PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO

"An Empirical Application of Stochastic Volatility Models to Latin-American Stock Returns using GH Skew Student´s *t*-Distribution"

Tesis para optar el grado de Magíster en Estadística

AUTOR

Patricia Lengua Lafosse

ASESOR

Cristian Bayes

JURADO

Luis Valdivieso

Oscar Millones

LIMA – PERÚ

2015

An Empirical Application of Stochastic Volatility Models to Latin-American Stock Returns using *GH* Skew Student's t-Distribution*

Patricia Lengua Lafosse†
Pontificia Universidad Católica del Perú

Thesis Advisor: Cristian Bayes‡
Pontificia Universidad Católica del Perú

March 12, 2015

Abstract

This paper represents empirical studies of stochastic volatility (SV) models for daily stocks returns data of a set of Latin American countries (Argentina, Brazil, Chile, Mexico and Peru) for the sample period 1996:01-2013:12. We estimate SV models incorporating both leverage effects and skewed heavy-tailed disturbances taking into account the *GH* Skew Student's t-distribution using the Bayesian estimation method proposed by Nakajima and Omori (2012). A model comparison between the competing SV models with symmetric Student´s t-disturbances is provided using the log marginal likelihoods in the empirical study. A prior sensitivity analysis is also provided. The results suggest that there are leverage effects in all indices considered but there is not enough evidence for Peru, and skewed heavy-tailed disturbances is confirmed only for Argentina, symmetric heavy-tailed disturbances for Mexico, Brazil and Chile, and symmetric Normal disturbances for Peru. Furthermore, we find that the *GH* Skew Student's t-disturbance distribution in the SV model is successful in describing the distribution of the daily stock return data for Peru, Argentina and Brazil over the traditional symmetric Student´s t-disturbance distribution.

JEL Classification: C11, C58.

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

*We are grateful to Gabriel Rodríguez for helpful comments and suggestions.
†patricia.lengua@pucp.pe (Author for Correspondence)
‡cbayes@pucp.edu.pe

1 Introduction

Returns from financial market variables such as stock and exchange rate are characterized by some empirical properties which are generally presented in financial time series. There are three important stylized facts or properties that are found in almost all set of daily returns: i) returns are not normally distributed; instead, the characteristics of the return distributions are excess of kurtosis (leptokurtic) and some degree of skewness compared with the Normal distribution[1], ii) there is almost no correlation between daily returns at different lags and iii) functions of returns can have substantial autocorrelations. For example, the autocorrelation of both absolute returns and squared returns are positive for many lags and statistically significative (Taylor, 2005). These properties are explained in most cases by the presence of time-varying volatility and volatility clustering over time.

Modelling time-varying volatility has been widely used in the literature of financial time series, as the demand for volatility forecasts has increase to assess the financial risk. Two approaches that have proven useful are the autoregressive conditional heteroskedasticity (ARCH) family, including ARCH model developed by Engle (1982) and generalized ARCH (GARCH) model of Bollerslev (1986), and the stochastic volatility (SV) model, first introduced by Taylor (1982), then Taylor (1986) was the first lengthy published treatment of the problem of volatility modelling in finance. For extensive reviews, see Bollerslev et al. (1994) and Engle (1995) for the ARCH family models and Shephard (2005) provide a comprehensive explanation of the SV models. Both approaches try to model and reproduce the principal properties of the asset returns; however, the difference is that ARCH models explicitly model and specify a process for the conditional variance of returns given past returns observed; while the SV models involve specifying a stochastic process for the volatility and this is modelled as an unobserved variable.

Departures from normality have originated propositions of other distributions in order to capture heavy-tailedness of the asset return distribution in the SV class of models. Heavy-tailed disturbances are often incorporated using distributions such as Student's t-distribution (see, for example, Harvey et al. (1994), Liesenfeld and Jung (2000), Chib et al. (2002), Berg et al. (2004), Jacquier et al. (2004), Omori et al. (2007), Asai (2008), Choy et al. (2008), Nakajima and Omori (2009), Asai and McAleer (2011), Wang et al. (2011), Nakajima (2012) and Delatola and Griffin (2013)), the Normal Inverse Gaussian distribution (*NIG,* see Barndorff-Nielsen (1997) and Andersson (2001)), the Generalized Error Distribution (*GED*, see Liesenfeld and Jung (2000)), the Generalized-t distribution (*GT,* see Wang (2012) and Wang et al. (2013)), a class of mixtures of Normal distributions (Abanto-Valle et al., 2010; Asai, 2009) and, to allow simultaneously treatment of skewness and heavy tails in the conditional distribution of returns, the Skew-GED distribution (Cappuccio et al., 2004, 2006), the Extended Generalized Inverse Gaussian (*EGIG*, see Silva et al. (2006)), the Skew Student's t-distribution (Tsiotas, 2012; Abanto-Valle et al., 2013) and the Generalized Hyperbolic (*GH*) Skew Student's t-distribution (Nakajima and Omori, 2012; Trojan, 2013)[2].

Another characteristic of the return distribution for financial variables is the asymmetric re-

[1] In most cases, it is a negative skewness and it can be viewed as the case where negative returns of a given magnitude are more likely than positive ones of the same magnitude. Regarding excess of kurtosis, it can be viewed as the case where extreme values are more likely than would be dictated by a Normal distribution.

[2] In fact, the *GT*-family nests a number of well-known distributions including Normal, Student-t, Laplace and *GED* distributions. The class of scale mixtures of normal distributions used by Abanto-Valle et al. (2010) include Normal, Student-t, Slash and Variance Gamma distributions. The Weibull and the Generalized Gamma distributions are particular cases of the *EGIG* family, used by Silva et al. (2006).

sponse of volatility known as the "leverage effect": negative past innovations on asset returns tend to increase the current volatility. First noted by Black (1976) and studied by Nelson (1991) and Yu (2005), leverage effect refers to the tendency for changes in asset prices to be negatively correlated with changes in asset volatility. Leverage effect is an important stylized fact of especially stock return indices and has motivated consideration of asymmetric extensions of the basic SV model.

Time-varying volatility for financial variables of developed economies have been studied extensively; however, empirical studies of the Latin American stock market indices so far are very scarce. The volatility characteristics of the financial markets in Latin America are far from being thoroughly analyzed despite their growth in recent years. The main aim of this paper is to estimate SV models incorporating both leverage effects and skewed heavy-tailed disturbances taking into account the *GH* Skew Student's t-distribution for the Latin American stock market indices using the Bayesian estimation method proposed by Nakajima and Omori (2012). The *GH* skew Student's t-distribution includes Normal and Student's t-distributions as special cases. Therefore, the SV model using the *GH* Skew Student's t-distribution (SVSKt model) can take a flexible form to fit the returns and volatility characteristics because the SVSKt model is able to model substantially skewed and heavy tailed data and includes the SV model with Normal disturbances (SV-Normal) and the SV model with symmetric Student's t-disturbances (SVt). We apply the SVSKt model to daily returns of five Latin American stock market indices: Peru, Argentina, Mexico, Chile and Brazil. We also include the U.S. S&P500 returns in order to perdorm some comparisons.

The *GH* Skew Student's t-distribution has been studied by Aas and Haff (2006) and briefly mentioned by Prause (1999) and Jones and Faddy (2003). It belongs to the class of *GH* distributions introduced by Barndorff-Nielsen (1977) and extensively discussed by Prause (1999). The *GH* distribution is a Normal variance-mean mixture and possesses a number of attractive properties: it is closed under conditioning, marginalization, and affine transformations, ii) *GH* distribution can be both symmetric and skew, and its tails are generally semiheavy and iii) *GH* distribution embraces many special cases including Normal, Hyperbolic, Normal Inverse Gaussian (*NIG*), Variance-Gamma, Student-t and skew Student's t-distributions (Aas and Haff, 2006; Nakajima and Omori, 2012). However, estimation and identification of its parameters can be difficult in general due to the flatness of the likelihood function, some parameters are hard to separate and the likelihood function may have several local maxima (Prause, 1999; Aas and Haff, 2006; Deschamps, 2012). Nevertheless Aas and Haff (2006) noted that the *GH* Skew Student's t-distribution is analytically tractable and it may considerably alleviate the identification problem mentioned above. Another advantage is that the *GH* Skew Student's t-distribution exhibit unequal thickness in both tails, unlike to other skewed extensions of the Student-t distribution. This distribution has the property that one tail has polynomial and the other exponential behavior and this offers more flexibility[3].

