The Quarterly Review of Economics and Finance 80 (2021) 272–286

Contents lists available at ScienceDirect

The Quarterly Review of Economics and Finance

journal homepage: www.elsevier.com/locate/qref

Stochastic Volatility in Mean: Empirical evidence from Latin-American stock markets using Hamiltonian Monte Carlo and Riemann Manifold HMC methods

Carlos A. Abanto-Valle [a], Gabriel Rodríguez [b,*], Hernán B. Garrafa-Aragón [c]

[a] Department of Statistics, Federal University of Rio de Janeiro, Caixa Postal 68530, CEP: 21945-970 Rio de Janeiro, Brazil
[b] Department of Economics, Pontificia Universidad Católica del Perú, 1801 Universitaria Avenue, Lima 32, Lima, Peru
[c] Escuela de Ingeniería Estadística de la Universidad Nacional de Ingenieria, Lima, Peru

A R T I C L E I N F O

Article history:
Received 27 May 2020
Received in revised form
14 December 2020
Accepted 4 February 2021
Available online 13 February 2021

JEL classification:
C11
C15
C22
C51
C52
C58
G12

Keywords:
Feed-back effect
Hamiltonian Monte Carlo
Markov Chain Monte Carlo
Non linear state space models
Riemannian Manifold Hamiltonian Monte Carlo
Stochastic Volatility in Mean
Stock Latin American markets

A B S T R A C T

The Stochastic Volatility in Mean (SVM) model of Koopman and Uspensky (2002) is revisited. An empirical study of five Latin American indexes in order to see the impact of the volatility in the mean of the returns is performed. Markov Chain Monte Carlo (MCMC) Hamiltonian dynamics is used to estimate latent volatilities and parameters. Our findings show that volatility has a negative impact on returns, indicating that volatility feedback effect is stronger than the effect related to the expected volatility. This result is clear and opposite to the finding of Koopman and Uspensky (2002).

1. Introduction

Stochastic volatility (SV) models have been considered a useful tool for modeling time-varying variances, primarily in financial applications where policy-makers or stockholders are constantly faced with decision-making problems that typically rely on volatility and risk measures.[1] An attractive feature of the SV model is its close theoretical appealing to financial-economic theories as documented by Melino and Turnbull (1990) and also its flexibility to capture the main stylized facts often observed in daily series of financial returns in a more appropriate way; see Carnero, Peña, and Ruiz (2004). As documented by Black (1976), Campbell and Entchel (2000) and Bekaert and Wu (2000), the daily asymmetrical relationship between equity market returns and volatility has garnered significantly increased attention in both theoretical and applied financial literature. The asymmetric behaviour of market volatility is important for al least three reasons. First, it is an important characteristic of the dynamics of market volatility, has implications for asset pricing and is a feature of priced risk factors. Secondly, it plays

* Corresponding author.
E-mail address: gabriel.rodriguez@pucp.edu.pe (G. Rodríguez).

[1] The other important branch of related models is GARCH models where time-varying variance is modeled as a deterministic function of past squared perturbations and lagged conditional variances. Details and explanations of the extensive GARCH literature may be found in Bollerslev, Chou, and Kroner (1992), Bollerslev, Engle, and Nelson (1994) and Diebold and Lopes (1995). SV models are reviewed in Taylor (1994), Ghysels, Harvey, and Renault (1994) and Shephard (1996).

https://doi.org/10.1016/j.qref.2021.02.005
1062-9769/

a key role in predicting risk, hedging and pricing options. Finally, asymmetric volatility implies negatively skewed returns distributions, i.e. it may help to explain some of the probability value losses in the markets.

In recent years, there has been extensive analysis of the relationship between expected returns and expected volatility. Usually, there seems to be stronger evidence of a negative link between unexpected returns and innovations to the volatility process. French, Schert, and Stambugh (1987) interpreted this fact as an indirect evidence of a positive correlation between the expected risk premium and *ex ante* volatility. If expected volatility and expected returns are positively linked and future cash flows are not affected, the current stock index price should fall. On the other hand, small shocks to the return process cause an increase in contemporary stock index prices. This fact is known as the volatility feedback effect. An alternative cause for asymmetric volatility in which causality runs in the opposite direction is the leverage effect proposed by Black (1976), who argues that a negative (positive) return shock results in an increase (decrease) in the financial leverage ratio of the firm that has an upward (downward) impact on the volatility of its stock returns. However, French et al. (1987) and Schwert (1989) argued that the intensity of the negative relation cannot be taken into account by leverage alone. For example, Campbell and Entchel (2000) found evidence of both volatility feedback and leverage effects, while Bekaert and Wu (2000) presented findings indicating that the leverage effect is dominated empirically by the volatility feedback.

SV models have frequently been used to model volatility in daily stock returns. However, the results have depended on a considerable pre-processing of these series, avoiding the problem of simultaneous estimation of the mean and variance. To address this problem, Koopman and Uspensky (2002) introduced the SV in mean (SVM) model, allowing the non-observed volatility to enter into the mean return equation as an explanatory variable. Based on Monte Carlo simulation methods, they derived an exact maximum likelihood estimation procedure by assuming normality of the innovations. Recently, Abanto-Valle, Migon, and Lachos (2012) extended the SVM model to the class of scale mixture of Normal distributions, sampling parameters and log-volatilities by MCMC techniques.

Traditionally, MCMC algorithms are based on the Metropolis-Hastings (Hastings, 1973; Metropolis, Rosenbluth, Rosenbluth, Teller, & Teller, 1953) and the Gibbs sampling (Gelfand & Smith, 1990; Geman & Geman, 1984) algorithms. Due to the random walk nature of these algorithms, they often take an unacceptably long time to converge to the target distribution (Neal, 1993). On the other hand, Hamiltonian Monte Carlo (HMC) (Duane, Kennedy, Pendleton, & Roweth, 1987; Neal, 2011) avoids random walk behavior by introducing auxiliary variables to simulate Hamiltonian dynamics. This enables HMC to produce distant proposals with high acceptance probability, resulting in fast exploration of the target density. Girolami and Calderhead (2011) proposed a modification of the HMC method by exploiting the Riemannian geometry of the target distribution to improve convergence and mixing of the chain. By adapting the variance-covariance matrix used in the HMC, the new method called Riemannian Manifold HMC (RMHM) takes into consideration the local structure of the joint probability density. HMC methods has been applied in the context of SV models. See for example Takaishi (2009), Nugroho and Morimoto (2015, 2016) and Zevallos, Gasco, and Ehlers (2017) among others. Recently Kreuzer and Czado (2019) use HMC in a single factor copula SV model[2] .

[2] Other references that use the HMC and RMHMC tools are Martin and Maheu (2013), Martin and M. (2013) and ? although they are applied to the multivariate context.

The principal contribution of this article is the application of the HMC and RMHMC methods for updating log volatilities and parameters of the SVM model, respectively. These algorithms allow to update the log-volatilities at once, and parameters from the mean and volatility equations at once in two separate blocks, respectively.

Empirical applications of time-varying modeling of volatility using SVM models in developed countries have been studied by Koopman and Uspensky (2002) and Leão, Abanto-Valle, and Chen (2017). However, empirical evidence in Latin American markets is scarse; see Abanto-Valle, Migon, and Lachos (2011). For this reason, we contribute to the empirical literature performing a detailed empirical study of five Latin American indexes: MERVAL (Argentina), IBOVESPA (Brazil), IPSA (Chile), MEXBOL (Mexico) and IGBVL (Peru) in the context of the SVM model. We also include the S&P 500 returns in order to perform some comparisons. We find empirically, that the coeficient estimate, which quantifies both the ex ante relationship between returns and volatility and the volatility feedback effect, is negative and significant for all the indexes considered here except for the IGBVL. Likewise, we show evidence of the superiority of the SVM model when comparing it with the performance of three other models: GARCH(1,1), GARCH-M(1,1) and the basic SV models.

The remainder of this paper is organized as follows. Section 2 gives a brief introduction about HMC and RMHMC based methods in a general context. Section 3 outlines SVM model as well as the Bayesian estimation procedure using HMC and RMHMC methods. Section 4 is devoted to application using five Latin American indexes and the S&P 500. Finally, some concluding remarks and suggestions for future developments are given in Section 5. An Appendix outlines details of the HMC and RHMC algorithms.

2. MCMC using Hamiltonan dynamics

HMC (Duane et al., 1987; Neal, 2011) and RMHMC (Girolami & Calderhead, 2011) have become established as powerful, general purpose MCMC algorithms for sampling from general, continuous distributions. The efficiency of both methods is due to the fact that they make use of gradient information from the target density to allow for an ergodic Markov chain capable of large transitions that are accepted with high probability. In the following subsections, we briefly describe both methods.

2.1. Hamiltonian Monte Carlo

To sample a random variable $\boldsymbol{\theta} \in \mathbb{R}^p$ from the density $\pi(\boldsymbol{\theta})$, we introduce an independent auxiliary variable $\boldsymbol{\omega} \in \mathbb{R}^d$, such that $\boldsymbol{\omega} \sim \mathcal{N}_d(\mathbf{0}, \mathbf{M})$. The negative of the joint log-probability density function is given by

$$H(\boldsymbol{\theta}, \boldsymbol{\omega}) = -\mathcal{L}(\boldsymbol{\theta}) + \frac{1}{2}\log\{(2\pi)^p|\mathbf{M}|\} + \frac{1}{2}\boldsymbol{\omega}'\mathbf{M}^{-1}\boldsymbol{\omega}, \tag{1}$$

where $\mathcal{L}(\boldsymbol{\theta}) = \log \pi(\boldsymbol{\theta})$. Using a physical analogy, $\boldsymbol{\theta}$ may be interpreted as a particle position in the system, $-\mathcal{L}(\boldsymbol{\theta})$ defines its potential energy, whereas $\boldsymbol{\omega}$ is interpreted as the momentum with kinetic energy $\frac{1}{2}\boldsymbol{\omega}'\mathbf{M}^{-1}\boldsymbol{\omega}$ and the variance-covariance matrix $\mathbf{M}$ denotes a mass matrix. Finally, the total energy of a closed system is the scalar-valued Hamiltonian function $H(\boldsymbol{\theta}, \boldsymbol{\omega})$; see, for further details, Duane et al. (1987) and Leimkuhler and Reich (2004).

Further, $(\boldsymbol{\theta}(\tau), \boldsymbol{\omega}(\tau))$ provides a complete explanation of the physical system and its time-dynamics, with respect to a fictitious time τ, solves the equations of Hamilton:

$$\frac{\partial\boldsymbol{\theta}}{\partial\tau} = \frac{\partial H(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial\boldsymbol{\omega}} = \mathbf{M}^{-1}\boldsymbol{\omega}, \tag{2}$$

$$\frac{\partial\boldsymbol{\omega}}{\partial\tau} = -\frac{\partial H(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial\boldsymbol{\theta}} = \nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}), \tag{3}$$

where, $\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta})$ denotes the gradient of $\mathcal{L}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. The time dynamics associated with Eqs. (2) and (3) preserves total energy of the system and also the Boltzmann distribution $g(\boldsymbol{\theta}, \boldsymbol{\omega}) \propto \exp\{-H(\boldsymbol{\theta}, \boldsymbol{\omega})\}$. It means that if $(\boldsymbol{\theta}(0), \boldsymbol{\omega}(0)) \sim g(\boldsymbol{\theta}, \boldsymbol{\omega})$, then also $(\boldsymbol{\theta}(\tau), \boldsymbol{\omega}(\tau)) \sim g(\boldsymbol{\theta}, \boldsymbol{\omega})$ for all $\tau > 0$. In addition, since $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$ are independent under the Boltzmann distribution associated with Eq. (1), it is clear that the original target is $\boldsymbol{\theta}$-marginal of the Boltzmann distribution. For practical applications of interest, Hamilton's equations cannot be solved analytically, so numerical solutions primarily based on a discrete approximation to the real continuous solution are necessary. We use the Stormer-Verlet leapfrog integrator of Leimkuhler and Reich (2004) which is widely used because it retains the reversibility and volume preservation properties required to obtain an accurate sampler and calculates the updates through the following expressions:

$$\boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})} = \boldsymbol{\omega}^{(\tau)} + \frac{\epsilon}{2}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}),$$

$$\boldsymbol{\theta}^{(\tau+\epsilon)} = \boldsymbol{\theta}^{(\tau)} + \epsilon\mathbf{M}^{-1}\boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})},$$

$$\boldsymbol{\omega}^{(\tau+\epsilon)} = \boldsymbol{\theta}^{(\tau)} + \frac{\epsilon}{2}\nabla_{\boldsymbol{\theta}}\mathcal{L}(\boldsymbol{\theta}^{(\tau)}),$$

for some small step-size $\epsilon > 0$, defined by the user. Starting with the current state $(\boldsymbol{\theta}, \boldsymbol{\omega})$ and after a given number of time steps this results in a proposal $(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*)$. Since total energy is preserved only approximately with the Stormer-Verlet integrator, a corresponding bias is introduced into the joint density which can be corrected with an accept-reject step. The proposal state is accepted as the next state of the Markov chain according the Metropolis-Hastings acceptance probability given by:

$$\alpha(\boldsymbol{\theta}, \boldsymbol{\omega}; \boldsymbol{\theta}^*, \boldsymbol{\omega}^*) = \min\{1, \exp(-H(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*) + H(\boldsymbol{\theta}, \boldsymbol{\omega})\}. \tag{4}$$

