An Application of an Stochastic Volatility Model with Leverage and Heavy-Tailed Errors to Latin-American Stock Returns using a GH Skew Student’s t-Distribution

Patricia Lengua Lafosse
Pontificia Universidad Católica del Perú
Gabriel Rodríguez*
Pontificia Universidad Católica del Perú

May 31, 2014

Abstract
Following the approach of Nakajima and Omori (2012), we use a Bayesian analysis of a stochastic volatility (SV) model with leverage effect and asymmetrically heavy-tailed error using generalized hyperbolic(GH) student’s t-distribution applied to the returns of a set of Latin-American countries (Argentina, Brazil, Chile, Mexico and Peru) for daily data for the period 1996-01-2013:12. The SVSKt model is a general model that contains the SV model with symmetric Student’s t error distribution and the standard SV model with normal error distribution. The results suggest that there are leverage effects in all analyzed countries except to IGBVL and the presence of asymmetric heavy-tailed errors is confirmed. A comparison with estimates of the US S&P500 returns is also provided. The models for stock returns are compared based on the marginal likelihood in the empirical study. Overall, the results show evidence that GH-skew Student’s t-error distribution in SV model is clearly successful in describing the distribution of the daily stock return data of the Latin American indexes considered over the traditional symmetrically heavy-tailed error using Student’s t-distribution.

KeyWords: Stochastic Volatility, Generalized Hyperbolic Skew Student’s t-Distribution, Markov Chain Monte Carlo, Latin American Stock Returns, Latin American Stock Markets

JEl Classification: C11, C15, C58, G17

*Address for Correspondence: Gabriel Rodríguez, Department of Economics, Pontificia Universidad Católica del Perú, Av. Universitaria 1801, Lima 32, Lima, Perú, Telephone: +511-626-2000 (4998), Fax: +511-626-2874. E-Mail Address: gabriel.rodriguez@pucp.edu.pe.
1 Introduction

The returns of financial stocks or foreign exchange rates are characterized by the presence of some factors which are generally presented in financial time series. These factors have the following three important characteristics: i) The returns does not have a Normal distribution, ii) There exists reduced correlations in the levels of the returns of different days and iii) The correlation between the magnitudes of the returns (the series of the square of the returns) is positive and statistically significative (Taylor, 2005). These characteristics are explained in most cases with the presence of changes in the volatility over time.

The model of time-variant volatility has been widely used in the literature of financial time series. There are two broad streams in the modeling of these financial time series where variance is changing over time. The first stream is the Autoregressive Conditional Heteroscedasticity (ARCH) models based on the initial proposal of Engle (1982) and later generalized by Bollerslev (1986). This kind of models have been extended to include integrated variance, asymmetries, fractional behavior, etc (see Laurent (2012) for further details). The second type of models are the so called stochastic volatility (SV) models surveyed in Shephard (2005), Taylor (1994) and Kim et al. (1998). Both kind of streams try to model and reproduce the principal properties of the financial time series which are excess of kurtosis, clustering in the volatility periods, correlation in the squared residuals. However, the difference between both set of model is in the statistical approach. While in the ARCH models the volatility is considered as a function of the past squared shocks and the past observed values of the own volatility; in the stochastic volatility models, the volatility is a non observable component where its logarithms are modeled using a linear autoregressive process.

Departures from Normality in the distribution of the errors have originated propositions of other distributions, in particular in order to capture heavy-tailedness of the asset return distribution in the SV context, heavy-tailed errors are often incorporated using distributions such as Student´s t-distribution (Chib et al. (2002), Berg et al. (2004), Yu (2005), Omori et al. (2007) and Nakajima and Omori (2009)) and Skew-GED distribution (Cappuccio et al (2003)). Nakajima and Omori (2012) use a distribution which offers a more flexible skewness and heavy-tailedness: the Generalized Hyperbolic (GH) distribution proposed by Barndorf-Nielsen (1977). An advantage of the GH distribution is that it includes a broad variety of class of distributions as Normal, Hyperbolic, Normal Inverse Gamma and Skew Student’s t-distributions (Nakajima and Omori (2012), Fernández and Steel (1998); Ass and Haff (2006)).

Nakajima and Omori (2012) have proposed an efficient Bayesian estimation method for the SV model incorporating asymmetrically heavy-tailed error, using the GH skew Student’s t-distribution. This model incorporates skewness and includes the SV model with normal errors (SV-Normal) and the SV model with
student’s t-errors (SVt). The GH skew Student’s t-distribution belongs to the class of GH distributions, which have been well studied in the literature; see Prause (1999), Jones and Faddy (2003), Ass and Haff (2006). See also Silva et al. (2006) for a stochastic volatility model without leverage effects using the Generalized Inverse Gaussian (GIG) distribution.

The GH skew Student’s t-distribution is difficult to implement in the context of SV models due to the large numbers of latent volatility variables (Nakajima and Omori, 2012). It is computational burden to repeat filtering many times to evaluate the likelihood function for each set of parameters until find a maximum. Nakajima and Omori (2012) develop a novel Markov chain Monte Carlo (MCMC) algorithm for the efficient estimation of the SV model with leverage and asymmetrically heavy-tailed error using the GH skew Student’s t-distribution. The GH skew Student’s t-distribution is simple, flexible and easily incorporated for a Bayesian estimation scheme using the MCMC proposed by Nakajima and Omori (2012)\(^1\). The principal point to implement the MCMC algorithm is to express de GH skew Student’s t-distribution as a Normal variance-mean mixture of the GIG distribution. In specific, Nakajima and Omori (2012) consider an Inverse Gamma distribution as a mixing distribution among the class of GIG distributions to nest and extend various existing stochastic volatility models.

The empirical evidence about the returns of financial stocks for Latin-American countries is very scarce. Therefore, the overall aim of this paper is to estimate a SV model taking into account asymmetric heavy tails given that the errors have asymmetric GH skew Student’s t-distribution using the algorithm proposed by Nakajima and Omori (2012).

This paper is organized as follows. In Section 2, we describe a standard SV model and introduce the GH Student’s heavy-tailedness in the error distribution. This section shows also the algorithm proposed by Nakajima and Omori (2012). Section 3 shows the empirical application. The data used is the returns of the stock markets of Argentina, Brazil, Chile, Peru, and Mexico. In order to compare the results, we also include results for the US. Conclusions are presented in Section 4.

### 2 Stochastic Volatility Model

The standard SV models assume that the variance of financial returns has been generated under a latent stochastic process. At each period, the process determines the volatilities which follows a latent process, that is, the realized variances are not observable. The representation of a simple SV model usually has the following structure:

\(^1\)Other types of skew t-distributions used in the literature may be found in Hansen (1994), Fernández and Steel (1998), Prause (1999), Jones and Faddy (2003), Azzalini and Capitanio (2003), and Aas and Haff (2006).
\begin{align*}
y_t &= \epsilon_t \exp(h_t/2), \quad (1) \\
h_{t+1} &= \mu + \phi(h_t - \mu) + \eta_t, \quad (2) \\
\epsilon_t &\sim N(0, 1), \quad (3) \\
\eta_t &\sim N(0, \sigma^2) \quad (4)
\end{align*}

where \( y_t \) is the asset return at time \( t = 1, \ldots, T \), \( h_t \) is the logarithm of the volatility of return at time \( t \) which is assumed to follow a stationary process type AR(1). To ensure the stationary of \( h_t \), the persistence parameter \( \phi \) is assumed that \( |\phi| \leq 1 \) and the initial value, \( h_1 \), is assumed to follow the stationary distribution by setting \( h_0 = \mu \) and \( \eta_0 \sim N(0, \sigma^2/(1 - \phi^2)) \). Finally, the error terms \((\epsilon_t, \eta_t)\) are independently and identically distributed according to a bivariate normal distribution with mean \( 0 = (0, 0)' \) and covariance matrix \( \Sigma = \begin{bmatrix} 1 & 0 \\ 0 & \sigma^2 \end{bmatrix} \) where \( \sigma^2 \) is the volatility of the log-volatility and \( \epsilon_t \) and \( \eta_t \) are uncorrelated shocks.