A main difficulty of the SV framework is the parameter estimation because no explicit expression for the likelihood function of SV model is directly available due to the fact that the variance is an unobserved component. It is possible to compute the likelihood function but this requires the use of simulation techniques, like simulated maximum likelihood, method of simulated moments or Markov Chain Monte Carlo (MCMC) techniques. For an overview of estimation methods of SV models, see Shephard (1996, 2005); Ghysels et al. (1996); Broto and Ruiz (2004). Simulation

[3]Several articles have studied different skew t-type distributions where distributions have two tails behaving as polynomials. This fact means that they fit heavy-tailed data well, but they do not handle substantial skewness. By substantial skewness, Aas and Haff (2006) mean cases with one heavy tail and one nonheavy tail. Their definition relates to the relative fatness of the two tails of the density rather than some threshold for the skewness coefficient.

techniques require a computational burden since we need to repeat the filtering procedure many times to evaluate the likelihood function for each set of parameters until it reaches the maximum (Nakajima, 2012). Computer-intesive methods are thus needed even for the simplest version of the model[4]. In addition, Nakajima and Omori (2012) noted that the *GH* Skew Student's t-distribution is difficult to implement in the SV context due to the large numbers of latent volatility variables. To overcome this difficulty, Nakajima and Omori (2012) have proposed a Bayesian estimation method using the MCMC algorithm for a precise and efficient estimation of the SV model including both leverage effects and skewed heavy-tailed disturbances using the *GH* Skew Student's t-distribution. The key point to implement an efficient MCMC algorithm in the SVSKt model is to express the *GH* Skew Student's t-distribution of the disturbance as a Normal variance-mean mixture of the Generalized Inverse Gaussian (*GIG*), specifically the Inverse Gamma (*IG*) distribution as a mixing distribution among the class of *GIG* distributions to nest and extend various existing SV models.

The paper is organized as follows. In Section 2, we describe a basic SV-Normal model and introduce the *GH* Skew Student's t-distribution in the SV context (SVSKt model). In addition, we describe the Bayesian estimation method using the MCMC algorithm proposed by Nakajima and Omori (2012). Section 3 presents empirical results based on five Latin American stock market indices: Peru, Argentina, Mexico, Chile and Brazil, where the SVSKt model is applied to daily return data using the estimation method proposed by Nakajima and Omori (2012) and the competing SVt models are compared. In order to compare results, the SVSKt is also applied to US S&P500 daily return data. A prior sensitivity analysis is also provided in this Section. Conclusions are presented in Section 4. In the Appendix, we present the properties of the *GH* Skew Student's t-distribution, the MCMC sampling procedure in detail and the Multi-move sampler for the SVSKt model used by Nakajima and Omori (2012).

2 Bayesian Inference for the SV Model with Leverage and Skewed Heavy-Tailed Disturbances using the *GH* Skew Student's t-Distribution

2.1 A Basic SV Model

The SV models assume that the volatility of stock returns has been generated under a latent stochastic process. The basic discrete-time SV model with Normal disturbances can be written as

$$\begin{aligned} y_t &= \exp(h_t/2)\epsilon_t, \qquad t = 1, \ldots, n, && (1)\\ h_{t+1} &= \mu + \phi(h_t - \mu) + \eta_t, \qquad t = 0, \ldots, n-1, && (2)\\ \epsilon_t &\sim N(0, 1), && (3)\\ \eta_t &\sim N(0, \sigma^2), && (4) \end{aligned}$$

where y_t is the asset return and h_t is the unobserved logarithm of the volatility. The volatility process is commonly assumed to follow a stationary AR(1) process by imposing that the persistence parameter satisfies the condition $|\phi| < 1$; this imply that the log-volatility process is stationary and the initial value, h_1, is assumed to follow a stationary distribution by setting $h_0 = \mu$ and $\eta_0 \sim N(0, \sigma^2/(1-\phi^2))$. Finally, ϵ_t and η_t are uncorrelated Normal distributed disturbances.

[4] Despite the computational costs that these techniques involving, increasing computer power and the further development of efficient sampling techniques weaken this drawback noticeably.

There are characteristics of the return distribution for financial variables that the basic SV model with Normal disturbances does not capture such as excess of kurtosis and heavy-tailedness, skewness and the leverage effects. The excess of kurtosis and skewness of the asset return distribution justify the introduction of skewed heavy-tailed disturbances such as the *GH* Skew Student's t-distribution. On the side of the leverage effects, the basic SV model does not allow that the volatility reacts with positive or negative movements in returns. These leverage effects can be incorporated in the SV model assuming that there is any association between the return shocks (ϵ_t) and volatility shocks (η_t).

2.2 A SV Model with Leverage and Skewed Heavy-Tailed Disturbances

According to Nakajima and Omori (2012), the SV model with leverage effects can be written as:

$$
\begin{aligned}
y_t &= \exp(h_t/2)\epsilon_t, & t &= 1,\ldots,n, & (5)\\
h_{t+1} &= \mu + \phi(h_t - \mu) + \eta_t, & t &= 0,\ldots,n-1, & (6)\\
\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} &\sim N(0,\Sigma), \quad \text{with} & \Sigma &= \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}. & (7)
\end{aligned}
$$

This model is similar to previous basic SV model, but now we allow that ϵ_t and η_t are correlated disturbances where the parameter ρ measures the correlations between ϵ_t and η_t. We have volatility asymmetry if $\rho \neq 0$ and specifically, when $\rho < 0$, this indicates a leverage effect: a negative return today will increase volatility tomorrow, and when $\rho = 0$, there is not this type of effects (Yu, 2005).

Regarding the SV model incorporating both leverage effects and skewed heavy-tailed disturbances using the *GH* Skew Student's t-distribution, skewed heavy tails in the return distribution is incorporated into the SV model by replacing the Normal disturbance ϵ_t in (5) by a disturbance from a *GH* Skew Student's t-distribution, denoted by ω_t. This *GH* Skew Student's t-distribution is a limiting case of the more general class of the *GH* distribution. Following Prause (1999) and Aas and Haff (2006), the probability density function of a *GH* random variable ω_t^* is given by:

$$
f_{GH}(\omega^*;\lambda,\delta,\alpha,\mu_\omega,\beta) = \frac{(\alpha^2-\beta^2)^{\lambda/2}\mathbf{K}_{\lambda-1/2}\left(\alpha\sqrt{\delta^2+(\omega^*-\mu_\omega)^2}\right)\exp\left(\beta(\omega^*-\mu_\omega)\right)}{\sqrt{2\pi}\alpha^{\lambda-1/2}\delta^\lambda\mathbf{K}_\lambda\left(\delta\sqrt{\alpha^2-\beta^2}\right)\left(\sqrt{\delta^2+(x-\mu_\omega)^2}\right)^{1/2-\lambda}}, \qquad (8)
$$

where $\mathbf{K}_j$ is the modified Bessel function of the third kind of order j and the parameters must fulfill certain conditions; for more details see Appendix A. The *GH* distribution may be represented as a Normal variance-mean mixture with the Generalized Inverse Gaussian (*GIG*) distribution as a mixing distribution. This means that the *GH* variable ω_t^* can be represented as:

$$
\omega_t^* = \mu_\omega + \beta z_t^* + \sqrt{z_t^*}\epsilon_t, \qquad \epsilon_t \sim N(0,1), \qquad z_t^* \sim GIG(\lambda,\delta,\gamma), \qquad (9)
$$

with ϵ_t and z_t independent and $\gamma = \sqrt{\alpha^2-\beta^2}$. The *GH* Skew Student's t-distribution is the special case where $\lambda = -\nu/2 (\nu > 0)$ and $\alpha \to |\beta|$ (the latter implies $\gamma = 0$) in equation (8). The probability density function of a *GH* Skew Student's t- random variable ω_t is given by:

$$f_{GHskewt}(\omega;\nu,\delta,\mu_\omega,\beta) = \frac{2^{\frac{1-\nu}{2}}\delta^{\nu}\left|\beta\right|^{\frac{v+1}{2}}\mathbf{K}_{\frac{v+1}{2}}\left(\sqrt{\beta^2\left(\delta^2+(\omega-\mu_\omega)^2\right)}\right)\exp\left(\beta\left(\omega-\mu_\omega\right)\right)}{\Gamma(\frac{v}{2})\sqrt{\pi}\left(\sqrt{\delta^2+(\omega-\mu_\omega)^2}\right)^{\frac{v+1}{2}}}, \beta\neq 0, \tag{10}$$

and

$$f_{GHskewt}(\omega;\nu,\delta,\mu_\omega) = \frac{\Gamma(\frac{v+1}{2})}{\sqrt{\pi}\delta\Gamma(\frac{v}{2})}\left[1+\frac{(\omega-\mu_\omega)^2}{\delta^2}\right]^{-(v+1)/2}, \beta = 0. \tag{11}$$

where $\Gamma(.)$ is the gamma function. The density $f_{GHskewt}(\omega;\nu,\delta,\mu_\omega)$ in (11) is known as the non-central Student's t-distribution with ν degrees of freedom.