2.2. Riemann Manifold Hamiltonian Monte Carlo

Girolami and Calderhead (2011) proposed a modification of the HMC method by exploiting the Riemannian geometry of the target distribution to improve convergence and mixing of the chain. By adapting the variance-covariance matrix $\mathbf{M}$ used in the HMC, the new method called Riemmanian Manifold HMC (RMHMC) takes into consideration the local structure of the joint probability density. The key idea is to reformulate the Hamiltonian function as

$$H(\boldsymbol{\theta}, \boldsymbol{\omega}) = -\mathcal{L}(\boldsymbol{\theta}) + \frac{1}{2}\log\{(2\pi)^p|\mathbf{M}(\boldsymbol{\theta})|\} + \frac{1}{2}\boldsymbol{\omega}'\mathbf{M}(\boldsymbol{\theta})^{-1}\boldsymbol{\omega},$$

where $M(\boldsymbol{\theta}) = -E(\frac{\partial^2\mathcal{L}()}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}'})$ is the expected Fisher information matrix plus the log prior negative Hessian. Hence, the Hamiltonian equations for the momentum and position variables, respectively are now defined by:

$$\frac{\partial\boldsymbol{\theta}}{\partial\tau} = \frac{\partial H(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial\boldsymbol{\omega}} = \mathbf{M}(\boldsymbol{\theta})^{-1}\boldsymbol{\omega},$$

$$\frac{\partial\boldsymbol{\omega}_i}{\partial\tau} = -\frac{\partial H(\boldsymbol{\theta}, \boldsymbol{\omega})}{\partial\theta_i} = \nabla_{\theta_i}\mathcal{L}(\boldsymbol{\theta}) - \frac{1}{2}\mathrm{tr}\left[\mathbf{M}(\boldsymbol{\theta})^{-1}\frac{\partial\mathbf{M}(\boldsymbol{\theta})}{\partial\theta_i}\right] + \frac{1}{2}\boldsymbol{\omega}'\mathbf{M}(\boldsymbol{\theta})^{-1}\frac{\partial\mathbf{M}(\boldsymbol{\theta})}{\partial\theta_i}\mathbf{M}(\boldsymbol{\theta})^{-1}\boldsymbol{\omega}.$$

Following Leimkuhler and Reich (2004), we adopt the generalized Stormer-Verlet solution to simulate values in a discrete time. The procedure is described as follows:

$$\boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})} = \boldsymbol{\omega}^{(\tau)} - \frac{\epsilon}{2}\nabla_{\boldsymbol{\theta}}H(\boldsymbol{\theta}^{(\tau)}, \boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})}),$$

$$\boldsymbol{\theta}^{(\tau+\epsilon)} = \boldsymbol{\theta}^{(\tau)} + \frac{\epsilon}{2}\left[\nabla_{\boldsymbol{\omega}}H(\boldsymbol{\theta}^{(\tau)}, \boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})}) + \nabla_{\boldsymbol{\omega}}H(\boldsymbol{\theta}^{(\tau+\epsilon)}, \boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})})\right],$$

$$\boldsymbol{\omega}^{(\tau+\epsilon)} = \boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})} - \frac{\epsilon}{2}\nabla_{\boldsymbol{\theta}}H(\boldsymbol{\theta}^{(\tau+\epsilon)}, \boldsymbol{\omega}^{(\tau+\frac{\epsilon}{2})}).$$

Repeated application of these steps at each iteration provides a new proposal $(\boldsymbol{\theta}^*, \boldsymbol{\omega}^*)$ that is accepted with probability $\alpha(\boldsymbol{\theta}, \boldsymbol{\omega}; \boldsymbol{\theta}^*, \boldsymbol{\omega}^*)$ given by Eq. (4).

3. The Stochastic Volatility in Mean model

The SVM model is defined by

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 e^{h_t} + e^{\frac{h_t}{2}}\epsilon_t, \tag{5a}$$

$$h_{t+1} = \mu + \phi(h_t - \mu) + \sigma\eta_t, \tag{5b}$$

where y_t and h_t are, respectively, the compounded return and the log-volatility at time t. We assume that $|\phi| < 1$, i.e., that the log-volatility process is stationary and that the initial value $h_1 \sim \mathcal{N}(\mu, \frac{\sigma_\eta^2}{1-\phi^2})$. The innovations ϵ_t and η_t are assumed to be mutually independent and Normally distributed with mean zero and unit variance. The SVM model incorporates the unobserved volatility as an explanatory variable in the mean equation. Therefore the SVM model aims at estimating the ex-ante relationship between returns and volatility and the volatility feedback effect (parameter β_2). Under the Bayesian viewpoint, MCMC methods used to conduct the posterior analysis are described in the next subsection.

3.1. Parameter estimation via MCMC

Let $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \mu, \phi, \sigma^2)'$ denotes the full parameter vector of the SVM model, $\mathbf{h}_{1:T} = (h_1, \ldots, h_T)'$ denotes the entire vector of the log volatilities and $\mathbf{y}_{0:T} = (y_0, \ldots, y_T)'$ is the information available up to time T. Under the Bayesian approach and following Tanner and Wong (1987), we use the data augmentation principle, which considers $\mathbf{h}_{1:T}$ as latent variables. The joint posterior density of parameters and latent variables is given by

$$\begin{aligned} p(\mathbf{h}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{0:T}) &\propto p(\mathbf{y}_{1:T} \mid y_0, \boldsymbol{\theta}, \boldsymbol{\lambda}_{1:T}, \mathbf{h}_{0:T})p(\mathbf{h}_{1:T} \mid \boldsymbol{\theta})p(\boldsymbol{\theta}) \\ &= \prod_{t=1}^{T}\left[p(y_t \mid y_{t-1}, \beta_0, \beta_1, \beta_2, h_t)\right] \\ &\quad p(h_1 \mid \mu, \phi, \sigma^2)\prod_{t=1}^{T-1}\left[p(h_{t+1} \mid h_t, \mu, \phi, \sigma^2)\right]p(\boldsymbol{\theta}), \end{aligned} \tag{6}$$

where $p(y_t \mid y_{t-1}, \beta_0, \beta_1, \beta_2, h_t)$ and $p(h_{t+1} \mid h_t, \mu, \phi, \sigma^2)$ are defined for Eqs. (5a) and (5b), and $p(\boldsymbol{\theta})$ denotes the prior distribution. We assume that $p(\boldsymbol{\theta})$ is prior independent. The priors distributions of parameters in the SVM model are set as: $\beta_0 \sim \mathcal{N}(\bar{\beta}_0, \sigma^2_{\beta_0})$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(a_{\beta_1}, b_{\beta_1})$, $\beta_2 \sim \mathcal{N}(\bar{\beta}_2, \sigma^2_{\beta_2})$, $\mu \sim \mathcal{N}(\mu_0, \sigma^2_\mu)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(a_\phi, b_\phi)$ and $\sigma^2 \sim \mathcal{IG}(a_\sigma, b_\sigma)$, where $\mathcal{N}(.,.)$, $\mathcal{B}e(.,.)$ and $\mathcal{IG}(.,.)$ denote the Normal, Beta and Inverse Gamma distributions, respectively. This choice of priors ensures that all parameters have the right support; in particular the Beta prior on β_1 and ϕ ensures that $-1 < \beta_1, \phi < 1$.

Table 1
Summary statistics for daily stock returns data.

INDEX	MERVAL	IBOVESPA	IPSA	MEXBOL	IGBVL	S&P500
Size	4651	4698	4737	4759	4597	4777
Mean	0.0701	0.0376	0.0296	0.0464	0.0478	0.0177
S. D.	2.2125	2.0262	1.0695	1.4276	1.4111	1.2418
Minimum	−14.2896	−17.2082	−7.6381	−10.3410	−13.2908	−9.4695
Maximum	16.1165	28.8325	11.8034	12.1536	12.8156	10.9572
Skewness	2.2091	0.5313	0.1372	0.1458	−0.3915	−0.2086
Kurtosis	7.3418	16.8094	11.6866	8.7449	13.5715	10.6576
Returns						
$\hat{\rho}_1$	0.0550	0.0130	0.1840	0.0910	0.1890	−0.0700
$\hat{\rho}_2$	0.0020	−0.0180	0.0220	−0.0300	0.0080	−0.0450
$\hat{\rho}_3$	0.0240	−0.0390	−0.0190	−0.0301	0.0680	0.0100
$\hat{\rho}_4$	0.0070	−0.0320	0.0250	−0.0030	0.0640	−0.0080
$\hat{\rho}_5$	−0.0090	−0.0170	0.0270	−0.0150	0.0250	−0.0460
Q(12)	33.37	44.51	190.52	54.57	240.53	66.95
Squared Returns						
$\hat{\rho}_1$	0.2580	0.1990	0.2320	0.1430	0.4210	0.2040
$\hat{\rho}_2$	0.2160	0.1640	0.2130	0.1780	0.3890	0.3720
$\hat{\rho}_3$	0.1780	0.1860	0.1720	0.2540	0.3920	0.1920
$\hat{\rho}_4$	0.1660	0.1170	0.1550	0.1300	0.2840	0.2880
$\hat{\rho}_5$	0.2130	0.0990	0.2910	0.2420	0.2140	0.3220
Q(12)	1763.40	1069.30	2086.00	2147.10	3960.20	4643.70

Since $p(\mathbf{h}_{1:T}, \boldsymbol{\theta} \mid \mathbf{y}_{0:T})$ is not available in a closed form, we sample from using the Gibbs sampling as described in Algorithm 1.

Algorithm 1.

(1) Set $i = 0$ and get starting values for the parameters $\boldsymbol{\theta}^{(i)}$ and the latent quantities $\mathbf{h}_{1:T}^{(i)}$;
(2) Generate $(\mu, \phi, \sigma)^{(i+1)} \sim p(\mu, \phi, \sigma \mid \mathbf{y}_{1:T}, \mathbf{h}_{1:T}^{(i)})$;
(3) Draw $(\beta_0, \beta_1, \beta_2)^{(i+1)} \sim p(\beta_0, \beta_1, \beta_2 \mid \mathbf{h}_{1:T}^{(i)}, \mathbf{y}_{0:T})$;
(4) Generate $\mathbf{h}_{1:T}^{(i+1)} \sim p(h_{1:T} \mid (\beta_0, \beta_1, \beta_2)^{(i+1)}, (\mu, \phi, \sigma)^{(i+1)})$;
(5) Set $i = i + 1$ and return to step (2) until convergence is achieved,

A critical issue of MCMC methods is to determine how long the simulation needs to be run in order to estimate the characteristics of the distribution of interest. As documented by Cowles and Carlin (1996), assessing convergence from a theoretical point of view is not a trivial task and relatively little of practical use in applied work. Consequently, most MCMC users address the convergence problem by applying diagnostic tools to the output produced by running their sampler. In this article, we use the convergence diagnostic of Geweke (1992).

4. Empirical application

We consider the daily closing prices of five Latin American stock market: MERVAL (Argentina), IBOVESPA (Brazil), IPSA (Chile), MEXBOL (Mexico) and IGBVL (Peru). We use the S&P 500 in order to compare the results with Latin American stock markets in the sense that the U.S. stock market could be considered as a good benchmark. The data sets were obtained from the Yahoo finance web site available to download at http://finance.yahoo.com. The period of analysis is from January 6, 1998, until December 30, 2016. We consider the compounded return computed as a percentage, $y_t = 100 \times (\log P_t - \log P_{t-1})$, where the P_t denotes the adjusted closing price on day t. Table 1 shows summary descriptive statistics. The sample sizes differ between countries because of holidays and non-trading days on the stock exchange markets. According to Table 1, the IGBVL and S&P 500 are negatively skewed whereas the rest are positively skewed. The IGBVL returns are the most negatively skewed with −0.3915 and the IBOVESPA returns are the most positively skewed with 0.5313. Regarding the kurtosis, all the daily returns of the five Latin American returns and the S&P 500 are leptokurtic, since all estimates of kurtosis exceed 3. Brazil, Peru and Chile are the markets with the highest degree of kurtosis with the USA near the value observed for Chile. Although there are high differences between the minimum and maximum values, the most outstanding values are those corresponding to Argentina and Brazil. These two countries show to be the most volatile too, which can be attributed to extreme minimum and maximum values.

