth record in seven trading days. Prices on June 27, 2008, touched \$141.71/barrel, for August delivery in the New York Mercantile Exchange. The most recent price per barrel maximum of \$147.02 was reached on July 11, 2008. Figure 8 (solid line), shows the posterior probabilities of high-volatility state from the MSSV model, There is a significant difference between the period from 08/29/2001-12/03/2001, in which occurred the September 11

attacks. The trading volume gives us information to modify the regimes in the MSSV model. We found similar results for other stocks as IBM, Coca Cola and Kodak, although the results are not reported here.

In Figure 9 we show the smoothed mean of  $e^{h_t}$  obtained from the MCMC output for both models. From a practical point of view, we are mainly interested in whether we find a significant difference between the two series. Therefore, we plot the difference between the smoothed means of the two series in the lower panel of Figure 9. This graph shows us that these series do not differ very much in most periods, but in some periods of high volatility, we observe difference in squared percentages of more than 4%. This can have a substantial impact, for instance, in the valuation of derivative instruments and several strategic or tactical asset allocation topics.

To assess the goodness of the estimated models, we calculate the deviance information criterion, DIC (for more details about the DIC criterion see for example Spiegelhalter, Best, Carlin & van der Linde, 2002; Berg, Meyer & Yu, 2004 and Celeux, Forbes, Robert & Titterton, 2006). In this context,  $p_D$  is a measure of model complexity. As the quantity of information available is different between models with and without trading volume, we compare the SV and MSSV models and the SV-VOL and MSSV-VOL models. From Table 5, according with the DIC the SV model better fit the return series of the BP stock. Model comparison between the SV-VOL and MSSV-VOL gives us the MSSV-VOL as the better model for joint modeling of return and trading volume series according to the DIC criterion. Table 5 shows that the  $p_D$  for SV-VOL and MSSV-VOL models are greater than the  $p_D$  for the SV and MSSV models.

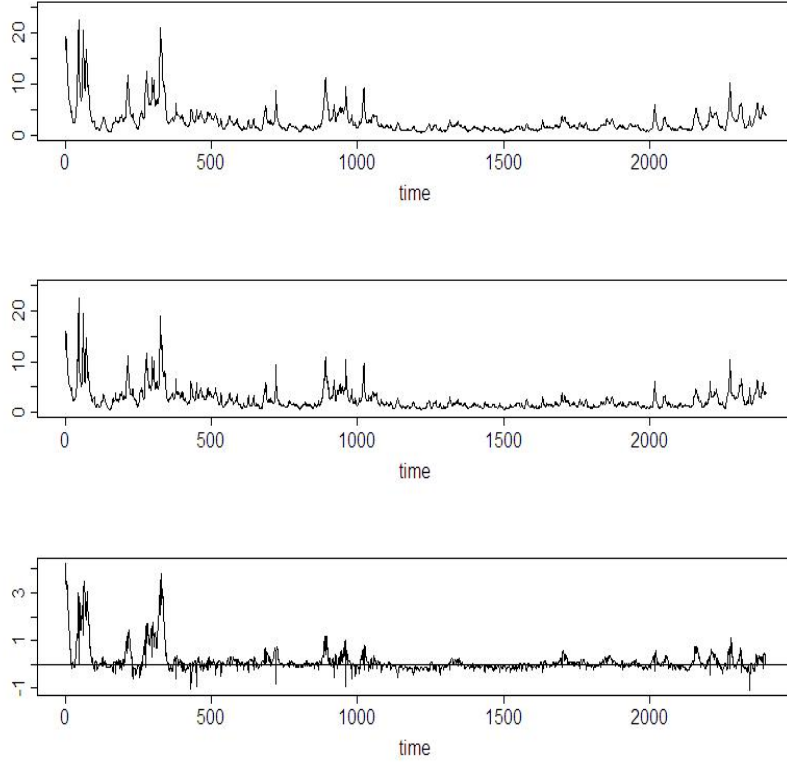


Figure 9: PB data set. Top: Smoothed mean of  $e^{h_t}$  from SV-VOL. Middle: smoothed mean of  $e^{h_t}$  from MSSV-VOL. Bottom: difference from both models.