As observed in the literature, estimation and identification of GH distribution parameters can be difficult in general (Prause, 1999; Aas and Haff, 2006; Deschamps, 2012) even for a GH Skew Student's t-distribution with $\lambda = -\nu/2$ ($\nu > 0$) and $\gamma = 0$ (Nakajima and Omori, 2012). In order to overcome these difficulties, Nakajima and Omori (2012) make the additional assumption that $\delta = \sqrt{\nu}$ and, show that their proposed parameterization is appropriate for the SV model with the GH Skew Student's t-distribution because it allows a parsimonious representation that is more amenable to estimation and leads an efficient MCMC sampling. This additional assumption yields $z_t^* = z_t \sim GIG(-\nu/2, \sqrt{\nu}, 0)$, or equivalently $IG(\nu/2, \nu/2)$ where IG denotes the Inverse Gamma distribution. Therefore, the GH Skew Student's t-disturbance, ω_t, can be express in the form of the normal variance-mean mixture as:

$$\omega_t = \mu_\omega + \beta z_t + \sqrt{z_t}\epsilon_t, \qquad \epsilon_t \sim N(0,1), \qquad z_t \sim IG(\nu/2, \nu/2), \tag{12}$$

where μ_ω and β are the location and skew parameters, respectively and the IG distribution is the mixing distribution among the class of GIG distributions. Nakajima and Omori (2012) argue that the structure of (12) leans itself well to the construction of a MCMC algorithm in the Bayesian inference context. To allow $E(\omega_t) = 0$, it is assumed that $\mu_\omega = -\beta\mu_z$, where $\mu_z \equiv E(z_t) = \nu/(\nu-2)$. The variance of ω_t is only finite when $\nu > 4$, as opposed to the symmetric Student's t-distribution which only requires $\nu > 2$, because of that an additional constraint is imposed, $\nu > 4$, in order to ensure existence of the second moment of ω_t.

Regarding the tails of the GH Skew Student's t-distribution, this distribution has the property that it exhibits unequal thickness in both tails where one tail has polynomial and the other exponential behavior. It is the only subclass of the GH family of distributions having this property. Thus, the GH Skew Student's t-distribution has one heavy and one semiheavy tail. This makes it unique for modeling substantially skewed and heavy-tailed data as found in financial markets (Aas and Haff, 2006; Trojan, 2013). The tails of the GH Skew Student's t-distribution are characterized uniquely by the parameters β and ν, which determine jointly the degree of skewness and heavy tailedness. A lower value of β (when ν fixed) implies a more negative skewness as well as heavier tails. On the other hand, as ν becomes larger (when β fixed) the density becomes less skewed and has lighter tails. The Figure 1 shows densities of the GH Skew Student's t-distribution using several combinations of the parameter values of β and ν, and demonstrates how both parameters determine jointly the skewness and kurtosis of the distribution.

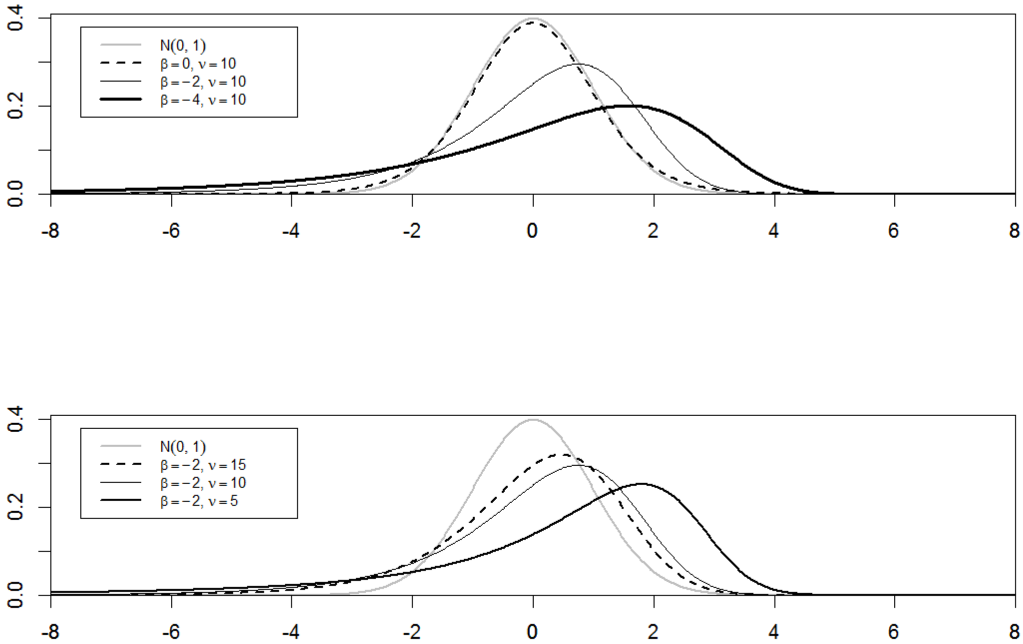

Figure 1. Densities of the GH Skew Student's t-distribution. Parameter β varying using $\beta = 0$ (symmetric t), -2 and -4 with $\nu = 10$ fixed (top); and parameter ν varying using $\nu = 5$, 10 and 15 with $\beta = -2$ fixed (bottom).

Taking into account the above issues, the SV model incorporating both leverage effects and skewed heavy-tailed disturbances using the *GH* Skew Student's t-distribution (SVSKt model) can be written as:

$$
\begin{aligned}
y_t &= \exp(h_t/2)\{\beta(z_t - \mu_z) + \sqrt{z_t}\epsilon_t\}, & t &= 1, \ldots, n, & (13)\\
h_{t+1} &= \mu + \phi(h_t - \mu) + \eta_t, & t &= 0, \ldots, n-1, & (14)\\
z_t &\sim IG(\nu/2, \nu/2), & & & (15)\\
\begin{pmatrix} \epsilon_t \\ \eta_t \end{pmatrix} &\sim N(0, \Sigma), \quad \text{with} & \Sigma &= \begin{bmatrix} 1 & \rho\sigma \\ \rho\sigma & \sigma^2 \end{bmatrix}. & (16)
\end{aligned}
$$

The degree of freedom $\nu > 4$ is unknown to be estimated. The SVSKt model includes the SV model with Normal disturbances (SV-Normal) when $\beta = 0$ and $z_t \equiv 1$ for all t and the SV model with symmetric Student's t-disturbances (SVt) when $\beta = 0$.

2.3 Bayesian Estimation of the SVSKt Model

We use the Bayesian estimation method proposed by Nakajima and Omori (2012) using the MCMC algorithm for the SVSKt model. In this subsection, we present some preliminaries issues about the estimation of SV models within the Bayesian context and a briefly discussion about the steps of the MCMC algorithm of Nakajima and Omori (2012).

Estimation of SV models consists of two stages: estimation of the set of parameters of the model, and estimation of the unobserved volatility time series. Techniques based on MCMC algorithms offer a framework both for estimating the parameters of the SV models and for assessing the latent volatilities. These methods have had a widespread influence on the theory and practice of Bayesian inference that are based on the posterior distributions of parameters given the observed data using the Bayes Theorem, where $\pi(\theta \mid y) \propto f(y \mid \theta)\pi(\theta)$ is the the posterior distribution of parameters conditional on the data y, θ is the vector that contains all parameters of the model, $f(y \mid \theta)$ is the likelihood function, and $\pi(\theta)$ are the priors which are beliefs about the distributions of the parameters. The idea behind the MCMC algorithms is to produce random variables from a given multivariate density (the posterior density in Bayesian applications) by repeatedly sampling a Markov chain whose invariant distribution is the target density of interest (Kim et al., 1998). There are typically many different ways of constructing a Markov chain with this property; but a key point is to isolate those that are simulation-efficient in the context of SV models, therefore the design of the MCMC algorithm is important for the speed of convergence of the chains.

In the SV context, the likelihood function to be maximized is given by:

$$f(y \mid \theta) = \int f(y \mid h, \theta) f(h \mid \theta) dh. \tag{17}$$

where θ is the vector that contains all parameters of the SV model. Jacquier et al. (1994) argue that the likelihood function has no analytical representation and is intractable. This fact precludes the direct analysis of the posterior density $\pi(\theta \mid y)$ by MCMC methods. This problem can be overcome by focusing instead on the density $\pi(\theta, h \mid y)$, where $h = (h_1, \ldots, h_n)$ is the vector of n latent log-volatilities (Kim et al., 1998). The MCMC procedures can be developed to sample this density without computation of the likelihood function $f(y \mid \theta)$. These draws can be used to make inferences by appealing to suitable ergodic Theorems for Markov chains. For example, posterior moments and marginal densities can be estimated by averaging the relevant function of interest over the sampled random variables. The posterior mean of θ is estimated by the sample mean of the simulated θ values.