We further observe that the IGBVL and IPSA returns show the highest level of first-order autocorrelation. These values decrease fast for the other orders of autocorrelation. In the case of returns, high first-order autocorrelation reflects the effects of non-synchronous or thin trading. The squared returns show high level of autocorrelation of order 1 which can be seen as an indication of volatility clustering. We further observe that high-order autocorrelations for squared returns are still high and decrease slowly[3] . The $Q(12)$ test statistic, which is a joint test for the hypothesis that the first twelve autocorrelation coefficients are equal to zero, indicates that this hypothesis has to be rejected at the 5% significance level for all returns and squared returns series.

We simulate the h_t's in one block using HMC method which was implemented by using 50 leapfrog steps and a step size of 0.015. The parameters $(\mu, \phi, \sigma)'$ and $(\beta_0, \beta_1, \beta_2)'$ are sampled in blocks using the RMHMC method. In the case of $(\mu, \phi, \sigma)'$, we use a step size of 0.5, 20 leafprog steps and the number of fixed point iterations is 5. For $(\beta_0, \beta_1, \beta_2)'$ we use a step size of 0.1, 20 leapfrog steps and 5 fixed point iterations. We set the prior distributions of the common parameters as: $\beta_0 \sim \mathcal{N}(0, 10)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(5, 1.5)$, $\beta_2 \sim \mathcal{N}(0, 10)$, $\mu \sim \mathcal{N}(0.0, 10)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(20, 1.5)$, and $\sigma^2 \sim \mathcal{IG}(2.5, 0.025)$. These values imply that the prior mean and standard deviation of β_1 and ϕ are (0.5385,0.3077) and (0.8605, 0.1074), respectively. For σ^2, the parameter setting implies a prior mean and prior standard deviation are (0.0167, 0.0236).

We conducted the MCMC simulation for 30,000 iterations. Table 2 reports the elapsed time to perform the 30,000 iterations. The first 10,000 draws are discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the

[3] This behavior has suggested that the literature considers that there is a long memory in the volatility of returns, as well as the possibility that infrequent level shifts cause such behavior. For a discussion on this, see Diebold and Inoue (2001) and Perron and Qu (2010), among others. For applications to different financial markets in Latin America, see Rodríguez (2017) and the references mentioned therein.

Table 2
Elapsed time in seconds to perform 30,000 iterations of the SVM for the market indexes.

MERVAL	IBOVESP	IPSA	MEXBOL	IGBVL	S&P 500
2848.34	2898.50	2873.25	2897.57	2874.33	2900.42

simulated chain, only every 10*th* values of the chain are stored. With the resulting 2000 values, we calculate the posterior means, the 95% credible intervals, the Monte Carlo error of the posterior means, the inefficiency factors (Inef) and the convergence diagnostic (CD) statistics as proposed by Geweke (1992). According to the CD values, the null hypothesis that the sequence of 2000 draws is stationary is accepted at 5% level for all the parameters and series considered here. All the calculations were performed running stand-alone code developed by the authors, by using open source C++ Armadillo library (Sanderson & Curtin, 2016) and the GNU Scientific Library (GSL, Gough, 2009). The inefficiency factor is defined as $1+2\sum_{s=1}^{\infty}\rho_s$ where ρ_s sis the sample auto-correlation at lag s, and are computed to measure how well the MCMC chain mixes (see e.g. Chib, 2001). It is the ratio of the numerical variance of the posterior sample mean to the variance of the sample mean from uncorrelated draws. The inverse of inefficiency factor is also known as relative numerical efficiency (Geweke, 1992). When the inefficiency factor is equal to m, we need to draw MCMC samples m times as many as uncorrelated samples.

Table 3 summarizes these results for the five Latin American stock market and the S&P 500 returns. The value of ϕ is very similar among all markets, suggesting similar degrees of persistence (ranging from 0.9501 for Argentina to 0.9803 for Mexico). The MEXBOL, IBOVESPA and S&P 500 are more persistent. In fact, for the two former volatilities, the half-lives of the shocks have a duration of 34.9 and 25.3 days, respectively. In the cases of the IPSA, IGBVL and MERVAL, the durations are 22.5, 17.8 and 13.5 days, respectively. For the S&P 500, the half-life duration of a shock is around 32.8 days, which is very close to the result of the MEXBOL.

The posterior mean estimates of σ show that all returns have similar estimates in the range from 0.1649 to 0.2725. The higest value is 0.2725 for the IGBVL which jointly with the estimate of ϕ indicates that IGBVL is the most volatile stock market index in the region. Regarding the posterior mean of μ, we find that the estimates are statistically significant for the MERVAL, IBOVESPA and IPSA indexes. For the MEXBOL, IGBVL and S&P 500 indexes, the parameter μ could be not significant because the credibility interval contains zero.

We observe that the posterior mean parameter β_0 is always positive and statistically significant for all series. The value of β_1 that measures the correlation of returns is as expected, small and very similar to the first-order autocorrelation coefficients reported in Table 1. The estimates of β_1 are statistically significant for all series with the exception of Brazil and, although in the cases of Chile and Peru these values are 0.188 and 0.1875, respectively, these values indicate a weak persistence with a rapid mean reversion. Regarding the parameter of interest (β_2), this is more negative in the cases of USA, Brazil and Chile. Intermediate values are observed in Argentina and Mexico while Peru presents the smallest value in absolute terms. Moreover, while all countries have a credibility interval that excludes the zero value, this does not happen in the case of Peru, so it is difficult to argue for an uncertainty effect in this market. It is important to note that the right side of the credibility interval is very close to zero in all markets except the U.S. Therefore, the posterior mean of β_2 parameter, which measures both the ex ante relationship between returns and volatility and the volatility feedback effect, is negative for all series and statistically significant for all the series with the exception of Peru. Following Koopman and Uspensky (2002), the volatility feedback effect (negative) dominates the positive effect which links the returns with the expected volatility. Our estimates are more negative compared to those of Koopman and Uspensky (2002) where the hypothesis that $\beta_2 = 0$ can never be rejected at the conventional 5% significance level. Therefore, the volatility feedback effect is clearly dominant in our results (except for Peru) in comparison to those of Koopman and Uspensky (2002). These results confirm the hypothesis that investors require higher expected returns when unanticipated increases in future volatility are highly persistent. This is consistent with our findings of higher values of ϕ combined with larger negative values for the in-mean parameter. We have indirect evidence of a positive intertemporal relation between expected excess market returns and its volatility as this is one of the assumptions underlying the volatility feedback hypothesis.

Figs. 1–6 show the MCMC output, the autocorrelation function (ACF) and the posterior densities of the parameters of the SVM model for all the indexes considered here. In all the cases, all the parameters showed good mixing properties with the exception of

Table 3
MCMC estimation of the SVM model. We report the posterior mean, the Monte Carlo (MC) error, the 95% credibility interval, the ineficiency factor (Inef) and the convergence diagnostic (CD), respectively.

Parameter	Mean	MC Error	95% interval	Inef	CD	Parameter	Mean	MC Error	95% interval	Inef	CD
MERVAL (Argentina)						IBOVESPA (Brazil)					
μ	1.1647	0.0020	(0.9931,1.3291)	1.16	−1.79	μ	1.2898	0.0040	(0.9807,1.6022)	1.36	1.17
ϕ	0.9501	0.0007	(0.9353,0.9629)	19.32	−0.98	ϕ	0.9730	0.0007	(0.9565,0.9863)	14.83	0.26
σ	0.2688	0.0025	(0.2399,0.3044)	44.67	1.17	σ	0.1649	0.0025	(0.1325,0.1968)	44.23	−0.64
β_0	0.2052	0.0008	(0.1291,0.2815)	1.00	0.81	β_0	0.2575	0.0018	(0.1101,0.4091)	1.10	−0.07
β_1	0.0478	0.0003	(0.0157,0.0783)	1.00	−1.26	β_1	0.0309	0.0005	(−0.0161,0.0743)	1.10	−0.07
β_2	−0.0287	0.0025	(−0.0510, −0.0073)	1.00	−0.50	β_2	−0.0259	0.0004	(−0.0510, −0.0018)	1.00	−0.04
IPSA (Chile)						MEXBOL (Mexico)					
μ	−0.3596	0.0025	(−0.5706, −0.1555)	1.00	−0.33	μ	0.2316	0.0033	(−0.0634,0.5200)	1.00	−0.33
ϕ	0.9697	0.0004	(0.9599,0.9791)	15.91	0.06	ϕ	0.9803	0.0003	(0.9725,0.9870)	11.24	−0.56
σ	0.2111	0.0021	(0.1880,0.2385)	55.41	−0.11	σ	0.1859	0.0016	(0.1655,0.2102)	42.70	0.05
β_0	0.0710	0.0004	(0.0385,0.1057)	1.00	0.88	β_0	0.1039	0.0004	(0.0655,0.2102)	1.00	−0.16
β_1	0.1885	0.0003	(0.1578,0.2184)	1.00	−0.79	β_1	0.0742	0.0004	(0.0438,0.1051)	1.00	1.50
β_2	−0.0442	0.0003	(−0.0868, −0.0021)	1.00	−0.90	β_2	−0.0301	0.0003	(−0.0581, −0.0020)	1.00	−0.41
IGBVL (Peru)						S&P 500 (USA)					
μ	−0.0095	0.0025	(−0.2342,0.2044)	1.00	−1.08	μ	0.1332	0.–0032	(−0.4096,0.1375)	1.07	0.01
ϕ	0.9618	0.0006	(0.9490,0.9732)	19.82	−0.98	ϕ	0.9791	0.0005	(0.9705,0.9864)	28.76	0.91
σ	0.2725	0.0026	(0.2351,0.3102)	37.22	0.86	σ	0.1968	0.0035	(0.1686,0.2311)	100.00	−1.08
β_0	0.0596	0.0041	(0.0232,0.0960)	1.00	−0.01	β_0	0.1085	0.0004	(0.0754,0.1404)	1.00	−1.05
β_1	0.1875	0.0004	(0.1563,0.2199)	1.00	−1.21	β_1	−0.0504	0.0003	(−0.0853,−0.0232)	1.00	−0.89
β_2	−0.0114	0.0004	(−0.0414,0.0179)	1.11	0.86	β_2	−0.0595	0.0004	(−0.0924,−0.0272)	1.07	0.68

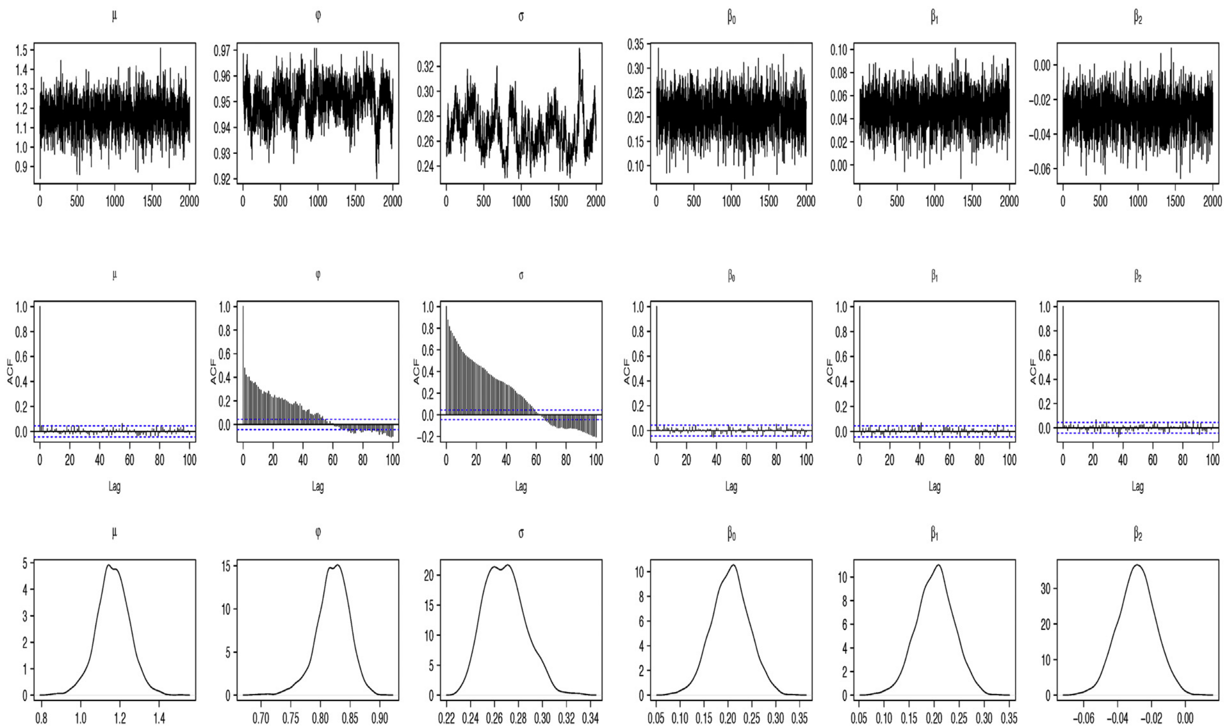

Fig. 1. MCMC estimation results of the SVM model for MERVAL (Argentina). Sample paths (top), autocorrelations (middle) and posterior densities (bottom).