There are factors that the simple SV model does not capture, these are (i) excess of kurtosis and (ii) the presence of correlation between \( \epsilon_t \) and \( \eta_t \). Other major characteristics of the return distribution for financial variables are their skewness, heavy-tailedness and volatility clustering with “leverage effects”: a negative past innovation on asset returns tends to increase the current volatility (Deschamps, 2012). The excess of kurtosis and heavy-tailedness in the distribution of the errors justifies the introduction of heavy-tailed disturbances and have originated propositions of other distributions in order to capture leverage effects and heavy-tailedness. The SV model with Student-t errors is one of the models to account for heavy-tailed returns (e.g., Berg et al. (2004), Chib et al. (2002), Harvey et al. (1994), Nakajima and Omori (2009), Omori et al. (2007) and Yu (2005)). On the side of the correlation, the simple SV model does not allow that the volatility reacts asymmetrically with positive or negative movements in returns since \( \sigma^2 \) is independent of previous returns signs. One factor considered important in financial time series is the “leverage effect”. These asymmetric volatility effects can be incorporated in the SV model assuming that there is any association between the shocks to the returns and volatility shocks (Nakajima, 2012).

2.1 A Stochastic Volatility Model with Leverage and Heavy-Tailed Errors

According to Nakajima and Omori (2012), the stochastic volatility model may be formulated in the following way introducing leverage effects and heavy-tailed errors:
\[ y_t = \exp(h_t/2)\epsilon_t, \quad t = 1, \ldots, T, \quad (5) \]
\[ h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \ldots, T - 1, \quad (6) \]
\[ \left( \begin{array}{c} \epsilon_t \\ \eta_t \end{array} \right) \sim N(0, \Sigma) \quad \text{and} \quad \Sigma = \begin{bmatrix} 1 & \rho \sigma \\ \rho \sigma & \sigma^2 \end{bmatrix}. \quad (7) \]

where \( y_t \) is the asset return, and \( h_t \) is the unobserved log-volatility. We assume that \(|\phi| < 1\) implying that the log-volatility process is stationary and the initial value \( h_1 \) is assumed to follow a stationary distribution fixing \( h_0 = \mu \) and \( \eta_0 \sim N(0, \sigma^2/(1 - \phi^2)) \). The parameter \( \rho \) measures the correlations between \( \epsilon_t \) and \( \eta_t \). When \( \rho < 0 \) we observe leverage effects and when \( \rho = 0 \) there is not this type of effects; see Yu (2005) and Omori et al. (2007).

In order to have a joint model of the leverage and asymmetric heavy-tailedness, Nakajima and Omori (2012) replace the random variable \( \epsilon_t \) in (5) by a random variable from a GH skew Student’s distribution, denoted by:

\[ \omega_t = \mu_\omega + \beta z_t + \sqrt{z_t} \epsilon_t, \quad (8) \]

where \( \epsilon_t \sim N(0, 1) \) and \( z_t \sim IG(\nu/2, \nu/2) \) where IG denotes the inverse Gamma distribution. Nakajima and Omori (2012) assumes that \( \mu_\omega = -\beta \mu_z \), where \( \mu_z \equiv \mathbb{E}(z_t) = \nu/(\nu - 2) \), for \( \mathbb{E}(\omega_t) = 0 \), and \( \nu > 4 \) in order to have a finite variance of \( \omega_t^2 \). When \( \beta = 0 \), the model reduces to the SV model with symmetric Student’s t-distribution (SVt model) and when \( \beta = 0 \) and \( z_t = 1 \) \( \forall t \), the model reduces to the standard SV with a normal error distribution. For an explanation of the relationship between \( (\beta, \nu) \) parameters in relation to the skewness and heavy-tailedness see Nakajima and Omori (2012) where a simulation evidence is presented. A lower value of \( \beta \) (\( \nu \) fixed) implies a more negative skewness or left-skewness as well as heavier tails. On the other hand, as \( \nu \) becomes larger (\( \beta \) fixed) the density becomes less skewed and has lighter tails. Hence the skewness and heavy-tailedness are determined jointly by the combination of the parameter values of \( \beta \) and \( \nu \).

Previous research has been shown that estimation of the parameters of the GH distribution are difficult to estimate due to different issues as flatness of the likelihood function, several local maxima; see Prause (1999), Aas and Haff (2006) for further details. Similar research argues that even the GH skew Student’s t-distribution model has several difficulties with some values of the parameters \( \lambda, \nu \) and \( \gamma \). In order to avoid some of the issues above mentioned, Nakajima and Omori (2012) make an additional assumption that \( \delta = \sqrt{\nu} \) as

\[ \epsilon_t \sim N(0, 1) \] and \( z_t^* \sim GIG(\lambda, \delta, \gamma) \). The GH skew Student’s t-distribution of (8) is the case where \( \lambda = -\nu/2(\nu > 0), \delta = \sqrt{\nu} \) and \( \gamma = 0 \) which yields \( z_t \sim GIG(-\nu/2, \sqrt{\nu}, 0) \), and equivalently \( IG(\nu/2, \nu/2) \).
formulated by (8). Therefore, using this GH Skew Student’s t-distribution, Nakajima and Omori (2012) propose the following stochastic volatility model (SVSKt model, hereafter):

\[
y_t = \{\beta(z_t - \mu_z) + \sqrt{\sigma_t} \epsilon_t\} \exp(h_t/2), \quad t = 1, \ldots, T;
\]

\[
h_{t+1} = \mu + \phi(h_t - \mu) + \eta_t, \quad t = 1, \ldots, T - 1;
\]

\[
z_t \sim IG(\nu/2, \nu/2),
\]

where \((\epsilon_t, \eta_t)\) are as in (7). The value of \(\nu > 4\) is the degree of freedom and unknown to be estimated. The SVSKt model is a general model that contains the SVt model and the standard SV model with normal error distribution. When \(\beta = 0\), the model reduces to the stochastic volatility model with the symmetric Student’s t-distribution (SVt) which has been widely used and analyzed in the previous literature; see Chib et al. (2002), Eraker et al. (2003), Yu (2005) and Omori et al. (2007).

3

The formulation in (9) is not simple but it possible to be estimated using the MCMC algorithm proposed by Nakajima and Omori (2012). The principal element of the model is to express the skew Student’s t-distribution in the form of the normal variance-mean mixture, as established in equation (8). Nakajima and Omori (2012) observe the variable \(z_t\), following the mixing distribution as a latent variable for a new form to implement the MCMC algorithm in the context of Bayesian inference. The conditional posterior distribution of each parameters is reduced to a more tractable form conditional on \(z_t\) that when the model is considered in the direct likelihood form of the skew Student’s t-distribution.

2.2 Bayesian Estimation of Stochastic Volatility Models

This paper is based on the approach suggested by Nakajima and Omori (2012) which follows a Bayesian Approach using MCMC methods. In a first moment, we introduce some elementary elements related to the Bayesian inference or Bayesian methods in general. After it, we introduce the algorithm of Nakajima and Omori (2012) which allow to introduce leverage jointly with the modeling of GH skew Student’s t-distribution in the errors. Description of the algorithm follows closely the description and notation used by Nakajima and Omori (2012).

2.2.1 Some Preliminaries

According to Nakajima (2012), the likelihood of the stochastic volatility model may be obtained based on methods of simulation for a given set of parameters.

\footnote{As Nakajima and Omori (2012) mention, there are other definitions for the skew t-distribution in the literature. Even there are other skew distributions with heavy tails as overviewed by Ass and Haaf (2006). It is possible to include other skew distributions or a more general GH distribution in the stochastic volatility model. However this would lead to an over parameterization because the second moment is already modeled as a latent stochastic process in the stochastic volatility model.}
This process is, however, computational intensive because need to repeat the process many times to evaluate the likelihood for each set of parameters until to achieve a maximum. To avoid this difficulty, we need the Bayesian approach based on MCMC methods.

The estimation of SV models involve estimating two sets of variables. First, the estimation of the set of parameters of the model. Second, the estimation of the unobservable processes or latent processes. In a general SV model, the estimations of the parameters \( \theta = (\mu, \phi, \sigma_\eta) \) requires the construction of the likelihood function which must be maximized. The likelihood function is given by

\[
f(y \mid \theta) = \int f(y \mid h, \theta) f(h \mid \theta) dh.
\] (12)

It is clear that under the stochastic volatility model, the set of observations (returns) is conditional to the vector of non observable volatilities. This expression (integral) should be performed in each point of the time of the sample \( h_t = [h = 1, 2, ..., T] \), that is, this procedure must be performed \( T \) times. Furthermore, Jaquier et al. (1994) argue that the likelihood function has no analytical representation.