Several approaches of MCMC algorithms have been suggested for the estimation of the SV model within the Bayesian context. Jacquier et al. (1994) use the single-move Gibbs sampling within the Metropolis–Hastings algorithm to sample from the log-volatilities $h = (h_1, \ldots, h_n)$. This algorithm consists of generating sample of one state, h_t, at a time given others, h_k $(k \neq t)$. Some researchers have argued that when parameters are correlated, the single-move procedure results in a slower speed of convergence of the Markov chain. Kim et al. (1998) developed the mixture sampler that approximates the distribution of log-squared returns by mixture of Normal distributions, allowing jointly drawing on the components of the whole vector of log-volatilities. Another approach, developed by Shephard and Pitt (1997) and Watanabe and Omori (2004) in the context of state space modeling, uses the multi-move sampler for generating the log-volatility in the SV model updating several variables at a time. This algorithm can produce efficient samples from the target conditional posterior distribution by dividing the process of $h = (h_1, \ldots, h_n)$ into several blocks and generate sample of each block given other blocks. Regarding the SV model with leverage, Omori and Watanabe (2008) developed the associated multi-move sampler and showed that it produces efficient samples. The mixture sampler and multi-move sampler are more efficient than the single-move sampler that generate sample of one state, h_t, at a time given others, h_k $(k \neq t)$ (Nakajima, 2012).

The Bayesian estimation method proposed by Nakajima and Omori (2012) for the SVSKt model extends the method developed by Omori and Watanabe (2008) for sampling h using the multi-move sampler. They noted that the key point to implement an efficient MCMC algorithm in the SVSKt model is to express the *GH* Skew Student's t-distribution of the disturbance as a Normal variance-mean mixture of the *GIG*, as stated in (12), specifically the *IG* distribution as a mixing distribution among the class of *GIG* distributions. They consider the variable z_t, following the mixing distribution, as a latent variable. The conditional posterior distribution of each parameter is reduced to a much more tractable form conditional on z_t than when the model is considered in the direct likelihood form of the *GH* Skew Student's t-distribution[5]. This treatment allows to draw sample from the conditional posterior distribution of z_t for $t = 1, \dots n$. Nakajima and Omori (2012) use the following sampling algorithm for the SVSKt model using the MCMC method.

Let $\theta = (\phi, \sigma, \rho, \mu, \beta, \nu)$, $\{y_t\}_{t=1}^n, \{h_t\}_{t=1}^n$, $\{z_t\}_{t=1}^n$ and $\pi(\phi)$, $\pi(\vartheta)$ and $\pi(\nu)$ are the prior probability densities of ϕ, $\vartheta \equiv (\sigma, \rho)'$ and ν respectively. Random samples are drawn from the posterior distribution of (θ, h, z) given y. The sampling steps are given by:

1. Initialize θ, 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.

The full algorithm describing more details of each sampling step can be found in Appendix B and the details of the multi-move sampler are described in the Appendix C.

3 Empirical Application to Stock Return Data

3.1 The Data

For Bayesian estimation of the SVSKt model, we consider the daily returns of five Latin American stock market indices: Peru, Argentina, Mexico, Chile and Brazil. The Latin American stock market indices are in Table 1. We use a sample from 1996/1/2 to 2013/12/30 for all stock market indices

[5] Nakajima and Omori (2012) noted that when $\rho = 0$, the closed form of the density $f(y_t \mid h_t)$, which is marginalized over z_t, is available. However, in the case $\rho \neq 0$, the closed form of the density $f(y_t \mid h_t, h_{t+1})$ is not available. Therefore, the latent variable z_t plays an important role in exploring the posterior distribution using the MCMC algorithm.

of Latin American except to Peru where the period is from 2001/1/2 to 2013/12/30 because there was a change in the methodology of the IGBVL index in November 1998 and this could affect the results. 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 stock market indices. One reason is that the U.S. stock market could be considered as a good benchmark.

Table 1. Latin American Stock Market Indices

Index	Country	Period
IGBVL	Peru	2001/1/2 - 2013/12/30
MERVAL	Argentina	1996/1/2 - 2013/12/30
MEXBOL	Mexico	1996/1/2 - 2013/12/30
IPSA	Chile	1996/1/2 - 2013/12/30
IBOVESPA	Brazil	1996/1/2 - 2013/12/30

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 2 shows the time series plots of the daily stock returns.

Table 2. Summary Statistics for Daily Stock Returns Data

Index	Obs.	Mean	S.D.	Skewness	Excess Kurtosis	Min.	Max.
IGBVL	3246	0.0008	0.0149	-0.5287	10.8286	-0.1329	0.1282
MERVAL	4439	0.0005	0.0215	-0.2801	5.3395	-0.1476	0.1612
MEXBOL	4529	0.0006	0.0151	0.0300	6.9916	-0.1431	0.1215
IBOVESPA	4456	0.0006	0.0213	0.2994	13.1430	-0.1723	0.2882
IPSA	4489	0.0003	0.0111	0.1332	7.9881	-0.0767	0.1180
S&P500	4531	0.0002	0.0127	-0.2272	7.4884	-0.0947	0.1096

Skewness statistics are sometimes used to assess the symmetry of distributions while kurtosis statistics are often interpreted as a measure of similarity to a Normal distribution. These statistics are sensitive to extreme observations because they make use of the third and fourth powers of the observations, respectively. The IGBVL and the MERVAL series are negatively skewed while the MEXBOL, IBOVESPA and the IPSA series are positively skewed. However, the skewness of the MEXBOL is very close to zero. The IGBVL index is the most negatively skewed with -0.5287 and the IBOVESPA index is the most positive skewed with 0.2994. Regarding the kurtosis, all the daily returns of Latin American indices considered have positive kurtosis where IBOVESPA has the highest value 13.1430. All five sets of returns of Latin American indices are leptokurtic, since all the estimates of kurtosis in Table 2 exceed 3, which is the kurtosis value for Normal distribution. Regarding the S&P500 daily returns, this index also has negative skewness and positive kurtosis. The summary statistics show that daily stock returns appear to be distributed with fat-tails for the five Latin American empirical returns distribution of the data and negative skewness for the IGBVL and the MERVAL. It is clear that the returns-generating process is not even approximately Gaussian.

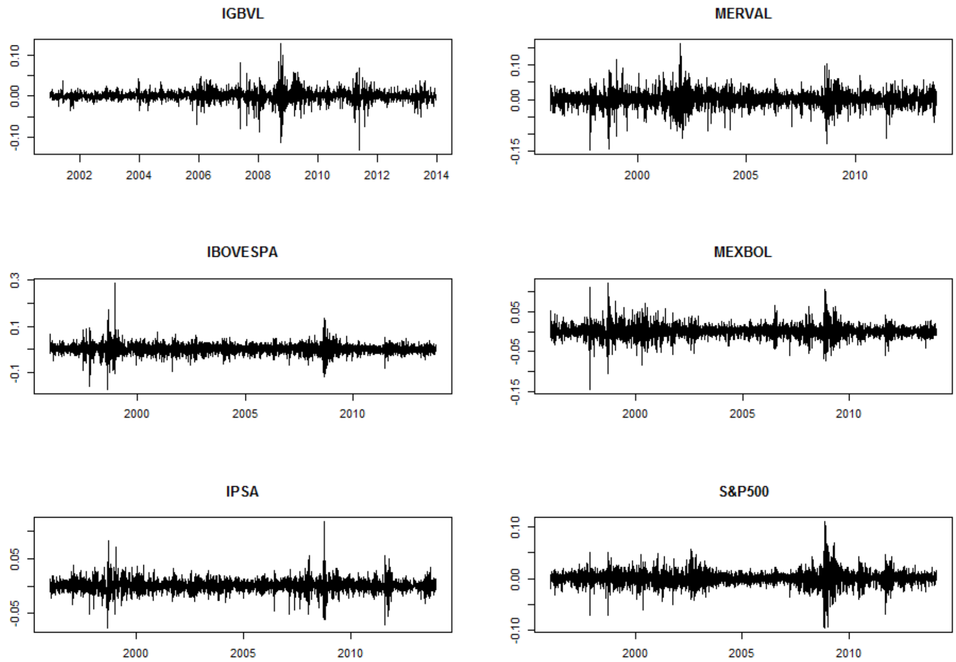

Figure 2. Times series plots for IGBVL (2001/01/02 - 2013/12/30) and MERVAL, IBOVESPA, MEXBOL, IPSA and S&P500 (1996/01/02 - 2013/12/30) daily returns.