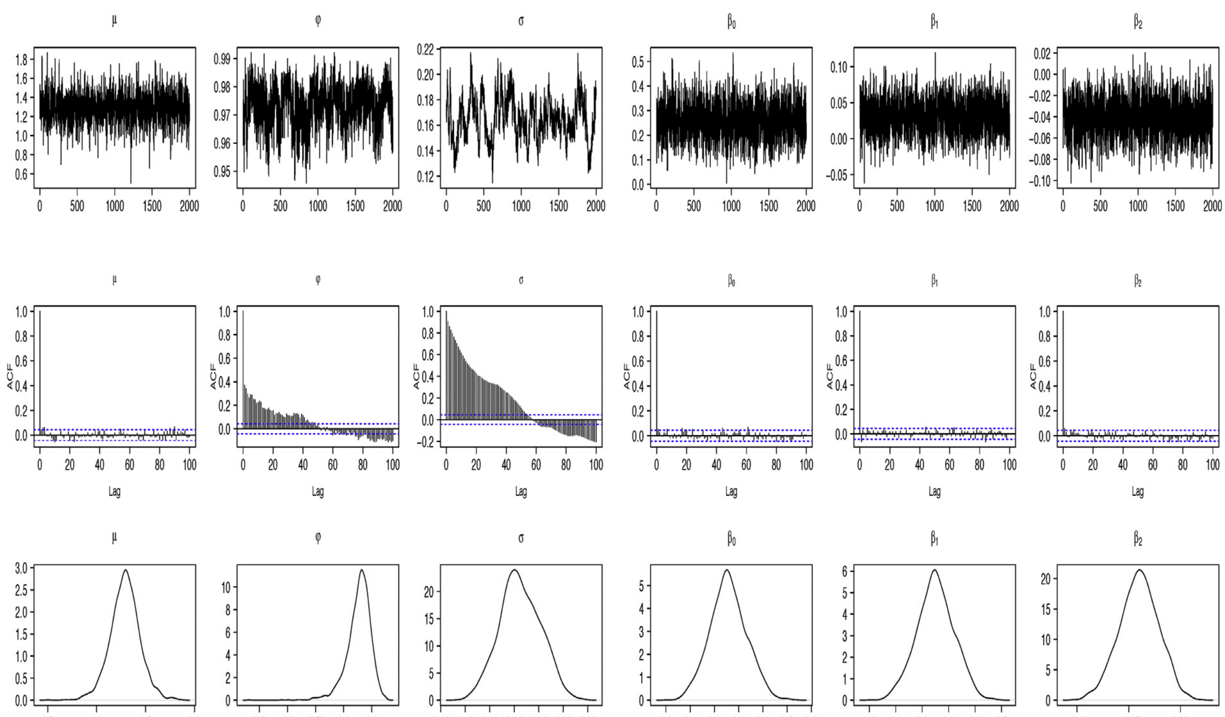

Fig. 2. MCMC estimation results of the SVM model for BOVESPA (Brazil). Sample paths (top), autocorrelations (middle) and posterior densities (bottom).

ϕ and σ. Regarding the inefficiency factors, showed in Table 3, these are higher in the ϕ and σ parameters. The USA shows the highest levels of inefficiency in the estimation of σ while in the case of Peru this is the lowest. A similar behavior is observed in the case of the ϕ parameter.

Now, we consider the correlation between all the parameters for all indexes obtained from the MCMC output. From Table 4, we found small correlations for almost all the parameters with exception of ϕ and σ, and β_2 and β_0. In both cases, we find negative correlations. This fact indicates that if ϕ (β_2) increases σ (β_0) decreases or vice versa. For ϕ and σ, we found the most negative correlation for the IGBVL, followed by MERVAL and S&P 500. For β_2 and β_0, the most negative value is found in the IBOVESPA returns, followed by MERVAL and IPSA returns, respectively.

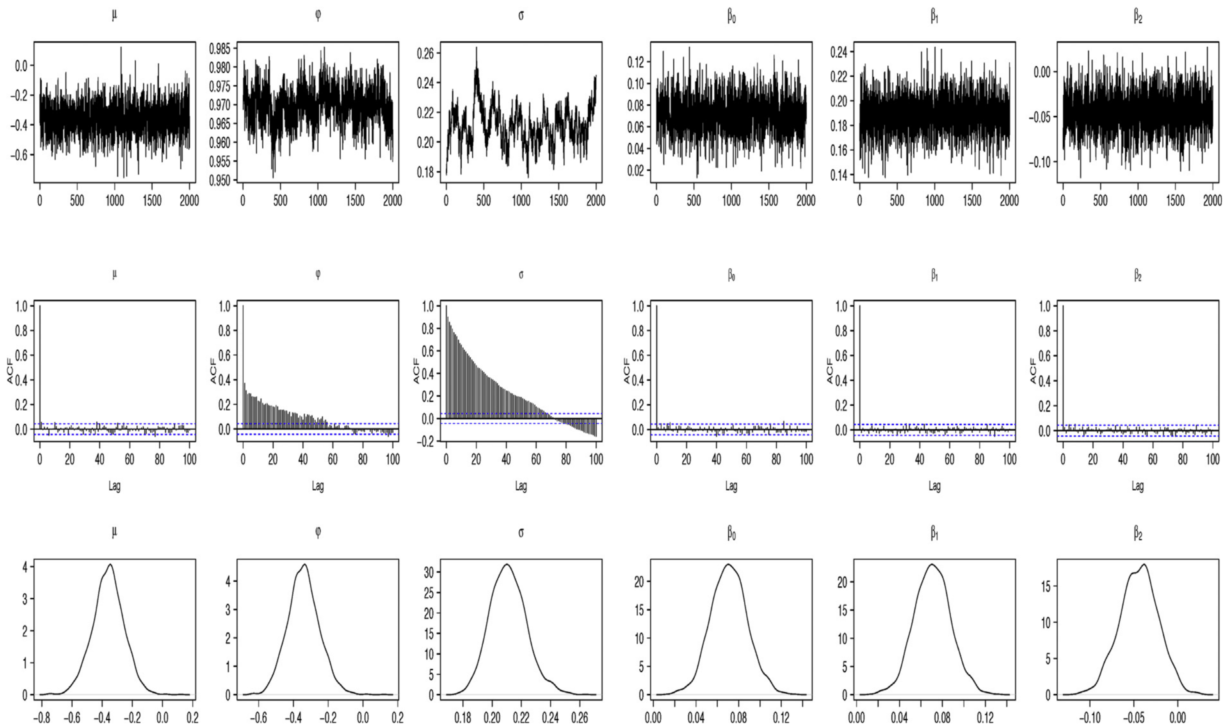

Fig. 3. MCMC estimation results of the SVM model for IPSA (Chile). Sample paths (top), autocorrelations (middle) and posterior densities (bottom).

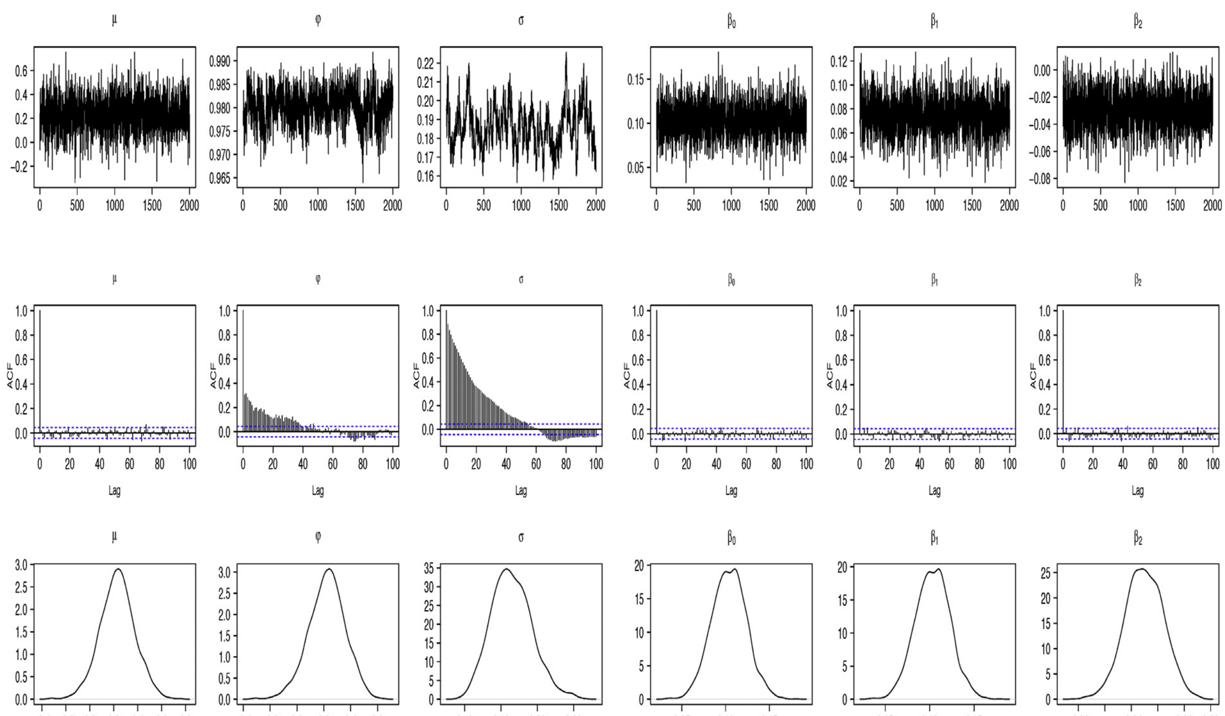

Fig. 4. MCMC estimation results of the SVM model for MEXBOL (Mexico). Sample paths (top), autocorrelations (middle) and posterior densities (bottom).

Fig. 7 show the smoothed mean of $e^{\frac{h_t}{2}}$ for all the indexes considered here. The smoothed mean is obtained as $\frac{1}{R}\sum_{i=1}^{R} e^{\frac{h_t^{(i)}}{2}}$, where $h_t^{(i)}$ is the value of h_t for the ith MCMC iteration and R is the number of iterations. Fig. 7 establishes a visual comparison of the evolutions of volatility and the absolute value of the returns. In both cases, the series show similar patterns at times of low, medium and high uncertainty, reflected in greater values for the estimated SVM model. It can be appreciated that the volatility of the Latin American stock market returns has been affected by the Russian and Brazilian financial crises in August 1998 and January 1999, reflected in higher volatility. Likewise, the financial crisis in the USA (2008) has had serious repercussions on the behavior of the volatility of the Latin American stock markets. Moreover, it can also be appreciated that in 2010 there were no major shocks in these markets. However, in 2011 the crisis was accentuated in the countries of the European

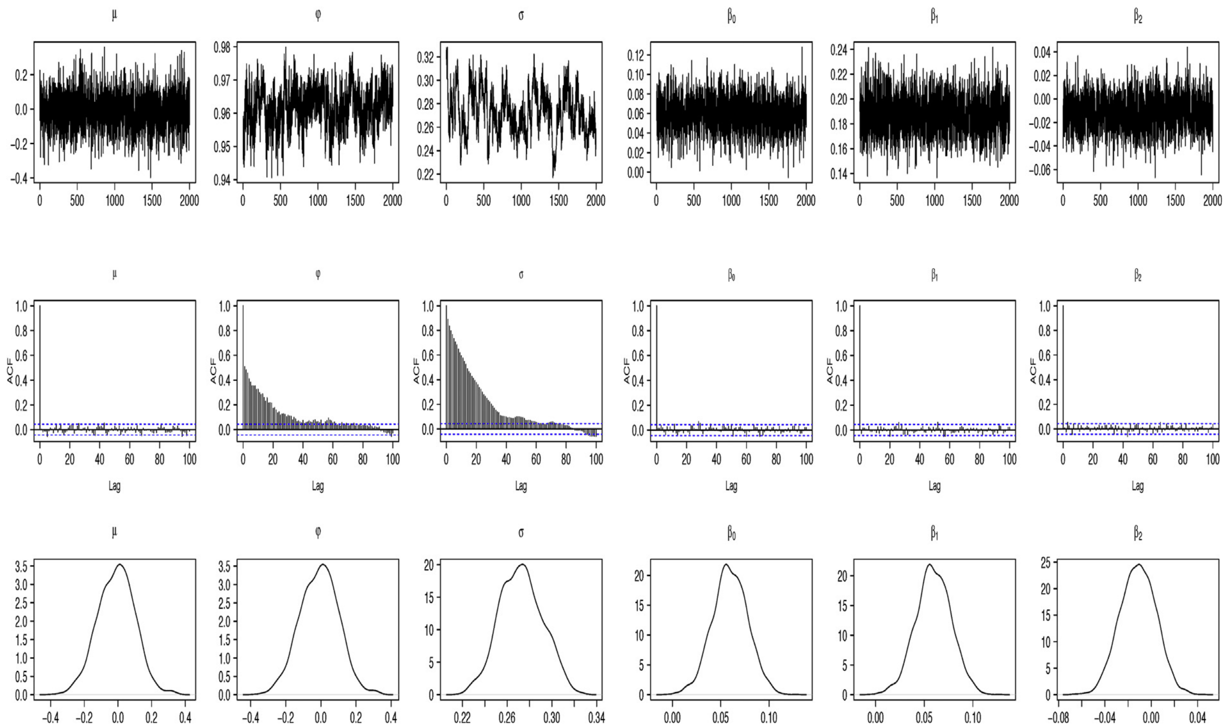

Fig. 5. MCMC estimation results of the SVM model for IGBVL(Peru). Sample paths (top), autocorrelations (middle) and posterior densities (bottom).