Model	DIC	$p_D$	Rank
SV	8614.05	183.60	1
MSSV	8615.98	249.91	2
SV-VOL	5138.63	303.92	2
MSSV-VOL	5074.81	328.25	1

Table 5: BP data set. DIC: deviance information criterion,  $p_D$ : effective number of parameters.

## 6 Conclusions

This article studies the joint distribution of daily returns and trading volume based on the modified mixture model with Markov switching volatility specification. 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. Our estimation result shows that the estimate of the persistence parameter drops and the estimate of the variance error rises in the volatility specification.

This article makes certain contributions, but several extensions are still possible. First, we focus on normal distributions for  $\epsilon_t$ , but other distributions, such as the Student-t, or skewed distributions, such as the skewed normal or the skewed t distribution can be used. Second, we specify the log of volatility as a simple AR(1) process with Markov switching, but more elaborate models, such as long memory models, may be required to specify volatility. On the other hand, 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.



## A Appendix: The Full conditionals

In this appendix, we described the full conditional distributions for the parameters and the latent  $\mathbf{S}_{0:T}$  indicators of the MSSV-VOL model.

### A.1 Full conditional distribution of $\phi$ , $\gamma$ and $\sigma^2$

The prior distributions are set as:  $\phi \sim \mathcal{N}_{[-1,1]}(\bar{\phi}, \sigma_\phi^2)$ ,  $\sigma_\eta^2 \sim \mathcal{IG}(\frac{T_0}{2}, \frac{M_0}{2})$ ,  $\gamma \sim \mathcal{N}_{2(\gamma_1 < 0)}(\bar{\gamma}, \bar{B})$ . According with (8), we have the following full conditional for  $\phi$ :

$$p(\phi \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{0:T}, \gamma) \propto Q_\phi \exp\left\{-\frac{a}{2\sigma_\eta^2}(\phi - \frac{b}{a})^2\right\} I_\phi$$

where  $Q_\phi = \sqrt{1 - \phi^2} \exp\left\{-\frac{1}{2\sigma_\eta^2}[(1 - \phi^2)(h_0 - \frac{\alpha_{S_0}}{1 - \phi})^2]\right\}$ ,  $a = \sum_{t=1}^T h_{t-1}^2 + \frac{\sigma_\eta^2}{\sigma_\phi^2}$ ,  $b = \sum_{t=1}^T h_{t-1}(h_t - \alpha_{S_t}) + \bar{\phi} \frac{\sigma_\eta^2}{\sigma_\phi^2}$  and  $I_\phi$  is an indicator variable. As  $p(\phi \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{0:T}, \gamma)$  does not have closed form, we sample from using the Metropolis-Hastings algorithm with proposal density the truncated  $\mathcal{N}_{[-1,1]}(\frac{b}{a}, \frac{\sigma_\eta^2}{a})$ .

To calculate the conditional posterior of  $\gamma$ , we define:

$$\mathbf{Z} = \mathbf{X}\gamma + \eta,$$

with  $\mathbf{Z}' = ([\frac{1-\phi^2}{\sigma^2}]^{1/2}h_0, \frac{1}{\sigma}[h_1 - \phi h_0], \dots, \frac{1}{\sigma}[h_T - \phi h_{T-1}])$

$$\mathbf{X}' = \begin{pmatrix} \left[ \frac{1-\phi^2}{\sigma^2} \right]^{1/2} & \frac{1}{\sigma} & \dots & \frac{1}{\sigma} \\ \left[ \frac{1-\phi^2}{\sigma^2} \right]^{1/2} S_0 & \frac{1}{\sigma} S_1 & \dots & \frac{1}{\sigma} S_T \end{pmatrix}$$

and  $\eta \sim \mathcal{N}(0, I_{K+1})$ . Then, we have that the full conditional for  $\gamma$  is given by  $\mathcal{N}_{2(\gamma_1 < 0)}(\boldsymbol{\mu}_\gamma, B_1)$ , where  $\boldsymbol{\mu}_\gamma = (\bar{\mathbf{B}}^{-1} + \mathbf{X}'\mathbf{X})^{-1}(\bar{\mathbf{B}}^{-1}\bar{\gamma} + \mathbf{X}'\mathbf{Z})$  and  $B_1 = (\bar{\mathbf{B}}^{-1} + \mathbf{X}'\mathbf{X})^{-1}$ .