3.2 Parameter Estimates

For parameter estimates of the SVSKt model, we use the same prior distributions as Nakajima and Omori (2012). The following prior distributions are assumed and commonly used in the literature (see, for example, Kim et al. (1998), Meyer and Yu (2000), Yu (2005), Omori et al. (2007), Nakajima and Omori (2009), Nakajima (2012), Trojan (2013)):

1. Let $\phi = 2\phi^* - 1$ and we specify a $Beta(\alpha, \beta)$ prior distribution for ϕ^* with $\alpha = 20$ and $\beta = 1.5$ which implies that the prior mean and prior standard deviation of ϕ are $(0.8605, 0.1074)$. Our prior on ϕ has the support on the interval $(-1, 1)$ and mirrors a belief in moderate volatility persistence with mean 0.86.

2. We assume a conjugate Inverse-Gamma prior for σ^2, that is $\sigma^2 \sim IG(\alpha, \beta)$ with shape parameter $\alpha = 2.5$ and scale parameter $\beta = 0.025$ which implies that the prior mean and prior standard deviation of σ^2 are $(0.0167, 0.0236)$.

3. We employ a Normal prior distribution for μ, that is $\mu \sim N(-10, 1)$[6] and a $U(-1, 1)$ prior distribution for ρ.

4. We specify a standard Normal prior distribution for β, that is $\beta \sim N(0, 1)$ and a $Gamma(\alpha, \beta)$ prior distribution for ν with shape parameter $\alpha = 16$ and rate parameter $\beta = 0.8$. We assume a additional constraint $\nu > 4$ in the prior distribution of ν for ensure existence of

[6] Kim et al. (1998) and Meyer and Yu (2000) employ a slightly informative prior for μ, $\mu \sim N(0, 10)$.

the second moment of ω_t, that is $E(\omega_t^2) < \infty$. Thus, the prior distribution of ν is $\nu \sim Gamma(16, 0.8)I(\nu > 4)$ which implies that the prior mean and prior standard deviation of ν are $(20, 5)$.

The MCMC simulation are conducted with 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. Using the 20 000 samples for each of the parameters, the posterior means, the standard deviations, the 95% intervals, and the inefficiency factor are obtained. The posterior means are computed by averaging the simulated samples. The 95% intervals are calculated using the 2.5th and 97.5th percentiles of the simulated samples. The MCMC sampler is initialized by setting $\phi = 0.97$, $\sigma = 0.2$, $\rho = -0.3$, $\mu = -10$, $\beta = -0.3$ and $\nu = 15$ for MERVAL, MEXBOL, IBOVESPA, IPSA and S&P500 and $\phi = 0.85$, $\sigma = 0.8$, $\rho = -0.05$, $\mu = -9$, $\beta = -0.015$ and $\nu = 30$ for IGBVL.

We compute the inefficiency factor to check the efficiency of the MCMC algorithm. The inefficiency factor is defined by $1 + 2\Sigma_{s=1}^{\infty}\rho_s$, where ρ_s is the sample autocorrelation at lag s. It measures how well the MCMC chain mixes (Chib, 2001; Nakajima and Omori, 2009, 2012). It is the estimated ratio of the numerical variance of the posterior sample mean to the variance of the hypothetical sample mean from uncorrelated draws. The inefficiency factor serves to quantify the relative efficiency from correlated versus independent samples. 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 $b_w = 1\ 000$.

Figures 3 - 8 show the MCMC estimation results of the SVSKt model for the IGBVL, MERVAL, MEXBOL, IBOVESPA, IPSA and S&P500 indices, respectively. Regarding the Latin American stock indices, 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 to the IGBVL, we obtain poor mixing (or slow convergence) of the Markov chain for some parameters (ϕ, σ and μ) and estimation results show high autocorrelation through iterations of ϕ, σ and μ with a slowly decay. Regarding S&P500, we obtain similar results as Nakajima and Omori (2012).

Table 3 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 daily returns data. The posterior means of ϕ, 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 daily returns data, it has a posterior mean of $\phi = 0.8618$, which indicates a low persistence in comparison to the volatility of the others indices above mentioned.

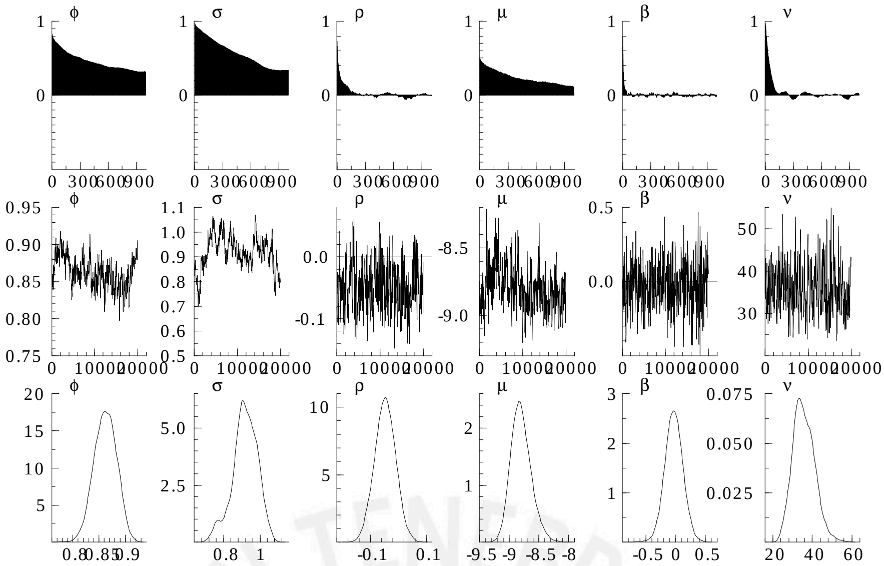

Figure 3. MCMC estimation results of the SVSKt model for IGBVL data (Peru). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

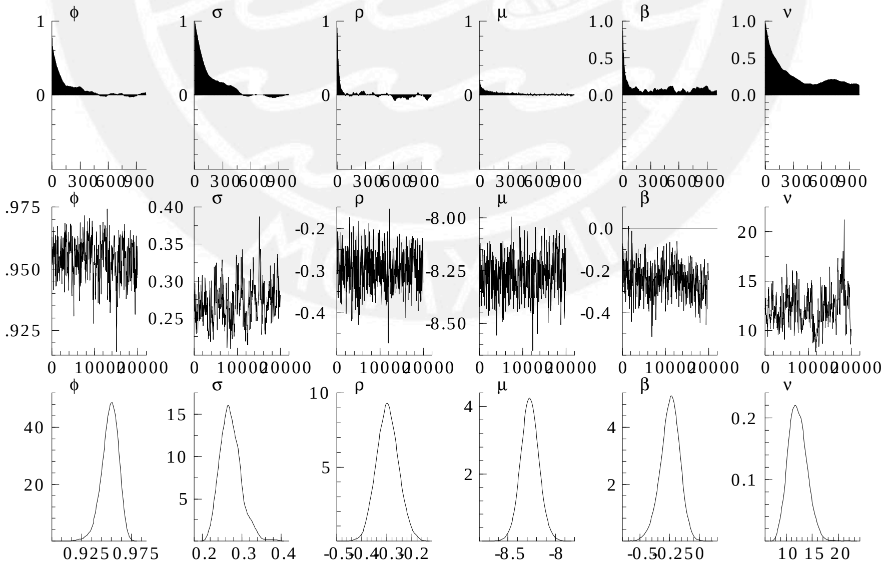

Figure 4. MCMC estimation results of the SVSKt model for MERVAL data (Argentina). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

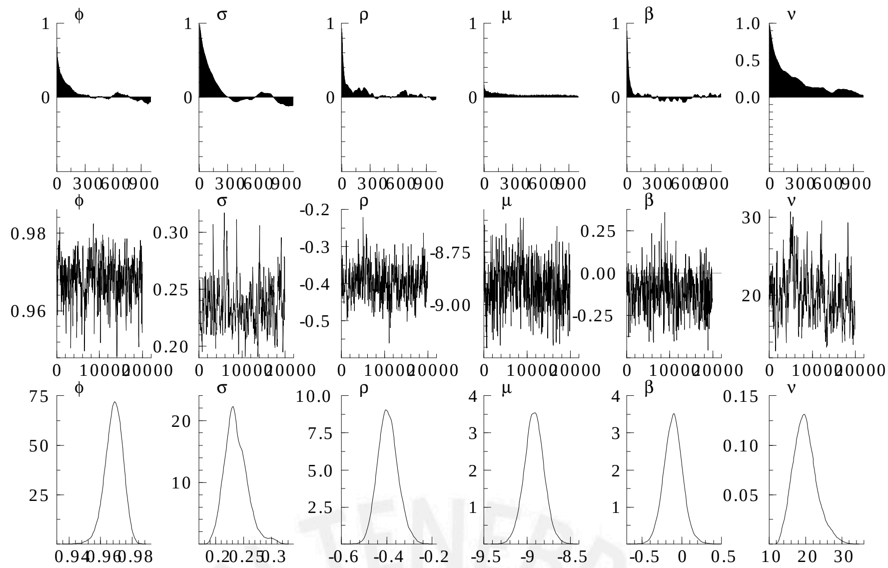

Figure 5. MCMC estimation results of the SVSKt model for MEXBOL data (Mexico). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