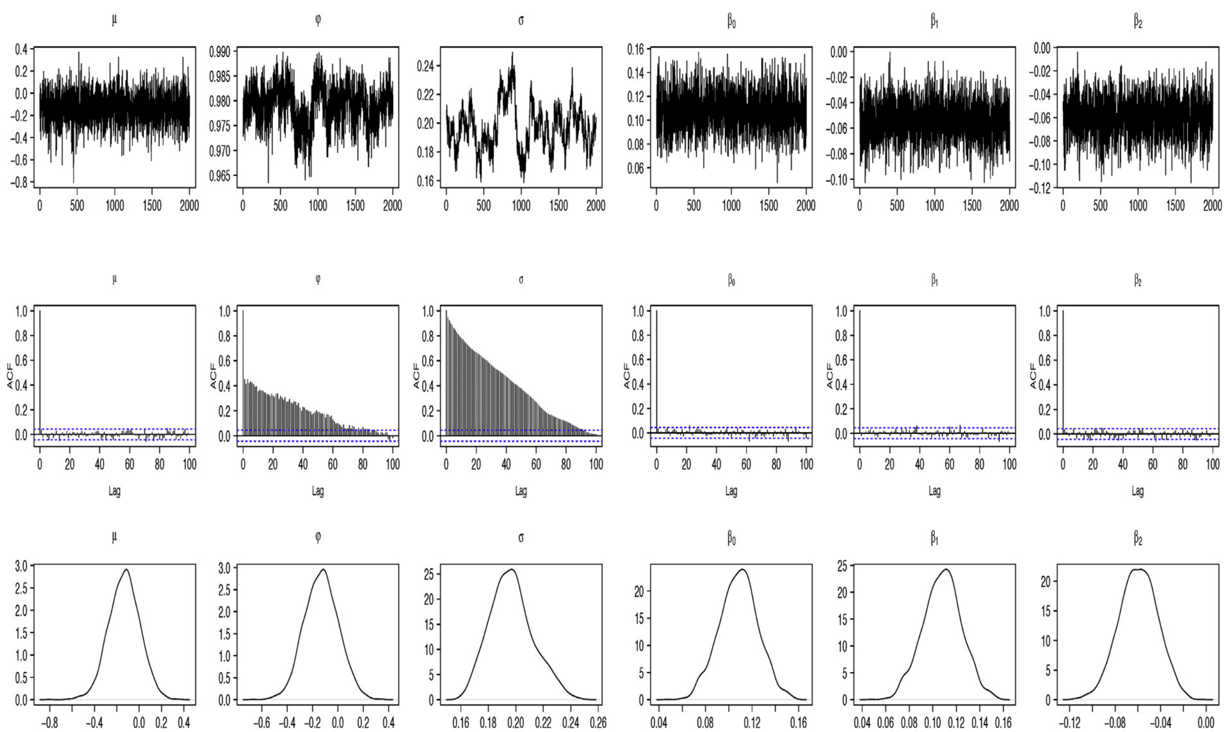

Fig. 6. MCMC estimation results of the SVM model for S&P 500 (USA). Sample paths (top), autocorrelations (middle) and posterior densities (bottom).

Union and North America, causing uncertainty in the markets of emerging economies such as Peru.

Next, we perform a sensitivity analysis of the prior distributions. We set the following set of priors: Prior # 1: $\mu \sim \mathcal{N}(0, 10)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(20, 1.5)$, $\sigma^2 \sim \mathcal{IG}(5, 0.025)$, $\beta_0 \sim \mathcal{N}(0, 10)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(5, 1.5)$ and $\beta_2 \sim \mathcal{N}(0, 10)$; Prior # 2: $\mu \sim \mathcal{N}(0, 25)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(22, 2.5)$, $\sigma^2 \sim \mathcal{IG}(2.5, 0.0125)$, $\beta_0 \sim \mathcal{N}(0, 25)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(4, 1.0)$ and $\beta_2 \sim \mathcal{N}(0, 25)$; Prior # 3: $\mu \sim \mathcal{N}(0, 50)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(22, 2.5)$, $\sigma^2 \sim \mathcal{IG}(5, 0.025)$, $\beta_0 \sim \mathcal{N}(0, 50)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(4, 1)$ and $\beta_2 \sim \mathcal{N}(0, 50)$; Prior # 4: $\mu \sim \mathcal{N}(0, 100)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(20, 1.5)$, $\sigma^2 \sim \mathcal{IG}(5, 0.025)$, $\beta_0 \sim \mathcal{N}(0, 100)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(5, 1.5)$ and $\beta_2 \sim \mathcal{N}(0, 100)$; Prior # 5: $\mu \sim \mathcal{N}(-0.5, 100)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(20, 1.5)$, $\sigma^2 \sim \mathcal{IG}(5, 0.025)$, $\beta_0 \sim \mathcal{N}(0.5, 100)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(5, 1.5)$ and $\beta_2 \sim \mathcal{N}(-0.5, 100)$; Prior # 6:

Table 4
Correlation matrix of posterior samples of the SVM model.

	μ	ϕ	σ	β_0	β_1	β_2		μ	ϕ	σ	β_0	β_1	β_2
MERVAL (Argentina)							IBOVESPA (Brazil)						
μ	1.0000						μ	1.0000					
ϕ	0.0438	1.0000					ϕ	−0.0153	1.0000				
σ	−0.0842	−0.6941	1.0000				σ	−0.0155	−0.5954	1.0000			
β_0	0.0229	−0.0517	0.0022	1.0000			β_0	−0.0434	−0.0452	0.0580	1.0000		
β_1	0.0247	−0.0331	0.0190	−0.0822	1.0000		β_1	0.0086	−0.0390	0.0015	−0.1493	1.0000	
β_2	−0.0007	0.0442	0.0121	−0.7504	0.0169	1.0000	β_2	0.0329	0.0340	−0.0235	−0.8114	0.0829	1.0000
IPSA (Chile)							MEXBOL (Mexico)						
μ	1.0000						μ	1.0000					
ϕ	−0.0079	1.0000					ϕ	0.0216	1.0000				
σ	−0.0390	−0.5806	1.0000				σ	−0.0290	−0.5456	1.0000			
β_0	−0.0086	−0.0556	0.0525	1.0000			β_0	0.0313	−0.0417	0.0588	1.0000		
β_1	0.0079	0.0381	−0.0525	−0.1391	1.0000		β_1	−0.0399	−0.0070	−0.0171	−0.0761	1.0000	
β_2	0.0001	0.0426	−0.0409	−0.7417	0.1021	1.0000	β_2	0.0151	0.0469	−0.0366	−0.6675	−0.0047	1.0000
IGBVL (Peru)							S&P 500 (USA)						
μ	1.0000						μ	1.0000					
ϕ	0.0105	1.0000					ϕ	0.0266	1.0000				
σ	−0.0495	−0.7258	1.0000				σ	−0.0619	−0.6820	1.0000			
β_0	0.0034	0.0593	−0.0603	1.0000			β_0	−0.0140	−0.0789	0.0826	1.0000		
β_1	−0.0052	0.0185	−0.0305	−0.0674	1.0000		β_1	0.0007	0.0364	−0.0599	−0.1315	1.0000	
β_2	0.0192	−0.0641	−0.0409	−0.6431	−0.0087	1.0000	β_2	0.0350	0.0273	0.0183	−0.6614	0.0764	1.0000

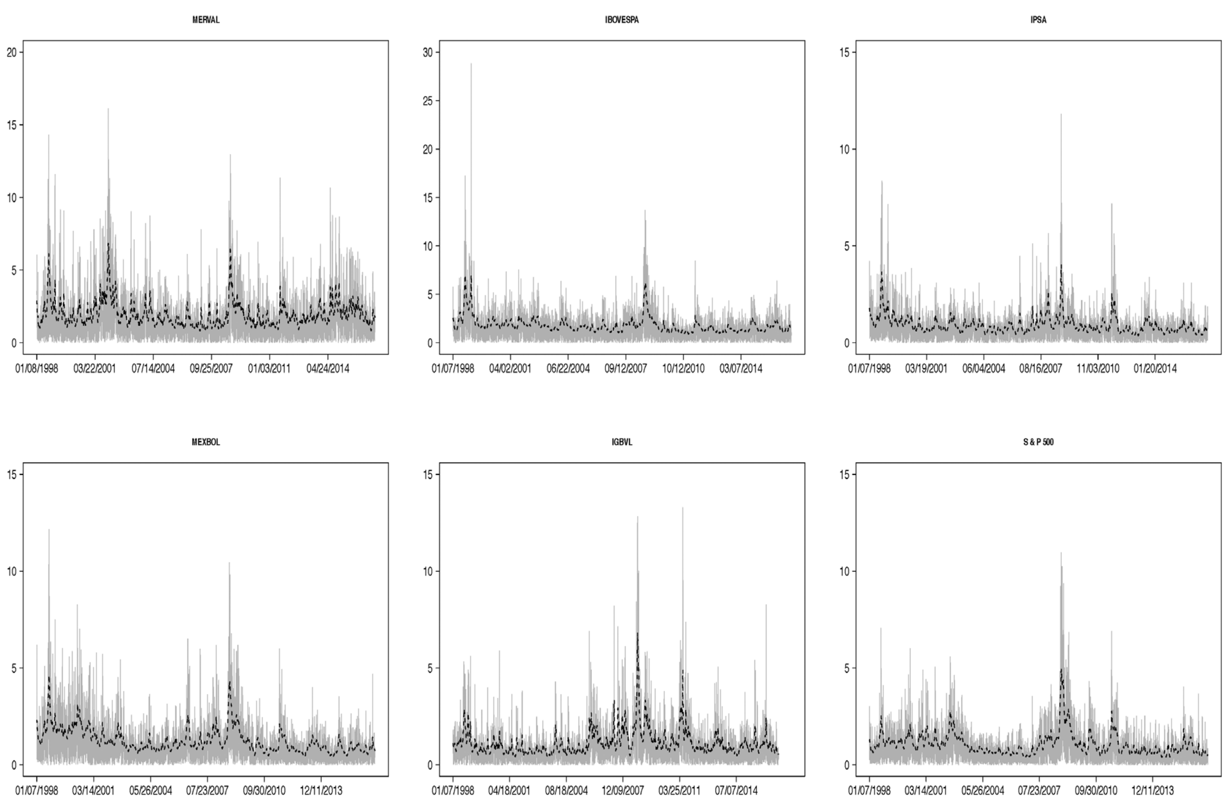

Fig. 7. MERVAL, IBOVESPA, IPSA, MEXBOL, IGBVL and S&P 500 returns data sets: absolute returns (full gray line) and posterior smoothed mean of $e^{\frac{h_t}{2}}$ of the SVM model (dotted black line).

$\mu \sim \mathcal{N}(0, 1000)$, $\frac{\phi+1}{2} \sim \mathcal{B}e(20, 1.5)$, $\sigma^2 \sim \mathcal{IG}(5, 0.025)$, $\beta_0 \sim \mathcal{N}(0, 1000)$, $\frac{\beta_1+1}{2} \sim \mathcal{B}e(3, 1)$ and $\beta_2 \sim \mathcal{N}(0, 1000)$. The prior #1 denotes the prior distribution assumed in the previous estimations (the baseline case). Table 5 shows the posterior estimates of ϕ, σ and β_2: the posterior means and 95% credibility interval. Our results indicate numerical stability for the posterior means and 95% credibility intervals for ϕ, σ and specially for the β_2 parameter that represents the feed back effect, for all the indexes considered here.

In order to compare the in-sample-fit of the SVM model, we estimate the GARCH-M(1,1) model defined by Eqs. (7) and (8) in Appendix B. We conduct an MCMC simulation for 50,000 iterations. The first 10,000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, only every 10*th* values of the chain are stored.