From (8), the conditional posterior of  $\sigma_\eta^2$  is  $\mathcal{IG}(\frac{T_1}{2}, \frac{M_1}{2})$ , where  $T_1 = T_0 + T + 1$  and  $M_1 = M_0 + [(1 - \phi^2)(h_0 - \frac{\alpha_{S_0}}{1 - \phi})^2] + \sum_{t=1}^T (h_t - \alpha_{S_t} - \phi h_{t-1})^2$ .

## A.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  $m_1 \sim \mathcal{G}(a_1, b_1)$ . Then the full the conditionals follows:

$$\begin{aligned} p(m_0 \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, m_1) &\propto \exp\{-(b_0 + T)m_0\} \\ &\times \exp\left\{\sum_{t=1}^T \log[m_0^{a_0-1/T} (m_0 + m_1 \exp(h_t))^{v_t}]\right\} \end{aligned}$$

$$\begin{aligned} p(m_1 \mid \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, m_0) &\propto \exp\left\{-m_1\left(b_1 + \sum_{t=1}^T \exp h_t\right)\right\} \\ &\times \exp\left\{\sum_{t=1}^T \log[m_1^{a_1-1/T} (m_0 + m_1 \exp(h_t))^{v_t}]\right\}. \end{aligned}$$

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 are  $\mathcal{N}_{(m_i > 0)}(\mu_{m_i}, \tau_{m_i}^2)$ , 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.

## A.3 Full conditionals of $p_{00}$ , $p_{11}$ and $\pi_0$

The prior distributions for  $p_{00}$ ,  $p_{11}$  and  $\pi_0$  are respectively given by  $p_{ii} \sim \mathcal{Be}(\lambda_{i_0}, \lambda_{i_1})$ ,  $i = 0, 1$  and  $\pi_0 \sim \mathcal{Be}(l_0, l_1)$ . Then, the full conditional posteriors are:  $p_{ii} \sim \mathcal{Be}(\lambda_{i_0}^*, \lambda_{i_1}^*)$  and  $\pi_0 \sim \mathcal{Be}(l_0^*, l_1^*)$ , where  $\lambda_{i_j}^* = \sum_{t=1}^T I(S_{t-1} = i, S_t = j) + \lambda_{i_j}$  for  $i, j = 0, 1$ . and  $l_i^* = l_i + I(S_0 = i)$ , for  $i = 0, 1$ .

## A.4 Full conditional of $\mathbf{S}_{0:T}$

The states  $\mathbf{S}_{0:T}$  are simulated as a block as proposed by Carter & Kohn (1994) and employed by So et al. (1998). The joint conditional distribution,  $p(\mathbf{S}_{0:T} \mid \theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T})$ , can be decomposed as:

$$\begin{aligned} p(\mathbf{S}_{0:T} \mid \theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}) &\propto p(S_T \mid \theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}) \\ &\times \prod_{t=0}^{T-1} p(S_t \mid \theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{t+1:T}) \end{aligned}$$

where  $\mathbf{S}_{t+1:T} = (S_{t+1}, \dots, S_T)'$ . This equation provides a scheme to draw  $S_T$  from  $p(S_T \mid \theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T})$  and then, recursively,  $S_t$  is generated from  $p(S_t \mid \theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}, \mathbf{h}_{0:T}, \mathbf{S}_{t+1:T})$ , for  $t = T - 1, \dots, 0$ . Suppressing for notational convenience the dependence on  $\theta, \mathbf{y}_{1:T}, \mathbf{v}_{1:T}$ , we have that

$$\begin{aligned} p(S_t \mid \mathbf{h}_{0:T}, \mathbf{S}_{t+1:T}) &= p(S_t \mid \mathbf{h}_{0:t}, S_{t+1}) \\ &= \frac{p(S_t \mid \mathbf{h}_{0:T})p(S_{t+1} \mid S_t)}{p(S_{t+1} \mid \mathbf{h}_{0:t})} \end{aligned}$$

Since  $p(S_T \mid \mathbf{h}_{0:T})$  and  $p(S_t \mid \mathbf{h}_{0:t}, S_{t+1})$  are discrete,  $\mathbf{S}_{0:T}$  can be obtained by simulating  $T + 1$  random numbers from the uniform distributions over  $[0, 1]$  (Ripley, 1987).