In the literature, three approaches have been developed in order to overcome the above mentioned difficulty. The first is the approach through the Kalman filter; the second is using techniques of quasi-likelihood or quasi-maximum method of moments; and the third alternative is the Bayesian approach.

Using the Kalman filter allows to estimate the unobservable components and the parameters by quasi-likelihood. The limitation consists in the approximation no lineal of the error \( \log(\sigma_t^2) \). The filter of Kalman works optimally with linear systems but if the system has other characteristics as in the current case, the parameters will be estimated with a clear bias. Another alternative is the method of moments as mentioned by Jacquier et al. (1994). However the principal drawback is the determination of the number of moments which is important because it could drive to an important loss of information from the data. The third alternative, the Bayesian approach, which is used in this work, has the advantage to be efficient in the estimation of the parameters and in the filter to the extraction of the latent processes.

In the Bayesian approach, the aim is to find the posterior distributions. By Bayes Theorem, if we have the set of parameters \( \theta \), the data and then the posterior distribution will be:

\[
\pi(\theta \mid y) \propto \pi(y \mid \theta) \pi(\theta),
\] (13)

where \( \pi(\theta \mid y) \) is the posterior distribution conditional on the data whose mean is the Bayesian estimation of the parameter. On the other hand, \( \pi(y \mid \theta) \) is the likelihood function, and \( \pi(\theta) \) are the priors, which are the beliefs about the
distributions of the parameters. However, as mentioned above, the calculus of the likelihood function is mostly unfeasible. This is why techniques of Monte Carlo Markov Chain (MCMC, hereafter) are used in the estimation process and they allow a direct estimation with multiple simulations of the posterior distributions. There are several algorithms to reach this goal. In the next subsection, we describe and follow the algorithm of Nakajima and Omori (2012) which is used in the present document.

2.2.2 The Algorithm of Nakajima and Omori (2012)

To describe the MCMC algorithm of Nakajima and Omori (2012), we follow same notation: \( \theta = (\phi, \sigma, \rho, \mu, \beta, \nu) \), \{\( y_t \)\}_{t=1}^{T}, \{\( h_t \)\}_{t=1}^{T}, \{\( z_t \)\}_{t=1}^{T}. \) For the prior distributions of \( \mu \) and \( \beta \) we assume \( \mu \sim N(\mu_0, \nu_0^2) \) and \( \beta \sim N(\beta_0, \sigma_0^2). \) We let \( \pi(\phi), \pi(\theta) \) and \( \pi(\nu) \) denote the prior probability densities of \( \phi, \theta \equiv (\sigma, \rho)^T \) and \( \nu \) respectively. Random samples are drawn from the posterior distribution of \((\theta, h, z)\) given \( y_t \) for the SVSKt model using the MCMC method as follows:

1. Initialize \( \theta, h \) and \( z. \)
2. Generate \( \phi|\sigma, \rho, \mu, \beta, \nu, h, z, y. \)
3. Generate \( (\sigma, \rho)|\phi, \mu, \beta, \nu, h, z, y. \)
4. Generate \( \mu|\phi, \sigma, \rho, \beta, \nu, h, z, y. \)
5. Generate \( \beta|\phi, \sigma, \rho, \mu, v, h, z, y. \)
6. Generate \( \nu|\phi, \sigma, \rho, \mu, \beta, h, z, y. \)
7. Generate \( z|\theta, h, y. \)
8. Generate \( h|\theta, z, y. \)
9. Go to 2.

In the following subsections, we present each sampling step in detail.

Generation of the parameters \((\phi, \sigma, \rho, \mu)\) (steps 2–4)

Step 2. The conditional posterior probability density \( \pi(\phi|\sigma, \rho, \mu, \beta, \nu, h, z, y) (\equiv \pi(\phi|.) \) is

\[
\pi(\phi|.) \propto \pi(\phi) \sqrt{1 - \phi^2} \exp \left\{ - \frac{(1-\phi^2)\bar{h}_t^2}{2\sigma^2} \right\}, \quad (14)
\]

\[
\propto \pi(\phi) \sqrt{1 - \phi^2} \exp \left\{ - \frac{(\phi - \mu_\phi)^2}{2\sigma_\phi^2} \right\}, \quad (15)
\]
where $\bar{h}_t = h_t - \mu$, $\bar{y}_t = \rho \sigma(y_t e^{-h_t/2} - \beta z_t)/\sqrt{z_t}$, $z_t = z_t - \mu_z$, and
\[
\begin{align*}
\mu_\phi &= \frac{\sum_{t=1}^T (\bar{h}_{t+1} - \bar{y}_t) \bar{h}_t}{\rho^2 \bar{h}_1^2 + \sum_{t=2}^T \bar{h}_t^2}, \quad (16) \\
\sigma_\phi^2 &= \frac{\sigma^2 (1 - \rho^2)}{\rho^2 \bar{h}_1^2 + \sum_{t=2}^T \bar{h}_t^2}. \quad (17)
\end{align*}
\]

In order to sample from this conditional posterior distribution, we implement the Metropolis–Hastings (MH, hereafter) algorithm (see, e.g., Chib and Greenberg (1995)). We propose a candidate, $\phi^* \sim TN(-1,1)(\mu_\phi, \sigma^2_\phi)$, where $TN(a,b)(\mu, \sigma^2)$ denotes the normal distribution with mean $\mu$ and variance $\sigma^2$ truncated on the interval $(a, b)$. Then, we accept it with probability given by
\[
\min \left\{ \frac{\pi(\phi^*) \sqrt{1 - \phi^2}}{\pi(\phi^*) \sqrt{1 - \phi^2}}, \quad (18) \right\}
\]

**Step 3.** Because the joint conditional posterior probability density $\pi(\theta|\phi, \mu, \nu, h, z, y)(\equiv \pi(\theta|::))$ of $\theta = (\sigma, \rho)'$ is given by
\[
\pi(\theta|::) \propto \pi(\theta) \sigma^T (1 - \rho^2)^{T-1} \exp \left\{ - \sum_{t=1}^T \frac{-(1-\phi^2) \bar{h}_1^2}{2 \sigma^2 (1 - \rho^2)} \right\}, \quad (19)
\]
which is a probability density from which it is not easy to sample. Nakajima and Omori (2012) apply the MH algorithm based on a normal approximation of the density around the mode. Because we have a constraint, $R = \{ \theta : \sigma > 0, |\rho| < 1 \}$, on the parameter space of the posterior distribution, we consider the transformation $\theta = (\omega_1, \omega_2)'$, where $\omega_1 = \log \sigma$, and $\omega_2 = \log(1 + \rho) - \log(1 - \rho)$, to generate a candidate using a normal distribution. We first search for $\hat{\theta}$ that maximizes (or approximately maximizes) $\pi(\theta|::)$, and obtain its transformed value $\hat{\omega}$. We next generate a candidate $\omega^* \sim N(\omega_*, \Sigma_*)$, where
\[
\omega_* = \hat{\omega} + \Sigma_* \frac{\partial \log \tilde{\pi}(\omega_*)}{\partial \omega} \bigg|_{\omega = \hat{\omega}}, \quad (20)
\]
\[
\Sigma_*^{-1} = \left. \frac{\partial^2 \log \tilde{\pi}(\omega_*)}{\partial \omega \partial \omega'} \right|_{\omega = \hat{\omega}}, \quad (21)
\]
where $\tilde{\pi}(\omega_*)$ is a transformed conditional posterior density. Then, we accept the candidate $\omega^*$ with probability
\[
\min \left\{ \frac{\pi(\theta^*|::) f_N(\omega|\omega_*, \Sigma_*)|J(\theta)|}{\pi(\theta|::) f_N(\omega^*|\omega_*, \Sigma_*)|J(\theta^*)|}, \quad (22) \right\}
\]
where $f_N(x|\mu, \Sigma)$ denotes the probability density function of a normal distribution with mean $\mu$ and covariance matrix $\Sigma$, and $J(\cdot)$ is the Jacobian for the
transformation, that is, $J(\theta) = |\frac{d\theta}{d\tilde{\theta}}| = \frac{2}{\pi(1-\rho^2)}$. The values of $(\theta, \theta^*)$ are evaluated at $(\omega, \omega^*)$, respectively.