The estimation results for the GARCH-M(1,1) model are given in Table 6, where we observe that estimates for the δ (lagged) parameter are always positive and statistically significant and very similar to the corresponding parameter (β_1) in the SVM model (see Table 3) and the first-order autocorrelations in Table 1. The magnitude of the

Table 5
Prior sensitivity.

Prior	INDEX								
	MERVAL parameters			BOVESPA parameters			IPSA parameters		
	ϕ	σ	β_2	ϕ	σ	β_2	ϕ	σ	β_2
1	0.9501	0.2688	−0.0287	0.9730	0.1649	−0.0259	0.9697	0.2111	−0.0442
	(0.9353,0.9629)	(0.2399,0.3044)	(−0.0510,−0.0070)	(0.9565,0.9863)	(0.1325,0.1968)	(−0.0510,−0.0018)	(0.9599,0.9791)	(0.1880,0.2395)	(−0.0868,−0.0021)
2	0.9441	0.2862	−0.0291	0.9732	0.1725	−0.0256	0.9681	0.2162	−0.0451
	(0.9273,0.9594)	(0.2504,0.3243)	(−0.0502, −0.0075)	(0.9649,0.9817)	(0.1484,0.1927)	(−0.0519, −0.0017)	(0.9569,0.9778)	(0.1904,0.2466)	(−0.0863, −0.0025)
3	0.9431	0.2898	−0.0295	0.9725	0.1749	−0.0257	0.9675	0.2177	−0.0448
	(0.9255,0.9594)	(0.2484,0.3325)	(−0.0478, −0.0091)	(0.9627,0.9810)	(0.1539,0.2011)	(−0.0507, −0.0021)	(0.9566,0.9774)	(0.1916,0.2464)	(−0.0866, −0.0013)
4	0.9495	0.2697	−0.0288	0.9746	0.1686	−0.0249	0.9705	0.2074	−0.0446
	(0.9332,0.9634)	(0.2365,0.3099)	(−0.0500, −0.0075)	(0.9654,0.9831)	(0.1456,0.1940)	(−0.0505, −0.0010)	(0.9602,0.9798)	(0.1822,0.2388)	(−0.0879, −0.0027)
5	0.9418	0.2932	−0.0290	0.9753	0.1728	−0.0257	0.9677	0.2173	−0.0439
	(0.9222,0.9586)	(0.2503,0.3409)	(−0.0505, −0.0081)	(0.9639,0.9817)	(0.1536,0.1934)	(−0.0515, −0.0018)	(0.9565,0.9777)	(0.1912,0.2473)	(−0.0856, −0.0009)
6	0.9489	0.2702	−0.0289	0.9739	0.1696	−0.0253	0.9694	0.2104	−0.0442
	(0.9327,0.9630)	(0.2338,0.3086)	(−0.0505, −0.0084)	(0.9644,0.9825)	(0.1461,0.1916)	(−0.0498, −0.0017)	(0.9596,0.9785)	(0.1859,0.2318)	(−0.0889, −0.0003)

Prior	INDEX								
	MEXBOL parameters			IGBVL parameters			S&P 500 parameters		
	ϕ	σ	β_2	ϕ	σ	β_2	ϕ	σ	β_2
1	0.9803	0.1859	−0.0301	0.9618	0.2725	−0.0114	0.9791	0.1968	−0.0595
	(0.9725,0.9870)	(0.1655,0.2102)	(−0.0581, −0.0020)	(0.9490,0.9732)	(0.2351,0.3102)	(−0.0414,0.0179)	(0.9705,0.9864)	(0.1686,0.2311)	(−0.0924, −0.0272)
2	0.9768	0.2009	−0.0306	0.9584	0.2854	−0.0112	0.9773	0.2047	−0.0596
	(0.9680,0.9848)	(0.1725,0.2313)	(−0.0599, −0.0020)	(0.9444,0.9703)	(0.2513,0.3262)	(−0.0410,0.0186)	(0.9687,0.9847)	(0.1811,0.2367)	(−0.0928, −0.0295)
3	0.9773	0.1944	−0.0311	0.9577	0.2887	−0.0108	0.9767	0.2086	−0.0592
	(0.9685,0.9847)	(0.1754,0.2268)	(−0.0596, −0.0022)	(0.9428,0.9699)	(0.2519,0.3397)	(−0.0409, −0.0204)	(0.9685,0.9844)	(0.1849,0.2370)	(−0.0921, −0.0260)
4	0.9804	0.1853	−0.0300	0.9620	0.2723	−0.0118	0.9793	0.1964	−0.0584
	(0.9731,0.9872)	(0.1646,0.2063)	(−0.0601, −0.0010)	(0.9500,0.9730)	(0.2393,0.3066)	(−0.0433,0.0177)	(0.9715,0.9864)	(0.1690,0.2244)	(−0.0905, −0.0257)
5	0.9781	0.1942	−0.0311	0,9583	0.2855	−0.0116	0.9769	0.2073	−0.0599
	(0.9699,0.9857)	(0.1659,0.2247)	(−0.0604, −0.0013)	(0.9445,0.9706)	(0.2464,0.3246)	(−0.0433,0.0180)	(0.9689,0.9840)	(0.1828,0.2342)	(−0.0927, −0.0263)
6	0.9798	0.1855	−0.0309	0.9591	0.2821	−0.0119	0.9789	0.1961	−0.0588
	(0.9728,0.9865)	(0.1633,0.2070)	(−0.0609, −0.0018)	(0.9452,0.9714)	(0.2434,0.3236)	(−0.0430,0.0118)	(0.9714,0.9863)	(0.1707,0.2193)	(−0.0920, −0.0267)

estimates of the in-mean parameter γ are similar to those obtained with the SVM model in absolute terms, so a large negative contemporaneous relationship in the SVM model is accompanied by a large positive ex-ante relationship in the GARCH-M model. However it is only observed for the cases of Brazil, Chile and Mexico because in the other cases, the null hypothesis of a zero ex ante relationship between excess returns and volatility can never be rejected at the 5% significance level. Compared to the SVM model, this result is consistent only for Peru. For the cases of USA and Argentina, results are in opposition with the in-mean parameter β_2 in the SVM models where it is statistically significant. The case of USA is surprising given that the SVM model suggests a statistically signifi-

Table 6
MCMC estimation of the GARCH-M(1,1) model. We report the posterior mean, the Monte Carlo (MC) error, the 95% credibility interval, the ineficiency factor (Inef) and the convergence diagnostic (CD), respectively.

Parameter	Mean	MC Error	95% interval	Inef	CD	Parameter	Mean	MC Error	95% interval	Inef	CD
MERVAL (Argentina)						IBOVESPA (Brazil)					
ω	0.0322	0.0017	(−0.0520,0.1185)	1.10	0.07	ω	−0.0263	0.0020	(−0.1209,0.0424)	3.95	−0.53
δ	0.0730	0.0005	(0.0502,0.1064)	1.06	0.15	δ	0.0041	0.0003	(0.0000,0.0244)	1.51	1.07
γ	0.0171	0.0004	(−0.0063,0.0373)	1.00	−1.69	γ	0.0341	0.0007	(0.0102,0.0610)	2.81	0.00
α	0.1570	0.0009	(0.1115,0.2041)	1.21	−0.77	α	0.0692	0.0008	(0.0434,0.0932)	4.72	−0.58
β	0.1106	0.0004	(0.0932,0.1330)	1.22	0.84	β	0.0837	0.0005	(0.0691,0.0993)	3.33	0.37
θ	0.8589	0.0005	(0.8316,0.8797)	1.21	−0.12	θ	0.8973	0.0007	(0.8794,0.9170)	4.75	0.48
$\beta+\theta$	0.9695	0.0002	(0.9540,0.9814)	1.27	−1.31	$\beta+\theta$	0.9810	0.0003	(0.9714,0.9898)	5.93	0.48
IPSA (Chile)						MEXBOL (Mexico)					
ω	0.0158	0.0014	(−0.0042,0.0445)	29.17	0.53	ω	0.0225	0.0001	(−0.0183,0.0673)	1.00	0.05
δ	0.1909	0.0012	(0.1642,0.2124)	33.69	−0.50	δ	0.0818	0.0001	(0.0562,0.1105)	1.00	1.01
γ	0.0408	0.0013	(0.0056,0.0632)	22.70	−0.45	γ	0.0359	0.0001	(0.0097,0.0625)	1.00	−1.55
α	0.0245	0.0003	(0.0172,0.0346)	42.67	0.41	α	0.0165	0.0001	(0.0105,0.0258)	1.00	−0.06
β	0.1347	0.0010	(0.1176,0.1573)	29.15	1.62	β	0.0821	0.0004	(0.0675,0.0984)	1.00	0.33
θ	0.8436	0.0011	(0.8143,0.8617)	31.85	−1.06	θ	0.9103	0.0001	(0.8923,0.9623)	1.00	0.33
$\beta+\theta$	0.9783	0.0008	(0.9630,0.9897)	16.51	0.52	$\beta+\theta$	0.9923	0.0001	(0.9820,0.9973)	1.00	−0.65
IGBVL (Peru)						S&P 500 (USA)					
ω	0.0398	0.0006	(0.0026,0.0767)	1.00	−0.05	ω	0.0356	0.0003	(0.0005,0.0713)	1.00	0.68
δ	0.2076	0.0004	(0.1757,0.2387)	1.00	−0.46	δ	0.0059	0.0005	(0.0000,0.0375)	1.00	−0.42
γ	0.0185	0.0004	(−0.0084,0.0439)	1.00	0.13	γ	0.0299	0.0005	(−0.0023,0.0616)	1.00	−1.13
α	0.06186	0.0002	(0.0436,0.0850)	1.00	0.42	α	0.0233	0.0001	(0.0163,0.0302)	1.00	0.97
β	0.1853	0.0006	(0.1542,0.2222)	1.00	0.21	β	0.1018	0.0003	(0.0837,0.1203)	1.00	−0.78
θ	0.7868	0.0005	(0.7435,0.8230)	1.00	−0.10	θ	0.8810	0.0003	(0.8588,0.9006)	1.00	−0.19
$\beta+\theta$	0.9722	0.0003	(0.9542,0.9866)	1.00	0.18	$\beta+\theta$	0.9828	0.0001	(0.9730,0.9898)	1.00	−0.19

Table 7
Model comparison criteria. Log-predictive score (LPS) and Deviance Information Criterion (DIC).

INDEX/MODEL	LPS				DIC			
	GARCH	GARCH-M	SVM	SV	GARCH	GARCH-M	SVM	SV
MERVAL	5.6223	2.0981	**2.0720**	2.0731	53,460.38	19,528.50	**19,285.50**	19,293.51
IBOVESPA	4.3952	1.9763	**1.9720**	1.9729	37,009.72	18,580.21	**18,540.66**	18,543.70
IPSA	2.1169	1.2997	**1.2931**	1.2936	18,816.81	12,325.90	**12,262.62**	12,265.85
MEXBOL	2.8946	1.6038	**1.5875**	1.5878	26,362.53	15,257.72	**15,121.16**	15,122.51
IGBVL	2.8712	1.5192	**1.4942**	1.4944	20,005.94	13,979.51	**13,749.80**	13,750.03
S&P 500	2.4344	1.4293	**1.4134**	1.4150	21,334.94	13,667.49	**13,516.47**	13,528.69

cant in-mean parameter whereas the GARCH-M model suggests not rejection of the zero value hypothesis. Overall, the results confirm the hypothesis that investors require higher expected returns when unanticipated increases in future volatility are highly persistent in the cases of Brazil, Chile and Mexico. According to the results of the GARCH-M model, these three markets present indirect evidence of a positive intertemporal relation between expected excess market returns and its volatility as this is one of the assumptions underlying the volatility feedback hypothesis. The SVM model suggests more and stronger evidence in favor of this hypothesis.

The volatility persistence parameters are comparable, but slightly more persistent to those found for the SVM model with near-unity values for $\beta+\theta$. It is more evident for Mexico.

We also fit the basic stochastic volatility model (SV) and the GARCH(1,1) models, but we do not report the results of parameter estimation. To assess the goodness of the estimated models (GARCH(1,1), GARCH-M, SV and SVM), we calculate the deviance information criteria (DIC) and the log-predictive score (LPS) proposed by (Spiegelhalter, Best, Carlin, & van der Linde, 2002) and Delatola and Griffin (2011), respectively. The DIC criterion is based on the posterior mean of the deviance, which can be approximated by $\overline{D}=\sum_{q=1}^{Q} D(\boldsymbol{\theta}_q)/Q$, where $D(\boldsymbol{\theta})=-2\log p(\mathbf{y}_T \mid \boldsymbol{\theta})=-2\log \mathcal{L}(\boldsymbol{\theta})$. The DIC can be estimated using the MCMC output by $\widehat{DIC}=\overline{D}+\widehat{p_D}=2\overline{D}-D(\bar{\boldsymbol{\theta}})$, where $\widehat{p_D}$ is the effective number of parameters, and can be evaluated as $\widehat{p_D}=\overline{D}-D(\bar{\boldsymbol{\theta}})$. The model that best fits a dataset is the model with the smallest DIC value. The LPS can be estimated as: $\widehat{LPS}=-\frac{1}{T}\sum_{t=1}^{T}\log p(y_t \mid \mathbf{y}_{t-1},\bar{\boldsymbol{\theta}})$. In both cases, the best model has the smallest DIC (LPS). In order to evaluate the likelihood approximation in the SV and SVM models, we apply the HMM where, in order to ensure a good approximation of the estimates, we use $b_m=-b_0=4.5$ and $m=200$; for further details, see Abanto-Valle, Langrock, Chen, and Cardoso (2017). Table 7 shows the LPS and DIC for all the indexes considered here. According to the LPS and DIC values (see values in bold in Table 7) the SVM model outperform the GARCH(1,1), GARCH-M, SV models for all the markets.