## References

- Abanto-Valle, C. A., Lopes, H. F. & Migon, H. S. (2007). *Stochastic volatility estimation: a modern dynamic model viewpoint* (Technical Report 210). Federal University of Rio de Janeiro, Department of Statistics.
- Andersen, T. (1994). Stochastic autoregressive volatility: A framework for volatility modeling. *Mathematical Finance*, 4, 75–102.
- Andersen, T. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance*, 51, 169–204.

- Berg, A., Meyer, R. & Yu, J. (2004). Deviance Information Criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics*, *22*, 107–120.
- Carter, C. K. & Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, *81*, 541–553.
- Carvalho, C. & Lopes, H. F. (2007). Simulation-based sequential analysis of markov switching stochastic volatility models. *Computational Statistics & Data Analysis*, *51*, 4526–4542.
- Celeux, G., Forbes, F., Robert, C. P. & Titterton, D. M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis*, *1*, 651–674.
- Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, *41*, 135–155.
- de Jong, P. & Shepard, N. (1995). The simulation smoother for time series models. *Biometrika*, *82*, 339–350.
- Foster, F. & Wiswanathan, S. (1995). Can speculative trading explain the tradin-volume relation. *Journal of Business and Economic Statistics*, *13*, 379–396.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. David & A. F. M. Smith (Eds.), *Bayesian Statistics. Vol. 4.* (pp. 169–193).
- Glosten, L. R. & Milgrom, P. R. (1985). Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics*, *14*, 71–100.
- Hamilton, J. D. (1988). Rational expectations econometrics analysis of changes in regime. *Journal of Economic Dynamics and Control*, *12*, 385–423.

- Harris, L. (1987). Transaction data test of the mixture distribution hypothesis. *The Journal of Financial and Quantitative Analysis*, *22*, 127–141.
- Heidelberger, P. & Welch, P. (1983). Simulation run length control in the presence of a initial transient. *Operations Research*, *31*, 1109–1144.
- Jacquier, E., Polson, N. & Rossi, P. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, *12*, 371–418.
- Karpoff, J. M. (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.
- Mahieu, R. & Bauer, R. (1998). A Bayesian analysis of stock return volatility and trading volume. *Applied Financial Economics*, *8*, 671–687.
- Mahieu, R. & Schotman, P. C. (1998). An empirical application of stochastic volatility models. *Journal of Applied Econometrics*, *13*, 333–360.
- Omori, Y. & Watanabe, T. (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models. *Computational Statistics & Data Analysis*, *52*, 2892–2910.
- Pemstein, D., Quinn, K. V. & Martin, A. D. (2007). The Scythe statistical library: An open source C++ library for statistical computation. *Journal of Statistical Software*, *5*, 1–29.
- 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.

- Ripley, B. D. (1987). *Stochastic Simulation*. New York: Wiley.
- Shephard, N. & Pitt, M. (1997). Likelihood analysis of non-Gaussian measurements time series. *Biometrika*, *84*, 653–667.
- Shibata, M. & Watanabe, T. (2005). Bayesian analysis of a Markov switching stochastic volatility model. *Journal of the Japan Statistical Society*, *35*, 205–219.
- So, M., Lam, K. & Li, W. (1998). A stochastic volatility model with Markov Switching. *Journal of Business and Economic Statistics*, *15*, 183–202.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, *64*, 583–640.
- Tauchen, G. E. & Pitts, M. (1983). The price variability-volume relationships in speculative markets. *Econometrica*, *51*, 485–506.
- 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.