**Step 4.** The conditional posterior probability density $\pi(\mu|\phi, \sigma, \beta, \nu, h, z, y)(\equiv \pi(\mu|\cdot))$ is given by

$$
\pi(\mu|\cdot) \propto \exp \left\{ -\frac{(\mu-\mu_0)^2}{2\sigma_0^2} - \frac{(\nu-\nu_0)^2}{2\sigma^2} \left( \frac{2}{1-\rho^2} - \frac{2}{2\sigma^2} \left( \frac{h_t+\mu-\phi(h_t+\mu)-\eta_t}{2\sigma^2(1-\rho^2)} \right)^2 \right) \right\}, \quad (23)
$$

from which we generate $\mu| \sim \mathcal{N}(\hat{\mu}, \sigma^2_\mu)$, where

$$
\begin{align*}
\sigma^2_\mu &= \left\{ \frac{1}{\sigma_0^2} + \frac{(1-\rho^2)(1-\phi^2) + (T-1)(1-\phi^2)}{\sigma^2(1-\rho^2)} \right\}^{-1}, \\
\hat{\mu} &= \sigma^2_\mu \left\{ \frac{\mu_0}{\sigma_0^2} + \frac{(1-\rho^2)(1-\phi^2)h_t + \sigma^2(1-\rho^2)}{(1-\rho^2)} \right\}.
\end{align*}
$$

**Step 5.** The posterior probability density $\pi(\beta|\phi, \sigma, \beta, \nu, h, z, y)(\equiv \pi(\beta|\cdot))$ is given by

$$
\pi(\beta|\cdot) \propto \exp \left\{ -\frac{(\beta-\beta_0)^2}{2\sigma^2_\beta} - \sum_{t=1}^{T-1} \left( \frac{h_t e^{\frac{h_t}{2}} - \beta_0 e^{\frac{h_t}{2}}}{2z_t e^{\frac{h_t}{2}}} \right)^2 \right\}, \quad (26)
$$

from which we generate $\beta| \sim \mathcal{N}(\hat{\beta}, \sigma^2_\beta)$ where

$$
\begin{align*}
\sigma^2_\beta &= \left\{ \frac{1}{\sigma_0^2} + \frac{1}{1-\rho^2} \sum_{t=1}^{T-1} \frac{z_t^2}{z_t^2 + z_n^2} \right\}, \\
\hat{\beta} &= \sigma^2_\beta \left\{ \frac{\beta_0}{\sigma_0^2} + \frac{1}{1-\rho^2} \sum_{t=1}^{T-1} \frac{z_t e^{\frac{h_t}{2}}}{z_t e^{\frac{h_t}{2}} + z_n e^{\frac{h_n}{2}}} \right\}.
\end{align*}
$$

**Step 6.** Because, as in Step 3, it is not easy to sample directly from the posterior probability density of $\nu$,

$$
\pi(\nu|\cdot) \propto \pi(\nu|\cdot) \prod_{t=1}^{T} \left( \frac{\nu/2}{\Gamma(\nu/2)} \right)^{\nu/2} z_t^{-\nu/2} \exp \left( \frac{\nu}{2z_t} \right) \left\{ -\sum_{t=1}^{T} \left( \frac{h_t e^{\frac{h_t}{2}}}{2z_t e^{\frac{h_t}{2}}} \right)^2 \right\}, \quad (29)
$$

for $\nu > 4$. The filter draws a sample of $\nu$ using the MH algorithm based on the normal approximation of the posterior probability density. We generate a candidate $\nu^*$ using a normal distribution truncated on $(4, \infty)$. 


Step 7. The conditional posterior probability density of the latent variable $z_t$ is

$$\pi(z_t | \theta, h, y) \propto g(z_t) \times z_t^{-(\nu + 1)/2} \exp\left(-\frac{\nu}{2z_t}\right),$$

(30)

$$g(z_t) = \exp\left\{\frac{-\left(y_t - \beta \pi_t e^{h_t/2}\right)^2}{2\sigma^2(1 - \rho^2)} - \frac{(\bar{h}_{t+1} - \bar{h}_h - \bar{y}_t)^2}{2\sigma^2(1 - \rho^2)} 1(t < n)\right\}.$$  

(31)

where $1(\cdot)$ is an indicator function. Using the MH algorithm, we generate a candidate $z^*_t \sim IG\left(\frac{\nu + 1}{2}, \frac{\nu}{2}\right)$ and accept it with probability \( \min\{g(z^*_t)/g(z_t), 1\} \).

**Generation of volatility latent variable $h$ (step 8)**

Step 8. An efficient strategy is to sample from the conditional posterior distribution of $h = \{h_t\}_{t=1}^n$ by dividing it into several blocks and sampling each block given the other blocks. This idea, called the block sampler or multi-move sampler, is developed by Shephard and Pitt (1997), and Watanabe and Omori (2004) in the context of state space modeling. They show that the sampler can produce efficient draws from the target conditional posterior distribution in comparison with a single-move sampler which primitively samples one state, say $h_t$, at a time given the others, $h_s(s \neq t)$. For the SV model with leverage, Omori and Watanabe (2008) developed the associated multi-move sampler and showed that it produces efficient samples (see also Takahashi et al. (2009) and Ishihara and Omori (2012)). We use their method for sampling $h_t$ in the SVSKt model.

### 3 Empirical Applications

#### 3.1 The Data

For Bayesian estimation of stochastic volatility model with leverage and asymmetric heavy tails using generalized hyperbolic skew Student’s t-error distribution, we consider five stock markets of Latin America: Peru, Argentina, Mexico, Chile and Brazil. The Latin American stock market indexes are in Table 1. We use a sample from 1996/1/2 to 2013/12/30 for all stock markets of Latin America except to Peru where the period is from 2001/1/2 to 2013/12/30 because there was a change in the methodology of IGBVL index in November 1998 and this could affect the results.

<table>
<thead>
<tr>
<th>Index</th>
<th>Country</th>
<th>Period</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGBVL</td>
<td>Peru</td>
<td>2001/1/2 - 2013/12/30</td>
</tr>
<tr>
<td>MERVAL</td>
<td>Argentina</td>
<td>1996/1/2 - 2013/12/30</td>
</tr>
<tr>
<td>MEXBOL</td>
<td>Mexico</td>
<td>1996/1/2 - 2013/12/30</td>
</tr>
<tr>
<td>IPSA</td>
<td>Chile</td>
<td>1996/1/2 - 2013/12/30</td>
</tr>
<tr>
<td>IBOVESPA</td>
<td>Brazil</td>
<td>1996/1/2 - 2013/12/30</td>
</tr>
</tbody>
</table>
In our application, we also analyze the U.S. S&P500 index from 1996/1/2 to 2013/12/30 to compare the results of literature with Latin American indexes. One reason is that the U.S. stock market could be considered as a good benchmark. Stock daily returns are computed as the log difference $y_t = \log P_t - \log P_{t-1}$, where $P_t$ is the closing stock price of day $t$. The data were obtained from Bloomberg and the sample size differs between countries because holidays and closed days of stock markets. Table 2 shows the number of observations and some descriptive statistics and Figure 1 shows the time series plots of the daily stock returns.

<table>
<thead>
<tr>
<th>Index</th>
<th>Obs.</th>
<th>Mean</th>
<th>Std. deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>IGBVL</td>
<td>3246</td>
<td>0.0008</td>
<td>0.0149</td>
<td>-0.5287</td>
<td>10.8286</td>
<td>-0.1329</td>
<td>0.1282</td>
</tr>
<tr>
<td>Merval</td>
<td>4439</td>
<td>0.0005</td>
<td>0.0215</td>
<td>-0.2801</td>
<td>5.3395</td>
<td>-0.1476</td>
<td>0.1612</td>
</tr>
<tr>
<td>MEXBOL</td>
<td>4529</td>
<td>0.0006</td>
<td>0.0151</td>
<td>0.0300</td>
<td>6.9916</td>
<td>-0.1431</td>
<td>0.1215</td>
</tr>
<tr>
<td>IBOVESPA</td>
<td>4456</td>
<td>0.0006</td>
<td>0.0213</td>
<td>0.2994</td>
<td>13.1430</td>
<td>-0.1723</td>
<td>0.2882</td>
</tr>
<tr>
<td>IPSA</td>
<td>4489</td>
<td>0.0003</td>
<td>0.0111</td>
<td>0.1332</td>
<td>7.9881</td>
<td>-0.0767</td>
<td>0.1180</td>
</tr>
<tr>
<td>S&amp;P500</td>
<td>4531</td>
<td>0.0002</td>
<td>0.0127</td>
<td>-0.2272</td>
<td>7.4884</td>
<td>-0.0947</td>
<td>0.1096</td>
</tr>
</tbody>
</table>

IGBVL and Merval are negatively skewed while MEXBOL, IBOVESPA and IPSA are positively skewed. However, the skewness of MEXBOL is very close to zero. The IGBVL index is the most negatively skewed with $-0.5287$ and IBOVESPA is the most positive skewed with $0.2994$. Regarding the kurtosis, all the stock returns have positive kurtosis where IBOVESPA has the highest value $13.1430$. The S&P500 index also has negative skewness and positive kurtosis. The summary statistics show heavy tailedness for the five Latin American empirical returns distribution of the data and negative skewness for IGBVL and Merval.
Figure 1. Times series plots for IGBVL (2001/01/02 - 2013/12/30) and Merval, Ibovespa, Mexbol, Ipsa and S&P500 (1996/01/02 - 2013/12/30) daily returns.