5. Discussion

This article presented a Bayesian implementation in order to estimate the SVM model as proposed by Koopman and Uspensky (2002), via HMC and RMHMC methods. The SVM model enabled us to investigate the dynamic relationship between returns and their time-varying volatility. We illustrated our methods through an empirical application of five Latin American return series and the S&P 500 returns. The β_2 estimate, which measures both the ex ante relationship between returns and volatility and the volatility feedback effect, is found to be negative and significant for all the indexes considered here with the exception of the IGBVL. The results are in line with those of French et al. (1987), who found a similar relationship between unexpected volatility dynamics and returns, and confirm the hypothesis that investors require higher expected returns when unanticipated increases in future volatility are highly persistent. This is consistent with our findings of higher values of ϕ combined with larger negative values for the in-mean parameter. We also estimated a GARCH-M model in order to compare estimates of the in-mean parameter. The results are consistent with those of the SVM model only for Brazil, Chile and Mexico. For the other countries the null hypothesis that the in-mean parameter is zero is not rejected which, at least for USA is surprising and rare.

Future research includes extending the model and algorithm to include time-varying parameters including the in-mean parameter. This would allow comparison with other algorithms such as the one proposed in Chan (2017). The second would be to incorporate heavy-tailedness as in Abanto-Valle et al. (2012) or skewness and heavy-tailedness simultaneously as in Leão et al. (2017).

6. Data availability statement

The data that support the findings of this study were obtained from the Yahoo finance web site available to download at http://finance.yahoo.com. The data is also available from the corresponding author upon reasonable request.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgements

We would like to thank the Associate Editor (Professor Narjees Boubakri) and two an anonymous referees for their useful comments, which improved the quality of this paper. The first author acknowledges partial financial support from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro (FAPERJ).

Appendix A.

A.1 Sampling scheme for $(\mu, \phi, \sigma)'$

We assume that $\mu \sim \mathcal{N}(\mu_0, \sigma^2_\mu)$, $(\phi+1)/2 \sim \mathcal{B}e(a_\phi, b_\phi)$, $\sigma^2 \sim \mathcal{IG}(a_\sigma, b_\sigma)$. As $|\phi| < 1$ and $\sigma > 0$, we make the transformations: $\omega = \text{arctanh}(\phi)$ and $\gamma = \log(\sigma)$ such that ω and γ are in the real line. Let $\boldsymbol{\theta}_1 = (\mu, \omega, \gamma)'$ and from Eq. (6), we have:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}_1) &= \text{constant} + \frac{1}{2}\log(1-\phi^2) - \frac{1-\phi^2}{2\sigma^2}(h_1-\mu)^2 - \frac{1}{2\sigma^2}\sum_{t=1}^{T-1}[h_{t+1}-\mu-\phi(h_t-\mu)]^2 \\
&\quad -T\log(\sigma) + (a_\phi - 1)(1+\phi) + (b_\phi - 1)\log(1-\phi) - 2(a_\sigma+1)\log(\sigma) - \frac{b_\sigma}{\sigma^2} \\
&\quad -\frac{1}{2\sigma^2_\mu}(\mu-\mu_0)^2.
\end{aligned}$$

As $\phi = \tanh(\omega)$ and $\sigma = \exp(\gamma)$, then $\frac{d\phi}{d\omega} = 1-\phi^2$ and $\frac{d\sigma}{d\gamma} = \sigma$, so we have that the gradient is given by

$$\triangledown_{\boldsymbol{\theta}_1}\mathcal{L}(\boldsymbol{\theta}_1) = \begin{bmatrix} \triangledown_\mu \mathcal{L}(\boldsymbol{\theta}_1) \\ \triangledown_\omega \mathcal{L}(\boldsymbol{\theta}_1) \\ \triangledown_\gamma \mathcal{L}(\boldsymbol{\theta}_1) \end{bmatrix},$$

where

$$\begin{aligned}
\triangledown_\mu \mathcal{L}(\boldsymbol{\theta}_1) &= \frac{1-\phi^2}{\sigma^2}(h_1-\mu) + \frac{1-\phi}{\sigma^2}\sum_{t=1}^{T-1}[h_{t+1}-\mu-\phi(h_t-\mu)] - \frac{1}{\sigma^2_\mu}(\mu-\mu_0), \\
\triangledown_\omega \mathcal{L}(\boldsymbol{\theta}_1) &= -\phi + \frac{\phi(1-\phi^2)}{\sigma^2}(h_1-\mu)^2 + \frac{1-\phi^2}{\sigma^2}\sum_{t=1}^{T-1}[h_{t+1}-\mu-\phi(h_t-\mu)][h_t-\mu] \\
&\quad +(a_\phi-1)(1-\phi) - (b_\phi-1)(1+\phi), \\
\triangledown_\gamma \mathcal{L}(\boldsymbol{\theta}_1) &= -T + \frac{1-\phi^2}{\sigma^2}(h_1-\mu)^2 + \frac{1}{\sigma^2}\sum_{t=1}^{T-1}[h_{t+1}-\mu-\phi(h_t-\mu)]^2 \\
&\quad -2(a_\sigma+1) + 2\frac{b_\sigma}{\sigma^2}.
\end{aligned}$$

Then, we have the tensor matrix:

$$\mathbf{M}(\boldsymbol{\theta}_1) = \begin{bmatrix} \frac{(1-\phi)^2(T-1)}{\sigma^2} + \frac{1-\phi^2}{\sigma^2} + \frac{1}{\sigma^2_\mu} & 0 & 0 \\ 0 & 2\phi^2 + (T-1)(1-\phi^2) + (a_\phi + b_\phi - 2)(1-\phi^2) & 2\phi \\ 0 & 2\phi & 2T + \frac{4b_\sigma}{\sigma^2} \end{bmatrix}$$

and

$$\begin{aligned}
\frac{\partial \mathbf{M}(\boldsymbol{\theta}_1)}{\partial \mu} &= \mathbf{0}_{3\times 3} \\
\frac{\partial \mathbf{M}(\boldsymbol{\theta}_1)}{\partial \omega} &= \begin{bmatrix} -\frac{2(1-\phi^2)}{\sigma^2}[(1-\phi)(T-1)+\phi] & 0 & 0 \\ 0 & 2\phi(1-\phi^2)[4-(T-1)-a_\phi-b_\phi] & 2(1-\phi^2) \\ 0 & 2(1-\phi^2) & 0 \end{bmatrix} \\
\frac{\partial \mathbf{M}(\boldsymbol{\theta}_1)}{\partial \gamma} &= \begin{bmatrix} -\frac{2}{\sigma^2}[(1-\phi)^2(T-1)+1-\phi^2] & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{-8b_\sigma}{\sigma^2} \end{bmatrix}.
\end{aligned}$$

A.2 Sampling scheme for $(\beta_0, \beta_1, \beta_2)'$

We assume the priors distribution as follows: $\beta_0 \sim \mathcal{N}(\ \bar{\beta}_0, \sigma^2_{\beta_0})$, $(\beta_1+1)/2 \sim \mathcal{B}e(a_{\beta_1}, b_{\beta_1})$ and $\mathcal{N}(\bar{\beta}_2, \sigma^2_{\beta_2})$. As $|\beta_1| < 1$, we make the transformation $\beta_1 = \tanh(\delta)$ and let $\boldsymbol{\theta}_2 = (\beta_0, \delta, \beta_2)'$ and from Eq. (6), we have:

$$\begin{aligned}
\mathcal{L}(\boldsymbol{\theta}_2) &= \text{constant} - \frac{1}{2}\sum_{t=1}^{T} e^{-h_t}[y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 e^{h_t}]^2 - \frac{1}{2\sigma_{\beta_0^2}}(\beta_0 - \bar{\beta}_0)^2 \\
&\quad +(a_{\beta_1}-1)\log(1+\beta_1) + (b_{\beta_1}-1)\log(1-\beta_1) - \frac{1}{2\sigma_{\beta_2^2}}(\beta_2 - \bar{\beta}_2)^2.
\end{aligned}$$

As, $\beta_1 = \tanh(\delta)$, $\frac{d\beta_1}{d\delta} = 1 - \beta_1^2$. Then, we have the gradient is given by

$$\triangledown_{\boldsymbol{\theta}_2}\mathcal{L}(\boldsymbol{\theta}_2) = \begin{bmatrix} \triangledown_{\beta_0}\mathcal{L}(\boldsymbol{\theta}_2) \\ \triangledown_{\delta}\mathcal{L}(\boldsymbol{\theta}_2) \\ \triangledown_{\beta_2}\mathcal{L}(\boldsymbol{\theta}_1) \end{bmatrix},$$

where

$$\begin{aligned}
\triangledown_{\beta_0}\mathcal{L}(\boldsymbol{\theta}_2) &= \sum_{t=1}^{T} \mathrm{e}^{-h_t}[y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 \mathrm{e}^{h_t}] - \frac{1}{\sigma_{\beta_0^2}}(\beta_0 - \bar{\beta}_0), \\
\triangledown_{\delta}\mathcal{L}(\boldsymbol{\theta}_2) &= (1-\beta_1^2)\sum_{t=1}^{T} \mathrm{e}^{-h_t}[y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 \mathrm{e}^{h_t}]y_{t-1} \\
&\quad + (a_{\beta_1} - 1)(1-\beta_1) - (b_{\beta_1} - 1)(1+\beta_1), \\
\triangledown_{\beta_2}\mathcal{L}(\boldsymbol{\theta}_2) &= \sum_{t=1}^{T}[y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 \mathrm{e}^{h_t}] - \frac{1}{\sigma_{\beta_2^2}}(\beta_0 - \bar{\beta}_2).
\end{aligned}$$

Then, we have the tensor matrix:

$$\mathbf{M}(\boldsymbol{\theta}_2) = \begin{bmatrix} \sum_{t=1}^{T}\mathrm{e}^{-h_t} + \frac{1}{\sigma_{\beta_0^2}} & (1-\beta_1^2)\sum_{t=1}^{T}\mathrm{e}^{-h_t}y_{t-1} & T \\ (1-\beta_1^2)\sum_{t=1}^{T}\mathrm{e}^{-h_t}y_{t-1} & (1-\beta_1^2)^2\sum_{t=1}^{T}\mathrm{e}^{-h_t}y_{t-1}^2 + (a_{\beta_1} + b_{\beta_1} - 2)(1-\beta_1^2) & (1-\beta_1^2)\sum_{t=1}^{T}y_{t-1} \\ T & (1-\beta_1^2)\sum_{t=1}^{T}y_{t-1} & \sum_{t=1}^{T}\mathrm{e}^{h_t} + \frac{1}{\sigma_{\beta_2^2}} \end{bmatrix}$$

and

$$\begin{aligned}
\frac{\partial \mathbf{M}(\boldsymbol{\theta}_2)}{\partial \mu} &= \mathbf{0}_{3\times 3}, \\
\frac{\partial \mathbf{M}(\boldsymbol{\theta}_2)}{\partial \omega} &= \begin{bmatrix} 0 & -2\beta_1(1-\beta_1^2)\sum_{t=1}^{T}\mathrm{e}^{-h_t}y_{t-1} & 0 \\ -2\beta_1(1-\beta_1^2)\sum_{t=1}^{T}\mathrm{e}^{-h_t}y_{t-1} & -4\beta_1(1-\beta_1^2)^2\sum_{t=1}^{T}\mathrm{e}^{-h_t}y_{t-1}^2 - 2\beta_1(a_{\beta_1} + b_{\beta_1} - 2)(1-\beta_1^2) & -2\beta_1(1-\beta_1^2)\sum_{t=1}^{T}y_{t-1} \\ 0 & -2\beta_1(1-\beta_1^2)\sum_{t=1}^{T}y_{t-1} & 0 \end{bmatrix}, \\
\frac{\partial \mathbf{M}(\boldsymbol{\theta}_1)}{\partial \beta_2} &= \mathbf{0}_{3\times 3}.
\end{aligned}$$