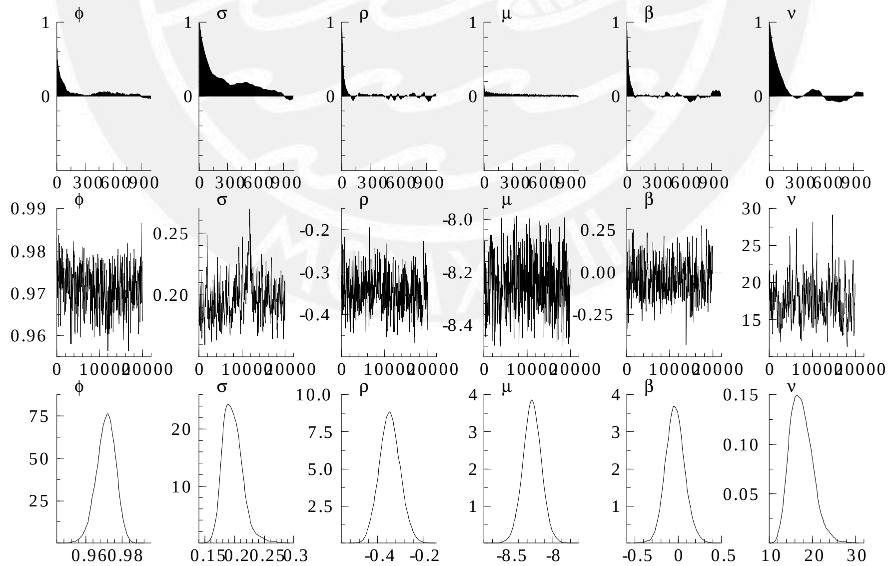

Figure 6. MCMC estimation results of the SVSKt model for IBOVESPA data (Brazil). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

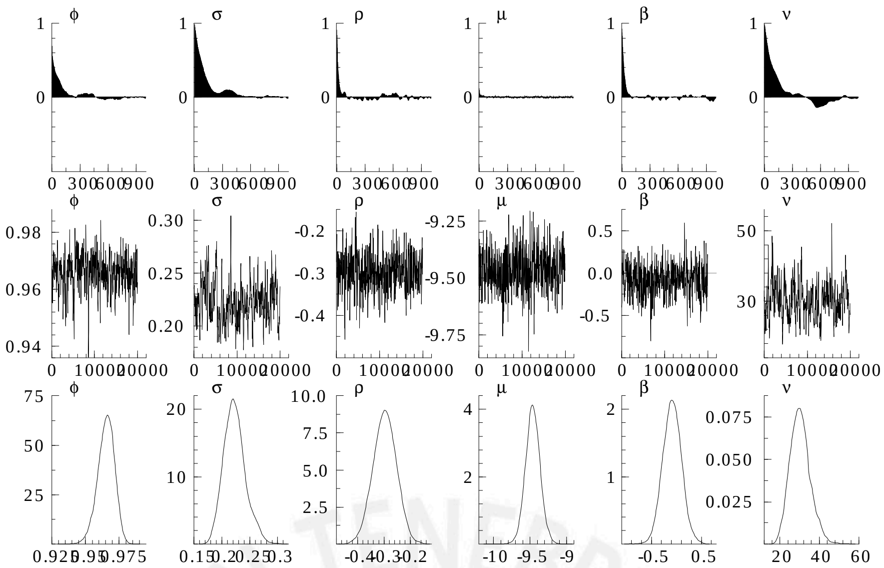

Figure 7. MCMC estimation results of the SVSKt model for IPSA data (Chile). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

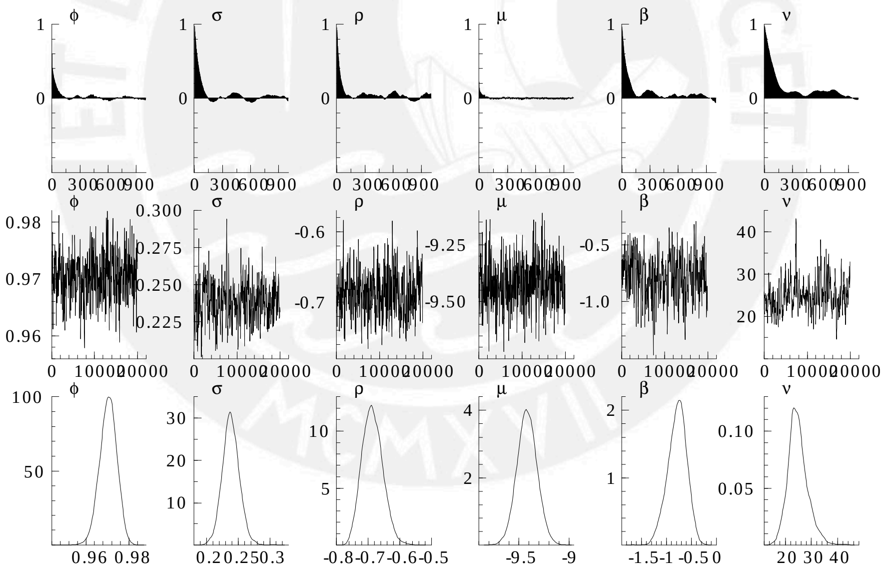

Figure 8. MCMC estimation results of the SVSKt model for S&P500 data (US). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

Table 3. MCMC Estimation Results of the SVSKt Model for Latin American Stock Return Data

Parameter	Mean	S.D.	95% interval		Inefficiency
(i) IGBVL					
ϕ	0.8618	0.0213	[0.8197,	0.9011]	417.43
σ	0.9173	0.0682	[0.7602,	1.0380]	539.46
ρ	-0.0475	0.0370	[-0.1197,	0.0245]	59.91
μ	-8.8002	0.1686	[-9.0995,	-8.4321]	233.32
β	-0.0286	0.1508	[-0.3247,	0.2678]	16.87
ν	35.6892	5.4920	[25.9746,	47.2584]	88.54
(ii) MERVAL					
ϕ	0.9535	0.0086	[0.9347,	0.9681]	129.12
σ	0.2717	0.0269	[0.2264,	0.3316]	226.88
ρ	-0.2977	0.0436	[-0.3807,	-0.2106]	35.47
μ	-8.2705	0.0951	[-8.4565,	-8.0816]	26.19
β	-0.2464	0.0811	[-0.4180,	-0.0952]	82.38
ν	12.2135	1.7827	[9.0445,	15.9292]	286.25
(iii) MEXBOL					
ϕ	0.9683	0.0057	[0.9562,	0.9785]	90.33
σ	0.2362	0.0206	[0.2015,	0.2847]	176.82
ρ	-0.3948	0.0444	[-0.4769,	-0.3014]	81.93
μ	-8.9186	0.1134	[-9.1434,	-8.6942]	30.52
β	-0.1076	0.1194	[-0.3389,	0.1326]	37.20
ν	19.8882	3.2235	[14.2966,	27.2299]	265.91
(iv) IBOVESPA					
ϕ	0.9711	0.0053	[0.9601,	0.9806]	67.00
σ	0.1969	0.0168	[0.1696,	0.2358]	246.63
ρ	-0.3464	0.0452	[-0.4333,	-0.2551]	37.82
μ	-8.2422	0.1057	[-8.4519,	-8.0325]	21.19
β	-0.0342	0.1107	[-0.2437,	0.1954]	38.07
ν	17.5026	2.6309	[13.1482,	23.3003]	159.85
(v) IPSA					
ϕ	0.9653	0.0062	[0.9520,	0.9763]	82.62
σ	0.2226	0.0196	[0.1878,	0.2656]	179.16
ρ	-0.2970	0.0440	[-0.3835,	-0.2120]	26.06
μ	-9.4626	0.0995	[-9.6562,	-9.2626]	4.26
β	-0.0852	0.1872	[-0.4475,	0.2858]	37.72
ν	30.2523	5.1130	[21.1299,	41.1597]	166.06

The posterior means of ρ, that measures the correlations between ϵ_t and η_t, are estimated to be negative for all indices considered. When ρ values are negative, it implies that there exist leverage effects. MEXBOL and IBOVESPA have in absolute value the highest posterior mean estimates of ρ (-0.3948 and -0.3464, respectively), which imply that the leverage effect is more notable for these indices. Also the 95% credible intervals are negative implying that the posterior probability that ρ is negative is greater than 0.95, and the negativity of ρ is credible. The same applies with MERVAL and IPSA (posterior mean estimates for ρ are -0.2977 and -0.2970, respectively) where the 95% credible intervals are negative, but it is a minor leverage effect than the previous indices. In the case of IGBVL, the posterior mean estimate of ρ 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 ρ, although mainly located in the negative range, 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 parameter σ, the posterior mean estimates of σ show that all indices 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 σ takes a very high value (0.9173) compared to the other indices. This implies that the variance of the shock η_t is large and the log-volatility has more variability than the other stock indices in Latin American. Regarding the posterior mean of μ, all indices show similar results in the range of -19.4626 to -8.2422.