3.2 Parameter Estimates

For parameter estimates of SVSKt model, we assume the same prior distributions as Nakajima and Omori (2012). These prior distributions are commonly used in the literature, see Kim et al. (1998), Nakajima and Omori (2009), Omori et al. (2007), Trojan (2013):

\[
\frac{\phi + 1}{2} \sim \text{Beta}(20, 1.5), \quad \sigma^{-2} \sim \text{Gamma}(2.5, 0.025), \quad \rho \sim U(-1, 1),
\]
\[
\mu \sim N(-10, 1), \quad \beta \sim N(0, 1), \quad \nu \sim \text{Gamma}(16, 0.8) I(\nu > 4).
\]

The beta prior distribution for \(\frac{\phi + 1}{2}\) implies that the mean and the standard deviation are (0.86, 0.11) for \(\phi\). The means and the standard deviations of Gamma (2.5, 0.025) and Gamma (16, 0.8) are (100, 63.2) and (20, 5), respectively.

We draw 20 000 samples after discarding the initial 5 000 samples as a burn-in period for Merval, Mexbol, Ibovespa, Ipsa and S&P500 and 9 000 samples as a burn-in period for IGBVL, so that the effect of initial values
on the posterior inference is minimized. We compute the inefficiency factor to check the efficiency of the MCMC algorithm. The inefficiency factor is defined by $1 + 2\sum_{s=1}^{\infty}\rho_s$ where $\rho_s$ is the sample autocorrelation at lag $s$. It measures how well the MCMC chain mixes (Chib, 2001; Nakajima and Omori (2009) and Nakajima and Omori (2012)). It is the estimated ratio of the numerical variance of the posterior sample mean to the variance of the sample mean from uncorrelated draws. When the inefficiency factor is equal to $m$, we need to draw MCMC samples $m$ times as many as uncorrelated samples. We compute the inefficiency factor using a Parzen window with bandwidth $b_w = 1000$.

Figures 2 - 7 show the MCMC estimation results of the SVSKt model for the IGBVL, MERVAL, MEXBOL, IBOVESPA, IPSA and S&P500 indexes. Regarding Latin American stock indexes, the sample paths appear to be stable and the proposed estimation scheme works well for MERVAL, MEXBOL, IBOVESPA and IPSA. In these cases, the autocorrelation over the iterations is decaying and there are convergence of the Markov chains of the parameters. Regarding IGBVL, we obtain poor mixing (or slow convergence) of the Markov chain of some parameters ($\phi$, $\sigma$ and $\mu$) and estimation the results show high autocorrelation through iterations of $\phi$, $\sigma$ and $\mu$ with a slowly decay. Regarding S&P500, we obtain similar results as Nakajima and Omori (2012).

![Figure 2](image)

**Figure 2.** MCMC estimation results of the SVSKt model for IGBVL data (Peru). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).
Figure 3. MCMC estimation results of the SVSKt model for MERVAL data (Argentina). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Figure 4. MCMC estimation results of the SVSKt model for MEXBOL data (Mexico). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).
Figure 5. MCMC estimation results of the SVSkT model for IBOVESPA data (Brazil). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Figure 6. MCMC estimation results of the SVSkT model for IPSA data (Chile). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).
Figure 7. MCMC estimation results of the SVSKt model for S&P500 data (US).

Table 2 shows the estimation results of the posterior estimates: the posterior means, the standard deviation, the 95% credible intervals and the inefficiency factors for the stock returns data. The posterior means of $\phi$, that measure persistence of the log-volatility, are close to one (in the range of 0.9535 to 0.9711) for MERVAL, MEXBOL, IBOVESPA and IPSA. These results are consistent with literature that indicate the high persistence of the volatility in stock returns. IBOVESPA and MEXBOL are more persistent, followed by IPSA and MERVAL. Regarding the IGBVL returns have a posterior mean of $\phi = 0.8618$, which indicates a low persistence in comparison to the volatility of the others indexes above mentioned.

The posterior means of $\rho$, that measures the correlations between $\epsilon_t$ and $\eta_t$, are estimated to be negative for all indexes considered. When $\rho$ values are negative imply that there exist leverage effects. MEXBOL and IBOVESPA have in absolute value the highest posterior mean estimates of $\rho$ ($-0.3948$ and $-0.3464$, respectively), which imply that the leverage effect is more notable in these indexes. Also the 95% credible intervals are negative implying that the posterior probability that $\rho$ is negative is greater than 0.95, and the negativity of $\rho$ is credible. The same applies with MERVAL and IPSA (posterior mean estimates of $\rho$ are $-0.2977$ and $-0.2970$, respectively) where the 95% credible intervals are negative, but it is a minor leverage effect than the previous indexes. In the case of IGBVL, the posterior mean estimate of $\rho$ is also negative although very close to zero and the 95% credible intervals contain zero and positive values.
This implies that the posterior distribution of $\rho$, although mainly located in the negative range as shown in the Figure 2, can take positive values or even zero, which would imply the non-existence of the leverage effect in IGBVL returns. These results support the evidence of the leverage effects in Latin American stock returns data.

Regarding the variance $\sigma^2$, the posterior mean estimates of $\sigma$ show that all indexes have similar estimates in the range from 0.1969 to 0.2717 with the exception of the IGBVL returns, where the posterior mean estimate of $\sigma$ takes a very high value (0.9173) compared to the other indexes. This implies that the variance of the shock $\eta_t$ is large and the log-volatility has more variability than the other stock indexes in Latin American. Regarding the posterior mean of $\mu$, all indexes show similar results in the range of $-19.4626$ to $-8.2422$. 
Table 3: MCMC estimation results of the SVSKt model for Latin American stock return data

(i) IGBVL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.8618</td>
<td>0.0213</td>
<td>[0.8197, 0.9011]</td>
<td>417.43</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.9173</td>
<td>0.0682</td>
<td>[0.7602, 1.0380]</td>
<td>539.46</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.0475</td>
<td>0.0370</td>
<td>[-0.1197, 0.0245]</td>
<td>59.91</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.8002</td>
<td>0.1686</td>
<td>[-9.0995, -8.4321]</td>
<td>233.32</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.0286</td>
<td>0.1508</td>
<td>[-0.3247, 0.2678]</td>
<td>16.87</td>
</tr>
<tr>
<td>$\nu$</td>
<td>35.6892</td>
<td>5.4920</td>
<td>[25.9746, 47.2584]</td>
<td>88.54</td>
</tr>
</tbody>
</table>

(ii) MERVAL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9535</td>
<td>0.0086</td>
<td>[0.9347, 0.9681]</td>
<td>129.12</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2717</td>
<td>0.0269</td>
<td>[0.2264, 0.3316]</td>
<td>226.88</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.2977</td>
<td>0.0436</td>
<td>[-0.3807, -0.2106]</td>
<td>35.47</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.2705</td>
<td>0.0951</td>
<td>[-8.4565, -8.0816]</td>
<td>26.19</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.2464</td>
<td>0.0811</td>
<td>[-0.4180, -0.0952]</td>
<td>82.38</td>
</tr>
<tr>
<td>$\nu$</td>
<td>12.2135</td>
<td>1.7827</td>
<td>[9.0445, 15.9292]</td>
<td>286.25</td>
</tr>
</tbody>
</table>