A.3 Sampling $\boldsymbol{h}_{1:T}$

Let $\mathcal{L}(\mathbf{h}_{1:T})$ be defined by

$$\begin{aligned}
\mathcal{L}(\mathbf{h}_{1:T}) &= \text{constant} + \sum_{t=1}^{T}\frac{h_t}{2} - \frac{1}{2\sigma^2}\sum_{t=1}^{T-1}[h_{t+1} - \mu - \phi(h_t - \mu)]^2 - \frac{1-\phi^2}{2\sigma^2}(h_1 - \mu)^2 \\
&\quad - \frac{1}{2}\sum_{t=1}^{T}\mathrm{e}^{-h_t}(y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 \mathrm{e}^{h_t})^2.
\end{aligned}$$

Then, we have that the gradient $\triangledown_{\mathbf{h}_{1:T}}\mathcal{L}(\mathbf{h}_{1:T})$ is given by

$$\triangledown_{\mathbf{h}_{1:T}}\mathcal{L}(\mathbf{h}_{1:T}) = \begin{bmatrix} \triangledown_{h_1}\mathcal{L}(\mathbf{h}_{1:T}) \\ \vdots \\ \triangledown_{h_t}\mathcal{L}(\mathbf{h}_{1:T}) \\ \vdots \\ \triangledown_{h_T}\mathcal{L}(\mathbf{h}_{1:T}) \end{bmatrix},$$

where

$$\begin{aligned}\nabla_{h_1}\mathcal{L}(\mathbf{h}_{1:T}) &= -\frac{1}{2}+\frac{1}{2}\mathrm{e}^{-h_1}(y_1-\beta_0-\beta_1 y_0-\beta_2\mathrm{e}^{h_1})^2+\beta_2(y_1-\beta_0-\beta_1 y_0-\beta_2\mathrm{e}^{h_1})\\ &\quad+\frac{\phi}{\sigma^2}[h_2-\mu-\phi(h_1-\mu)]-\frac{1-\phi^2}{\sigma^2}(h_1-\mu),\\ \nabla_{h_t}\mathcal{L}(\mathbf{h}_{1:T}) &= -\frac{1}{2}+\frac{1}{2}\mathrm{e}^{-h_t}(y_t-\beta_0-\beta_1 y_{t-1}-\beta_2\mathrm{e}^{h_t})^2+\beta_2(y_t-\beta_0-\beta_1 y_{t-1}-\beta_2\mathrm{e}^{h_t})\\ &\quad+\frac{\phi}{\sigma^2}[h_{t+1}-\mu-\phi(h_t-\mu)]-\frac{1}{\sigma^2}[h_t-\mu-\phi(h_{t-1}-\mu)]\mathrm{for} 1<t<T,\\ \nabla_{h_T}\mathcal{L}(\mathbf{h}_{1:T}) &= -\frac{1}{2}+\frac{1}{2}\mathrm{e}^{-h_T}(y_T-\beta_0-\beta_1 y_{T-1}-\beta_2\mathrm{e}^{h_T})^2+\beta_2(y_T-\beta_0-\beta_1 y_T-\beta_2\mathrm{e}^{h_T})\\ &\quad-\frac{1}{\sigma^2}[h_T-\mu-\phi(h_{T-1}-\mu)]\end{aligned}.$$

Appendix B. Sampling scheme of the parameters in the GARCH-M(1,1) Model

The GARCH(1,1) in mean model is defined by

$$y_t=\omega+\delta y_{t-1}+\gamma h_t+h_t^{1/2}\epsilon_t,\ \epsilon_t\sim\mathcal{N}(0,1) \tag{7}$$

$$h_t=\alpha+\beta(h_t\epsilon_{t-1}^2)+\theta h_{t-1}, \tag{8}$$

where $|\delta|<1$, $\alpha\geq 0$, $\beta,\theta>0$ and $\beta+\theta<1$. We use the following reparameterization $\psi=\log(\delta/(1-\delta))$, $\kappa=\log(\alpha)$, $\lambda=\log(\beta/(1-\beta))$ and $\rho=\log(\theta/(1-\beta-\theta))$. Let $\boldsymbol{\theta}=(\omega,\psi,\gamma,\kappa,\lambda,\rho)'$. We assume $\boldsymbol{\theta}\sim\mathcal{N}_6(\boldsymbol{\theta}_0,\boldsymbol{\Sigma}_0)$, where $\mathcal{N}_q(.,.)$ denotes the q-variate normal distribution. Then, the posterior distribution of $\boldsymbol{\theta}$ is given by

$$\pi(\boldsymbol{\theta}\mid y_{-1},y_0,y_1,\ldots,y_t)\propto\prod_{t=1}^{T}h_t^{-1/2}\exp\{-\frac{1}{2h_t}(y_t-\omega-\delta y_{t-1}-\gamma h_t)^2\}\times\exp\{-\frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)'\boldsymbol{\Sigma}_0^{-1}(\boldsymbol{\theta}-\boldsymbol{\theta}_0)\}.$$

As $\pi(\boldsymbol{\theta}\mid y_{-1},y_0,y_1,\ldots,y_t)$ is not available in closed form, we sample from using a Metropolis-Hastings step with proposal $N_6(\boldsymbol{\theta}_1,\boldsymbol{\Sigma}_1)$, where $\boldsymbol{\theta}_1$ is the maximum a posteriori and $\boldsymbol{\Sigma}_1$ the inverse of the Hessian matrix evaluated at $\boldsymbol{\theta}_1$.

References

Abanto-Valle, C. A., Migon, H. S., & Lachos, V. H. (2011). Stochastic volatility in mean models with scale mixtures of normal distributions and correlated errors: A bayesian approach. *Journal of Statistical Planning and Inference*, *141*, 1875–1887.

Abanto-Valle, C. A., Migon, H. S., & Lachos, V. H. (2012). Stochastic volatility in mean models with heavy-tailed distributions. *Brazilian Journal of Probability and Statistics*, *26*, 402–422.

Abanto-Valle, C. A., Langrock, R., Chen, M.-H., & Cardoso, M. V. (2017). Maximum likelihood estimation for stochastic volatility in mean models with heavy-tailed distributions. *Applied Stochastic Models in Business and Industry*, *33*, 394–408.

Bekaert, G., & Wu, G. (2000). Asymmetric volatility and risk in equity markets. *Review of Financial Studies*, *13*, 1–42.

Black, F. (1976). Studies of stock price volatility changes. *Proceedings of the 1976 meetings of the American statistical association, business and economical statistics section*, 177–181.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). Arch modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics*, *52*, 5–59.

Bollerslev, T., Engle, R. F., & Nelson, D. B. (1994). Arch models. In R. F. Engle, & D. L. MCFadden (Eds.), *Handbook of econometrics* (Vol. 4) (pp. 2959–3038). Amsterdam: Elsevier Science.

Campbell, J. Y., & Entchel, L. H. (2000). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics*, *31*, 281–318.

Carnero, M. A., Peña, D., & Ruiz, E. (2004). Persistence and kurtosis in GARCH and Stochastic volatility models. *Journal of Financial Econometrics*, *2*, 319–342.

Chan, J. C. (2017). The stochastic volatility in mean model with time-varying parameters: An application to inflation modeling. *Journal of Applied Econometrics*, *35*, 17–28.

Chib, S. (2001). *Markov Chain Monte Carlo methods: Computation and inference*. North-Holland, Amsterdam

Cowles, M. K., & Carlin, B. P. (1996). Markov Chain Monte Carlo convergence diagnostics: A comparative review. *Journal of the American Statistical Association*, *91*, 883–904.

Delatola, E.-I., & Griffin, J. E. (2011). Bayesian nonparametric modelling of the return distribution with stochastic volatility. *Bayesian Analysis*, *6*, 901–926.

Diebold, F. X., & Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics*, *105*, 131–159.

Diebold, F. X., & Lopes, J. A. (1995). Modeling volatility dynamics. In H. K (Ed.), *Macroeconometrics: Developments, tensions and prospects* (pp. 427–472). Amsterdam: Kluwer Academic Publishers.

Duane, S., Kennedy, A. D., Pendleton, B. J., & Roweth, D. (1987). Hybrid Monte Carlo. *Physics Letters B*, *195*, 216–222.

French, K. R., Schert, W. G., & Stambugh, R. F. (1987). Expected stock return and volatility. *Journal of Financial Economics*, *19*, 3–29.

Gelfand, A. E., & Smith, A. F. (1990). ampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, *85*, 398–409.

Geman, S., & Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis Machine Intelligence*, *6*, 721–741.

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. Dawid, & A. F. M. Smith (Eds.), *Bayesian statistics* (Vol. 4) (pp. 169–193).

Ghysels, E., Harvey, A. C., & Renault, E. (1994). Stochastic volatility. In G. S. Madala, & C. R. Rao (Eds.), *Handbook of statistics, Statistical Methods in Finance* (Vol. 14) (pp. 128–198). Amsterdam: Elsevier Science.

Girolami, M., & Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of Royal Statistical Society, Series B*, *73*, 122–214.

Gough, B. (2009). *GNU scientific library reference manual*. Network Theory Ltd.

Hastings, W. K. (1973). Monte Carlo sampling methods using markov chains and their applications. *Biometrika*, *57*, 97–100.

Koopman, S. J., & Uspensky, E. H. (2002). The stochastic volatility in mean model: Empirical evidence from international tock markets. *Journal of Applied Econometrics*, *17*, 667–689.

Kreuzer, A., & Czado, C. (2019). *Bayesian inference for a single factor copula stochastic volatility model using Hamiltonian Monte Carlo.*. , Technical Report. https://arxiv.org/abs/1808.08624

Leão, W. L., Abanto-Valle, C. A., & Chen, M.-H. (2017). Stochastic volatility in mean model with leverage and asymmetrically heavy-tailed error using generalized hyperbolic skew student's t distribution. *Statistics and its Interface*, *10*, 29–541.

Leimkuhler, B., & Reich, S. (2004). *Simulating Hamiltonian dynamics*. Cambridge: Cambridge University Press.

Martin, B., & M, M. J. (2013). Constrained Hamiltonian Monte Carlo in bekk garch with targeting. *Journal of Time Series Econometrics*, *7*, 95–113.

Martin, B., & Maheu, J. M. (2013). Bayesian adaptively updated Hamiltonian Monte Carlo with an application to high-dimensional bekk garch models. *Studies in Nonlinear Dynamics & Econometrics*, *17*, 345–372.

Melino, A., & Turnbull, S. M. (1990). Pricing foreign options with stochastic volatility. *Journal of Econometrics*, *45*, 239–265.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, *21*, 1087–1092.

Neal, R. M. (1993). *Probabilistic inference using Markov Chain Monte Carlo methods*. Canada: University of Toronto. Technical Report CRG-TR-93-1 Department of Computer Science.

Neal, R. M. (2011). Mcmc using Hamiltonian dynamics. In S. Brooks, A. Gelman, G. Jones, & X.-L. Meng (Eds.), *Handbook of Markov chain Monte Carlo* (pp. 113–162). Boca Raton: Chapmall and Hall/CRC Press.

Nugroho, D. B., & Morimoto, T. (2016). Box-cox realized asymmetric stochastic volatility models with generalized student's t-error distributions. *Journal of Applied Statistics*, *43*, 1906–1927.

Nugroho, D. B., & Morimoto, T. (2015). Estimation of realized stochastic volatility models using Hamiltonian Monte Carlo-based methods. *Computational Statistics*, *30*, 491–516.

Perron, P., & Qu, Z. (2010). Long-memory and level shifts in the volatility of stock market return indices. *Journal of Business and Economic Statistics*, *28*, 275–290.

Rodríguez, G. (2017). Modeling latin-american stock and forex markets volatility: Empirical application of a model with random level shifts and genuine long memory. *The North American Journal of Economics and Finance*, *42*, 393–420.

Sanderson, C., & Curtin, R. (2016). Armadillo: A template-based c++ library for linear algebra. *Journal of Open Source Software*, *1*, 26.

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance*, *44*, 1115–1153.

Shephard, N. G. (1996). Statistical aspects of arch and stochastic volatility. In D. R. Cox, D. V. Hinkley, & O. E. Barndorff-Nielsen (Eds.), *Time series models in* *econometrics, finance and other fields, monographs on statistics and applied probability 65* (pp. 1–67). London: Chapman and Hall.

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.

Takaishi, T. (2009). Bayesian inference of stochastic volatility nodel by hybrid Monte Carlo. *Journal of Circuits, Systems, and Computers*, *18*, 1381–1396.

Tanner, M. A., & Wong, W. H. (1987). The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association*, *82*, 528–550.

Taylor, S. J. (1994). Modelling stochastic volatility: A review and comparative study. *Mathematical Finance*, *4*, 183–204.

Zevallos, M., Gasco, L., & Ehlers, R. (2017). Riemann manifold langevin methods on stochastic volatility estimation. *Communications in Statistics – Simulation and Computation*, *46*, 7942–7956.