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 β and ν. With ν fixed, a lower value of β implies a more negative skewness or left-skewness as well as heavier tails. On the other hand, with β fixed, as ν becomes larger the density becomes less skewed and has lighter tails. The posterior means of β are estimated to be negative for all indices returns data considered. MERVAL has the less value of posterior mean estimate of β with -0.2464 and the 95% credible intervals are negative implying that the posterior probability that β is negative is greater than 0.95, and the negativity of β is credible. However, the posterior mean estimates of β 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 β are very close to zero for IGBVL, IBOVESPA and IPSA, it could imply the case of symmetric heavy tailed disturbances. Finally, the posterior means of ν are around 35.6892 for IGBVL, 12.2135 for MERVAL, 19.8882 for MEXBOL, 17.5026 for IBOVESPA and 30.2523 for IPSA returns.

Figures 9-13 show the density of the *GH* Skew Student's t-distribution with the estimates parameters of Table 3 for the indices considered. Four points are worth mentioning: (i) The distributions of the IGBVL, MEXBOL, IBOVESPA and IPSA appear to be symmetrical, (ii) In the cases of symmetric distributions, the MEXBOL and IBOVESPA have heavier tails than the IGBVL and IPSA, (iii) The distribution of the IGBVL is similar to the Normal distribution and (iv) The distribution of the MERVAL have negative skewness (asymmetric) and has heavier tails than the others indices considered. These results support the evidence of skewed heavy-tailed disturbances only for the MERVAL, symmetric heavy-tailed disturbances for the MEXBOL, IBOVESPA and IPSA, and symmetric Normal disturbances for the IGBVL.

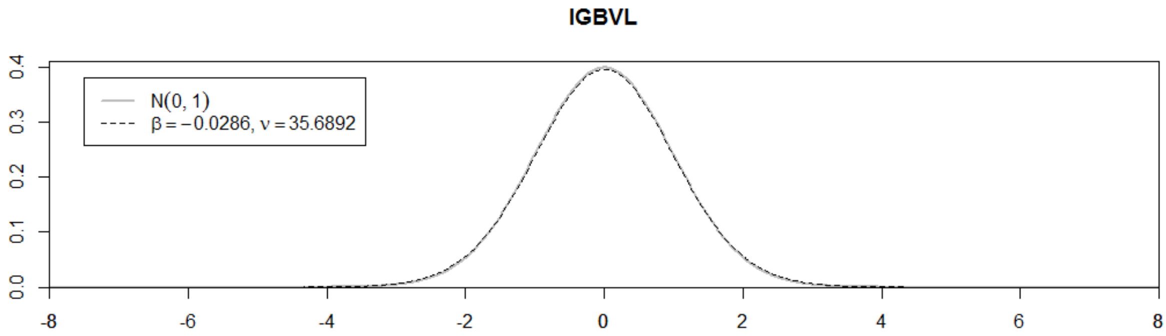

Figure 9. Density of the GH Skew Student's t-distribution with estimates parameters $\beta = -0.0286$ and $\nu = 35.6892$ for IGBVL data (Peru).

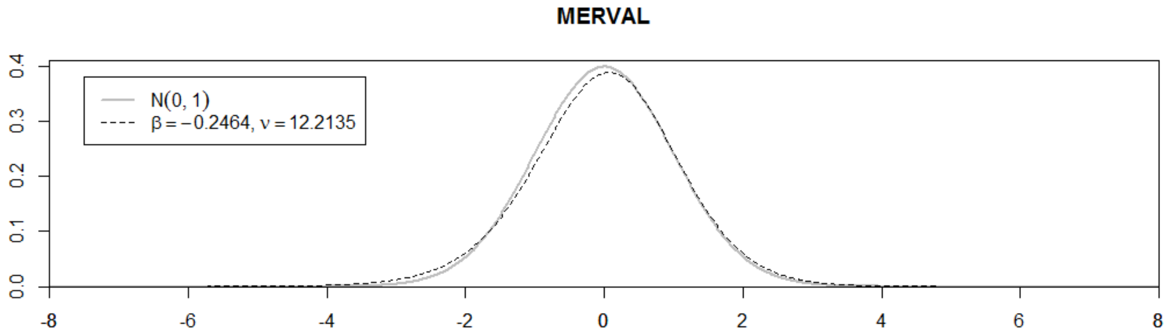

Figure 10. Density of the GH Skew Student's t-distribution with estimates parameters $\beta = -0.2464$ and $\nu = 12.2135$ for MERVAL data (Argentina).

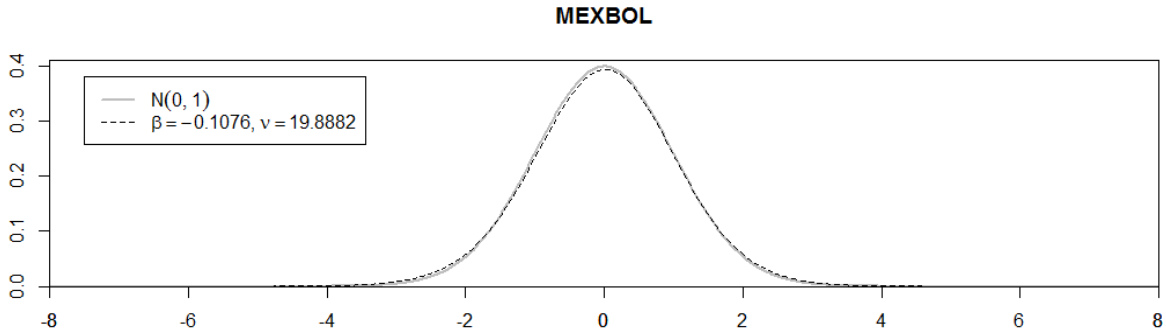

Figure 11. Density of the GH Skew Student's t-distribution with estimates parameters $\beta = -0.1076$ and $\nu = 19.8882$ for MEXBOL data (Mexico).

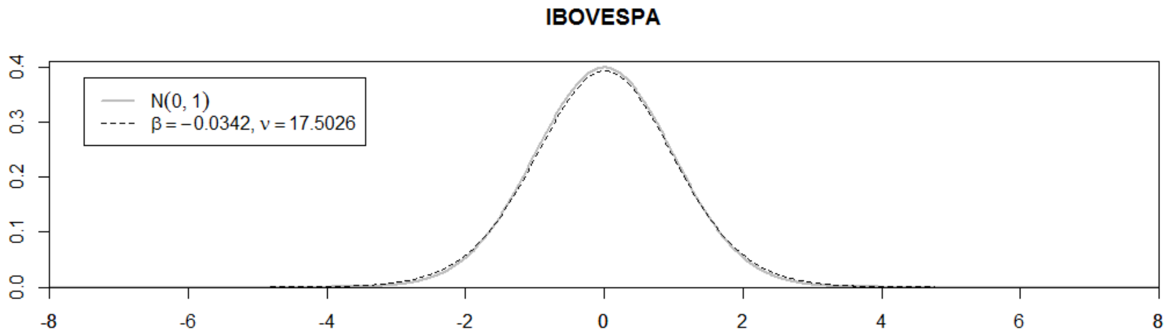

Figure 12. Density of the GH Skew Student's t-distribution with estimates parameters $\beta = -0.0342$ and $\nu = 17.5026$ for IBOVESPA data (Brazil).

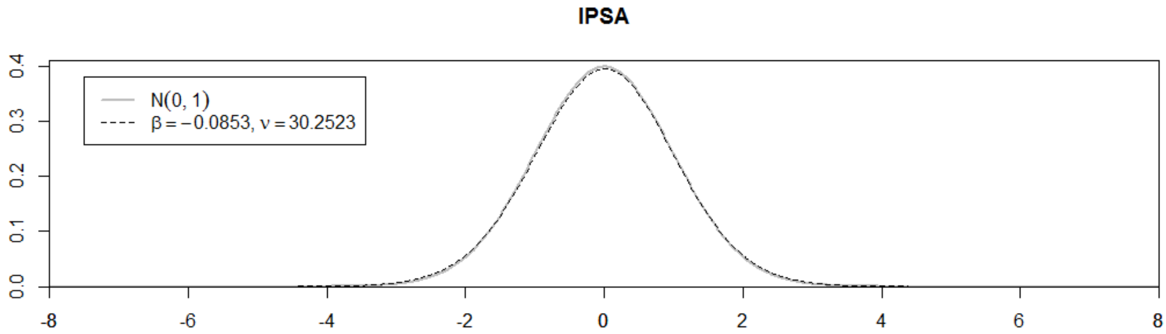

Figure 13. Density of the GH Skew Student's t-distribution with estimates parameters $\beta = -0.0853$ and $\nu = 30.2523$ for IPSA data (Chile).