(iii) MEXBOL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9683</td>
<td>0.0057</td>
<td>[0.9562, 0.9785]</td>
<td>90.33</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2362</td>
<td>0.0206</td>
<td>[0.2015, 0.2847]</td>
<td>176.82</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.3948</td>
<td>0.0444</td>
<td>[-0.4769, -0.3014]</td>
<td>81.93</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.9186</td>
<td>0.1134</td>
<td>[-9.1434, -8.6942]</td>
<td>30.52</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.1076</td>
<td>0.1194</td>
<td>[-0.3389, 0.1326]</td>
<td>37.20</td>
</tr>
<tr>
<td>$\nu$</td>
<td>19.8882</td>
<td>3.2235</td>
<td>[14.2966, 27.2999]</td>
<td>265.91</td>
</tr>
</tbody>
</table>

(iv) IBOVESPA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9711</td>
<td>0.0053</td>
<td>[0.9601, 0.9806]</td>
<td>67.00</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.1969</td>
<td>0.0168</td>
<td>[0.1696, 0.2358]</td>
<td>246.63</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.3464</td>
<td>0.0452</td>
<td>[-0.4333, -0.2551]</td>
<td>37.82</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.2422</td>
<td>0.1057</td>
<td>[-8.4519, -8.0325]</td>
<td>21.19</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.0342</td>
<td>0.1107</td>
<td>[-0.2437, 0.1954]</td>
<td>38.07</td>
</tr>
<tr>
<td>$\nu$</td>
<td>17.5026</td>
<td>2.6309</td>
<td>[13.1482, 23.3003]</td>
<td>159.85</td>
</tr>
</tbody>
</table>

(v) IPSA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9653</td>
<td>0.0062</td>
<td>[0.9520, 0.9763]</td>
<td>82.62</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2226</td>
<td>0.0196</td>
<td>[0.1878, 0.2656]</td>
<td>179.16</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.2970</td>
<td>0.0440</td>
<td>[-0.3835, -0.2120]</td>
<td>26.06</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-9.4626</td>
<td>0.0995</td>
<td>[-9.6562, -9.2626]</td>
<td>4.26</td>
</tr>
<tr>
<td>$\beta$</td>
<td>-0.0852</td>
<td>0.1872</td>
<td>[-0.4475, 0.2858]</td>
<td>37.72</td>
</tr>
<tr>
<td>$\nu$</td>
<td>30.2523</td>
<td>5.1130</td>
<td>[21.1299, 41.1597]</td>
<td>166.06</td>
</tr>
</tbody>
</table>

As mentioned previously, the skewness and the heavy tailedness of the GH Skew Student’s t-Distribution are determined jointly by the combination of the parameter values of $\beta$ and $\nu$. A lower value of $\beta$ implies a more negative skewness or left-skewness as well as heavier tails. On the other hand, as $\nu$ becomes larger
the density becomes less skewed and has lighter tails. The posterior means of $\beta$ are estimated to be negative for all index returns considered. MERVAL has the less value of posterior mean estimate of $\beta$ with $-0.2464$ and the 95% credible intervals are negative implying that the posterior probability that $\beta$ is negative is greater than 0.95, and the negativity of $\beta$ is credible. These results support the evidence of the skewness, however the posterior mean estimates of $\beta$ for IGBVL, MEXBOL, IBOVESPA and IPSA are also negative but the 95% credible intervals contain zero and positive values. We know that when $\beta = 0$ in the SVSKt model correspond to symmetric student’s $t$-density. The estimates of $\beta$ very close to zero for IGBVL, IBOVESPA and IPSA index returns could imply the case of symmetric heavy tailed errors. Finally, the posterior means of $\nu$ are around 35.6892 for IGBVL, 12.2135 for MERVAL, 19.8882 for MEXBOL, 17.5026 for IBOVESPA and 30.2523 for IPSA returns.

Regarding S&P500 daily returns, Table 4 shows the estimation results of the posterior estimates. These results are very similar to Nakajima and Omori (2012). The posterior mean of $\phi$ are close to one (0.9703) and imply high persistence, more than Latin American stock returns considered. The posterior mean of $\rho$ is estimated to be negative (--0.6864) which imply the evidence of the leverage effect in the S&P500 index. Also the 95% credible intervals are negative implying that the posterior probability that $\rho$ is negative is greater than 0.95, and the negativity of $\rho$ is credible. Regarding the variance $\sigma^2$ of $\eta_t$, the posterior mean estimates of $\sigma$ is 0.2382, similar parameter estimates to MERVAL, MEXBOL, IPSA and IBOVESPA indexes. The posterior mean of $\mu$ is $-9.4186$. The posterior means of $\beta$, that measures the skewness, are estimated to be negative ($-0.7842$). Also the 95% credible intervals are negative implying that the posterior probability that $\beta$ is negative is greater than 0.95, and the negativity of $\beta$ is credible. These results support the evidence of the skewness. Finally, the posterior means of $\nu$ are around 24.8685, which indicates a heavy tailedness in the stock returns distribution.

Table 4: Estimation results of the SVSKt model for US S&P500 stock return data

<table>
<thead>
<tr>
<th>S&amp;P500</th>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.9703</td>
<td>$\phi$</td>
<td>0.0040</td>
<td></td>
<td>[0.9622, 0.9778]</td>
<td>40.86</td>
</tr>
<tr>
<td>0.2382</td>
<td>$\sigma$</td>
<td>0.0135</td>
<td></td>
<td>[0.2125, 0.2652]</td>
<td>89.75</td>
</tr>
<tr>
<td>-0.6864</td>
<td>$\rho$</td>
<td>0.0323</td>
<td></td>
<td>[-0.7449, -0.6199]</td>
<td>76.70</td>
</tr>
<tr>
<td>-9.4186</td>
<td>$\mu$</td>
<td>0.1001</td>
<td></td>
<td>[-9.6135, -9.2207]</td>
<td>7.79</td>
</tr>
<tr>
<td>-0.7842</td>
<td>$\beta$</td>
<td>0.1899</td>
<td></td>
<td>[-1.1837, -0.4380]</td>
<td>116.81</td>
</tr>
<tr>
<td>24.8685</td>
<td>$\nu$</td>
<td>3.8282</td>
<td></td>
<td>[18.2556, 33.5375]</td>
<td>189.21</td>
</tr>
</tbody>
</table>

The indicator of how well MCMC chain mixes is measured by the inefficiency factors of the MCMC algorithm defined by $1 + 2\sum_{s=1}^{\infty}p_s$ where $p_s$ is the sample autocorrelation at lag $s$. The inefficiency factor of IGBVL has high values for $\phi$, $\sigma$ and $\mu$ for IGBVL. These results is supporting by the initial MCMC estimation
that show high autocorrelation through iterations of parameters $\phi$, $\sigma$ and $\mu$ of IGBVL that decay slowly (see Figure 2). MERVAL, MEXBOL, IBOVESPA and IPSA returns show low values of inefficiency factor in all parameters estimates but parameter $\sigma$ and $\nu$ have higher values of inefficiency factor compared with the other parameters. In general, the inefficiency factor for the parameters of the S&P500 returns have low values.

The Figure 8 shows the log-volatility estimates for the Latin American stock returns considered. The results show that there is a similar pattern between periods of higher volatility between the five Latin American stock returns considered. Most of times, these periods of high volatility are associated with international crisis. For example, all indexes considered had a rise in log-volatility in the period from August to November 1998 due to the Asian crisis, which cause an contagion effect. Also, all stock returns show a considerable increase in the level of log-volatility for the period September - October 2008 associated with the outbreak of the international financial crisis. Another example is in July and September 2011 by the Europe crisis.

![Figure 8. Log-volatility for IGBVL (2001/01/02 - 2013/12/30) and MERVAL, MEXBOL, IPSA, IBOVESPA and S&P500 (1996/01/02 - 2013/12/30) daily returns.](image)

### 3.3 Model Comparison

In this subsection, model comparisons between competing models for the daily stock returns are provided. We make a comparison between the SVSKt model with the SV model with symmetric Student’s error distribution ($\beta = 0$).

In Bayesian framework, we can compare the fit of several competing models using their posterior probabilities to select the one that is the best supported.
by the data. The idea is that the model that yields the largest posterior probability has the best fit between competing models. The posterior probability of each model is proportional to the product of prior probability of the model and the marginal likelihood. The ratio of two posterior probabilities is also known as a Bayes factor. If the prior probabilities are assumed to be equal between competing models, we choose the model that yields the largest marginal likelihood.

The marginal likelihood is defined as the integral of the likelihood with respect to the prior density of the parameter. The approach adopted here for the computation of the marginal likelihood relies on the method developed by Chib (1995) as well as Kim et al. (1998), Nakajima (2012) and Nakajima and Omori (2012) use the same approach. We estimate the log of marginal likelihood of a model, denoted by \( \log m(y) \), as

\[
\log m(y) = \log f(y|\Theta) + \log \pi(\Theta) - \log \pi(\Theta|y),
\]  

where \( \Theta \) is a parameter set in the model, \( f(y|\Theta) \) is a likelihood of the model, \( \pi(\Theta) \) is a prior probability density and \( \pi(\Theta|y) \) is a posterior probability density. The equality holds for any values of \( \Theta \), but we usually use the posterior mean of \( \Theta \) to obtain a stable estimate of \( m(y) \). The prior probability density can be easily calculated, although the likelihood and posterior part must be evaluated by simulation (Nakajima and Omori, 2012). The likelihood \( f(y|\Theta) \) can be estimated by the particle filter (e.g. Pitt and Shepard, 1999; Chib et al., 2002; Omori et al., 2007). For the posterior probability density, we use the method of Chib (1995) and Chib and Jeliazkov (2001) to compute \( \pi(\Theta|y) \) using the sample obtained through the reduced iteration of the MCMC algorithm. We evaluated these densities at the posterior mean for \( \Theta \).

In order to compare the competing models, we estimate the marginal likelihood as follows: (i) The likelihood is going to be estimated using the auxiliary particle filter with 10000 particles. It is replicated 10 times to obtain the standard error of the likelihood estimate as in Nakajima and Omori (2012) and (ii) The posterior probability density at \( \Theta \) is evaluated through additional but reduced MCMC runs. We use 5000 iterations for the reduced runs.

Figures 9 - 13 show the MCMC estimation results of the SVt model for IGBVL, Merval, MEXBOL, IBOVESPA and IPSA returns.
Figure 9. MCMC estimation results of the SVt model for IGBVL data (Peru). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Figure 10. MCMC estimation results of the SVt model for MERVAL data (Argentina). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).
Figure 11. MCMC estimation results of the SVt model for MEXBOL data (Mexico). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Figure 12. MCMC estimation results of the SVt model for IBOVESPA data (Brazil). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).
Figure 13. MCMC estimation results of the SVt model for IPSA data (Chile). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Table 5 shows the MCMC estimation results of the posterior estimates of the SVt model: the posterior means, the standard deviation, the 95% credible intervals and the inefficiency factors for the IGBVL, MERVAL, MEXBOL, IPSA and IBOVESPA returns data. The posterior means of estimates parameter are very similar to the SVSKt model.
Table 5: MCMC estimation results of the SVt model for Latin American stock return data

(i) ICBVL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.8613</td>
<td>0.0209</td>
<td>[0.8183, 0.9024]</td>
<td>407.89</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.9823</td>
<td>0.0964</td>
<td>[0.7655, 1.1565]</td>
<td>634.21</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.0382</td>
<td>0.0386</td>
<td>[-0.1122, 0.0401]</td>
<td>104.68</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.7154</td>
<td>0.1639</td>
<td>[-9.0253, -8.3828]</td>
<td>156.46</td>
</tr>
<tr>
<td>$\nu$</td>
<td>36.1646</td>
<td>5.7485</td>
<td>[26.2375, 48.8401]</td>
<td>105.50</td>
</tr>
</tbody>
</table>

(ii) MERVAL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9533</td>
<td>0.0081</td>
<td>[0.9358, 0.9674]</td>
<td>103.72</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2707</td>
<td>0.0244</td>
<td>[0.2279, 0.3256]</td>
<td>197.98</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.2810</td>
<td>0.0434</td>
<td>[-0.3652, -0.1945]</td>
<td>52.60</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.2351</td>
<td>0.0948</td>
<td>[-8.4186, -8.0462]</td>
<td>20.95</td>
</tr>
<tr>
<td>$\nu$</td>
<td>12.3573</td>
<td>1.9107</td>
<td>[9.2705, 16.8398]</td>
<td>253.81</td>
</tr>
</tbody>
</table>

(iii) MEXBOL

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9694</td>
<td>0.0054</td>
<td>[0.9584, 0.9788]</td>
<td>74.32</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2282</td>
<td>0.0178</td>
<td>[0.1973, 0.2661]</td>
<td>160.80</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.4037</td>
<td>0.0471</td>
<td>[-0.4951, -0.3126]</td>
<td>94.09</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.9333</td>
<td>0.1122</td>
<td>[-9.1537, -8.7119]</td>
<td>17.26</td>
</tr>
<tr>
<td>$\nu$</td>
<td>17.2837</td>
<td>3.2351</td>
<td>[12.0182, 24.5254]</td>
<td>280.95</td>
</tr>
</tbody>
</table>

(iv) IBOVESPA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9564</td>
<td>0.0067</td>
<td>[0.9423, 0.9687]</td>
<td>59.67</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2528</td>
<td>0.0176</td>
<td>[0.2219, 0.2912]</td>
<td>169.99</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.3244</td>
<td>0.0447</td>
<td>[-0.4091, -0.2330]</td>
<td>46.67</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-8.2198</td>
<td>0.0907</td>
<td>[-8.3970, -8.0400]</td>
<td>11.09</td>
</tr>
</tbody>
</table>

(v) IPSA

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>95% interval</th>
<th>Inefficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\phi$</td>
<td>0.9649</td>
<td>0.0061</td>
<td>[0.9521, 0.9759]</td>
<td>85.17</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.2253</td>
<td>0.0205</td>
<td>[0.1899, 0.2718]</td>
<td>283.75</td>
</tr>
<tr>
<td>$\rho$</td>
<td>-0.2928</td>
<td>0.0438</td>
<td>[-0.3754, -0.2059]</td>
<td>43.86</td>
</tr>
<tr>
<td>$\mu$</td>
<td>-9.4536</td>
<td>0.1004</td>
<td>[-9.6466, -9.2478]</td>
<td>23.10</td>
</tr>
<tr>
<td>$\nu$</td>
<td>29.8385</td>
<td>4.8740</td>
<td>[21.3740, 40.5813]</td>
<td>165.48</td>
</tr>
</tbody>
</table>

Table 6 shows the logarithm of the estimated marginal likelihoods and their standard errors. The SVSKt model outperforms for all datasets regardless of the sample periods except to MEXBOL where the estimated marginal likelihoods of the SVt model is larger than the SVSKt. We can see that GH-skew Student’s $t$-error distribution in SV model is clearly successful in describing the distribution of the daily stock return data of the Latin American indexes considered.
### Table 6: Estimated log of marginal likelihoods (Log-ML)

<table>
<thead>
<tr>
<th>Model</th>
<th>IGBVL</th>
<th>MERVAL</th>
<th>MEXBOL</th>
<th>IBOVESPA</th>
<th>IPSA</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVSKt</td>
<td>9812.000</td>
<td>11469.963</td>
<td>13377.385</td>
<td>11624.206</td>
<td>14583.355</td>
</tr>
<tr>
<td></td>
<td>(1.582)</td>
<td>(0.806)</td>
<td>(0.580)</td>
<td>(0.996)</td>
<td>(0.557)</td>
</tr>
<tr>
<td>SVt</td>
<td>9781.776</td>
<td>11464.903</td>
<td>13380.232</td>
<td>11614.349</td>
<td>14586.065</td>
</tr>
<tr>
<td></td>
<td>(1.557)</td>
<td>(0.684)</td>
<td>(0.704)</td>
<td>(0.627)</td>
<td>(0.653)</td>
</tr>
</tbody>
</table>

*Standard errors of the log-ML in parentheses.

## 4 Conclusions

This paper uses a Bayesian estimation of SV model with leverage and GH skew Student’s t-errors proposed by Nakajima and Omori (2012) to assess the asymmetrically heavy-tailed distributions of stock returns for IGBVL, MERVAL, MEXBOL, IPSA, and S&P500 daily returns. The SVSKt model is a general model that contains the SVt model with symmetric Student’s t error distribution and the standard SV model with normal error distribution.

The MCMC estimation results of the SVSKt model with the algorithm proposed by Nakajima and Omori (2012) show that the sample paths of the iterations of parameters are stable, and the proposed estimation scheme works well for all indexes except for IGBVL where the Markov chains do not converge and there is high autocorrelation between iterations. The posterior mean parameter estimates are consistent with literature that indicate the high persistence of the volatility in stock returns. Also the results show that the IGBVL returns have low persistence in comparison to the volatility of the others indexes of Latin American considered.

The results support the evidence of the leverage effects in Latin American stock returns data except to IGBVL returns. The estimates show that the leverage effect is more notable in MEXBOL and IBOVESPA indexes, followed by MERVAL and IPSA. In the case of IGBVL, the posterior mean estimate of $\rho$ is negative but very close to zero, which would imply the non-existence of the leverage effect in IGBVL returns. Another important result is that the log-volatility of IGBVL returns have more variability than the other stock returns in Latin American considered. Finally, the results show a similar pattern of volatility between the indexed considered and show evidence of skewness and heavy-tailedness for Latin American daily returns. The posterior means of $\beta$ are estimated to be negative for all index returns considered.

Finally, model comparisons between SVSKt and SVt models for the daily stock returns are provided. The results show that the SVSKt model outperforms for all datasets regardless of the sample periods except to MEXBOL where the estimated marginal likelihoods of the SVt model is larger than the SVSKt.
Appendix

A. The GH Skew Student- $t$ Distribution

This appendix includes some important properties of the Generalized Hyperbolic skew Student-t distribution. For a more complete treatment, see Aas and Haff (2006) and Prause (1999). The GH skew-t distribution is a limiting case of the GH distribution, which was introduced in Barndorff-Nielsen (1977). Starting with the latter, it can be parameterized in the following way,

$$f_{GH}(x) = \frac{(\beta^2 - \gamma^2)^{\lambda/2}K_{\lambda-1/2}(\beta\sqrt{\delta^2 + (x-\mu)^2}) \exp(\gamma(x-\mu))}{\sqrt{2\pi}\beta^{1-1/2}\delta^\lambda K_{\lambda}(\delta\sqrt{\beta^2 + \gamma^2})(\delta + (x-\mu)^2)^{1/2-\lambda}}$$ \hspace{1cm} (A.1)

where $K_i$ is the modified Bessel function of the third kind of order $i$ and $\omega \in \mathbb{R}$. The parameters must satisfy the conditions:

$$\delta \geq 0, |\gamma| < \beta \text{ if } \lambda > 0,$$
$$\delta > 0, |\gamma| < \beta \text{ if } \lambda = 0,$$
$$\delta > 0, |\gamma| \leq \beta \text{ if } \lambda < 0.$$

The tails of the GH distribution behave as

$$f_{GH}(x) \sim \text{const} |x|^{\lambda-1} \exp(-\beta|x| + \gamma x) \text{ as } x \to \pm\infty \forall \lambda.$$ \hspace{1cm} (A.3)

It follows that as long as $|\gamma| \neq \beta$, both tails are semi-heavy.

The GH distribution can be represented as a normal mean-variance mixture with Generalized Inverse Gaussian (GIG) distributed mixture variable

$$X = \mu + \gamma Z + \sqrt{\psi} Y, \quad Y \sim N(0,1), \quad Z \sim \text{GIG}(\lambda, \delta, \psi)$$ \hspace{1cm} (A.4)

where $\psi = \sqrt{\beta^2 - \gamma^2}$. It follows from equation (A.4) that $X | Z = z \sim N(\mu + \gamma z, z)$. However, parameters are difficult to estimate due to the flatness of the likelihood function (see e.g. Aas and Haff, 2006). A parsimonious representation that is more amenable to estimation can be obtained by letting $\lambda = -v/2 (v > 0), \delta = \sqrt{v}$ and $\beta \to |\gamma|$ in equation (A.1) ($\psi = 0$). Then $Z \sim \text{GIG}(-v/2, \sqrt{v}, 0)$ or equivalently $Z \sim \text{IG}(v/2, v/2)$. This limiting case of the GH distribution is referred to as the GH skew-t distribution, with skew parameter $\gamma$, degrees of freedom parameter $v$ and mixture representation

$$X = \mu + \gamma Z + \sqrt{Z} Y, \quad Y \sim N(0,1), \quad Z \sim \text{IG}(v/2, v/2).$$ \hspace{1cm} (A.5)
Its closed form density is given by

\[
\begin{align*}
    f_{\text{GH skewt}}(x) &= \frac{2^{\frac{1-v}{2}} v^{\frac{3}{2}} \gamma^{\frac{\gamma+1}{2}} \frac{K_{\frac{\gamma+1}{2}}(\sqrt{\gamma^2 v + x^2})}{\Gamma(\frac{\gamma}{2}) \sqrt{\gamma^2 v + x^2}}}{\gamma \exp(\gamma(x - \mu))}, \quad \gamma \neq 0, \\
    \text{and} \\
    f_{\text{GH}}(x) &= \frac{\Gamma(\frac{\gamma+1}{2})}{\sqrt{\pi} \Gamma(\frac{\gamma}{2})} \left[ 1 + \frac{(x - \mu)^2}{v} \right]^{-(\gamma+1)/2}, \quad \gamma = 0.
\end{align*}
\]  

(A.6)

The latter density is commonly referred to as the non-central Student-\( t \) distribution with \( v \) degrees of freedom, expectation \( \mu \) and variance \( v/(v-2) \).

The first four moments of a GH skew-\( t \) distributed random variate \( X \) are (Aas and Haff, 2006):

\[
\begin{align*}
    E(X) &= \mu + \frac{\gamma v}{v - 2}, \\
    \text{Var}(X) &= \frac{2\gamma^2 v^2}{(v-2)(v-4)} + \frac{v}{v-2}, \\
    \text{Skewness}(X) &= \frac{2\gamma \left[v(v-4)\right]^{1/2}}{[2\gamma^2 v + (v-2)(v-4)]^{3/2}} \left[ 3(v-2) + \frac{8\gamma^2 v}{v-6} \right], \\
    \text{Kurtosis}(X) &= \frac{6}{[2\gamma^2 v + (v-2)(v-4)]^2} \left[ (v-2)^2(v-4) + \frac{16\gamma^2 v(v-2)(v-4)}{v-6} + \frac{8\gamma^4 v^2(5v-22)}{(v-6)(v-8)} \right].
\end{align*}
\]

(A.8) \quad (A.9) \quad (A.10) \quad (A.11)

We observe that for the mean and variance to exist, \( v > 2 \) and \( v > 4 \), respectively. Skewness and (excess) kurtosis are defined only if \( v > 6 \) and \( v > 8 \), respectively.

From equation (A.3) it follows that the tails of the GH skew-\( t \) distribution are given by

\[
    f_{\text{GH skewt}}(x) \sim \text{const} |x|^{-v/2-1} \exp(-|\gamma x| + \gamma x) \text{ as } x \to \pm \infty.
\]

(A.12)

Thus we have a heavy tail decaying as

\[
    f_{\text{GH skewt}}(x) \sim \text{const} |x|^{-v/2-1} \quad \text{if} \quad \begin{cases} 
    \gamma < 0 \text{ and } x \to -\infty \\
    \gamma > 0 \text{ and } x \to +\infty
\end{cases}
\]

(A.13)

and a light tail as

\[
    f_{\text{GH skewt}}(x) \sim \text{const} |x|^{-v/2-1} \exp(-2|\gamma x|) \quad \text{if} \quad \begin{cases} 
    \gamma < 0 \text{ and } x \to +\infty \\
    \gamma > 0 \text{ and } x \to -\infty
\end{cases}
\]

(A.14)
The heavy tail shows polynomial and the light tail exponential behavior. Thus the GH skew-$t$ distribution is able to model substantially skewed and heavy tailed data, as found for example in financial markets. The tails of the GH skew-$t$ distribution are characterized uniquely by parameters $\gamma$ and $\nu$, which determine jointly the degree of skewness and heavy tailedness. Finally, note that the heavy tail of the GH skew-$t$ distribution is heavier than the tails of the symmetric Student-$t$ distribution, which decay as

$$f_{GHt}(x) \sim \text{const} \cdot |x|^{-\nu-1} \text{ as } x \to \pm \infty.$$ 

References