Regarding the S&P500 daily returns, Table 4 shows the results of the posterior estimates. These results are very similar to Nakajima and Omori (2012). The posterior mean of ϕ are close to one (0.9703) and this fact implies high persistence, more than Latin American stock returns considered. The posterior mean of ρ is estimated to be negative (-0.6864) which imply the evidence of leverage effects for the S&P500. Also the 95% credible intervals are negative implying that the posterior probability that ρ is negative is greater than 0.95, and the negativity of ρ is credible. The posterior mean estimate of σ is 0.2382, similar parameter estimates to MERVAL, MEXBOL, IPSA and IBOVESPA indices. The posterior mean of μ is -9.4186. The posterior mean of β is estimated to be negative (-0.7842). Also the 95% credible intervals are negative implying that the posterior probability that β is negative is greater than 0.95, and the negativity of β is credible. Finally, the posterior means of ν are around 24.8685. Figure 14 shows the density of the *GH* Skew Student's t-distribution with the estimates parameters of Table 4 for the US S&P500. The distribution of the S&P500 has more negative skewness (asymmetric) and heavier tails than the Latin American indices considered. The negative skewness and heavy tails are more notable in this case.

Table 4. Estimation Results of the SVSKt Model for US S&P500 Stock Return Data

		S&P500		
Parameter	Mean	S.D.	95% interval	Inefficiency
ϕ	0.9703	0.0040	[0.9622, 0.9778]	40.86
σ	0.2382	0.0135	[0.2125, 0.2652]	89.75
ρ	-0.6864	0.0323	[-0.7449, -0.6199]	76.70
μ	-9.4186	0.1001	[-9.6135, -9.2207]	7.79
β	-0.7842	0.1899	[-1.1837, -0.4380]	116.81
ν	24.8685	3.8282	[18.2556, 33.5375]	189.21

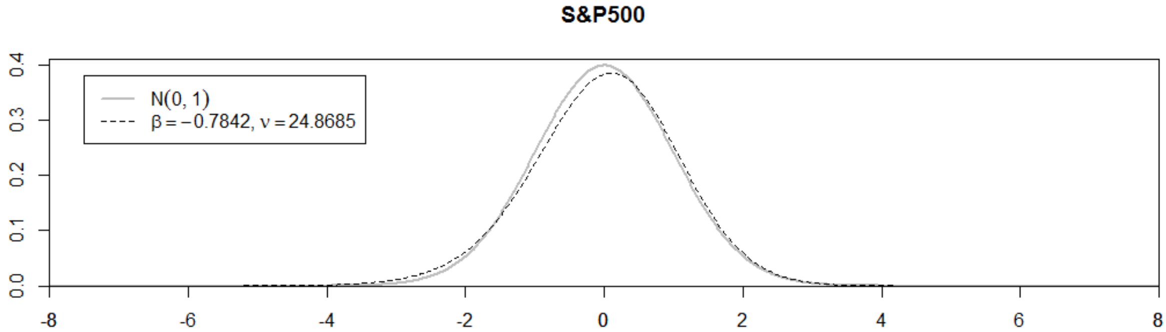

Figure 14. Density of the GH Skew Student's t-distribution with estimates parameters $\beta = -0.7842$ and $\nu = 24.8685$ for S&P500 data (US).

The indicator of how well MCMC chain mixes is measured by the inefficiency factor of the MCMC algorithm defined by $1 + 2\Sigma_{s=1}^{\infty}\rho_s$ as mentioned before. The inefficiency factor shows high values for parameters ϕ, σ and μ for IGBVL. These results are supporting by the initial MCMC Figure 3 that shows high autocorrelation through iterations of parameters ϕ, σ and μ for IGBVL that decay slowly. MERVAL, MEXBOL, IBOVESPA and IPSA returns show low values of inefficiency factor in all parameters estimates but parameter σ and v 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.

Figure 15 shows the log-volatility estimates for the Latin American stock indices returns considered. The results show that there is a similar pattern between periods of higher volatility between the five Latin American indices. Most of times, these periods of high volatility are associated with international crisis. For example, all indices 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.

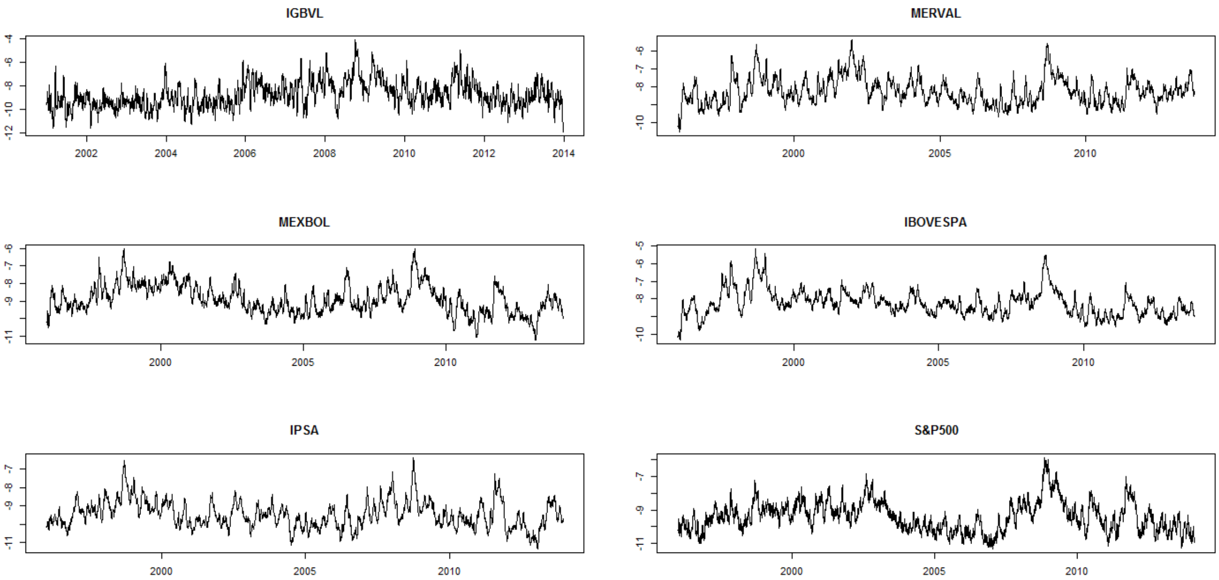

Figure 15. Log-volatility for IGBVL (2001/01/02 - 2013/12/30) and MERVAL, MEXBOL, IPSA, IBOVESPA and S&P500 (1996/01/02 - 2013/12/30) daily returns.

3.3 Model Comparison

In this subsection, model comparison between competing models for the daily stock returns are provided. We make a comparison between the SVSKt model with the SVt model (with symmetric Student's t-disturbances, or equivalently the SVSKt model with $\beta = 0$). All models compared are allowed to include leverage effects. Model comparison in a Bayesian framework can be performed using posterior odds. If $y = \{y_t\}_{t=1}^n$ denote the returns observation vector; then, the posterior odds in favor of model A, M_A, to model B, M_B, is given by:

$$\frac{f(M_A|y)}{f(M_B|y)} = \frac{f(y|M_A)}{f(y|M_B)} \frac{f(M_A)}{f(M_B)}, \tag{18}$$

where $f(M_i|y)$ is the posterior probability of the model i with $i = A, B$, $f(M_i)$ is the prior probability of the model, $f(y|M_i)$ is the marginal likelihood. $\frac{f(y|M_A)}{f(y|M_B)}$ and $\frac{f(M_A)}{f(M_B)}$ are called Bayes factor and prior odds, respectively. As is the usual practice, the prior odds is assumed to be 1, that is the prior probabilities are assumed to be equal between competing models, so that the posterior odds ratio is equal to the Bayes factor (Asai, 2009). The idea is compare the competing models using their posterior probabilities to select the one that is the best supported by the data. We choose the model that yields the largest posterior probability, or equivalently the largest marginal likelihood. Thus, we choose the model A if the posterior odds or Bayes factor is greater than 1, and we choose the model B if it is less than 1.

The marginal likelihood is defined by:

$$f(y|M_i) = \int f(y|M_i, \theta_i) f(\theta_i|M_i) d\theta_i, \tag{19}$$

this is, the integral of the likelihood with respect to the prior density of the parameter. To compute the logarithm of the marginal likelihood, we follow the log marginal likelihood identity of model M_i which is developed in Chib (1995), and that can be written as:

$$\log f(y|M_i) = \log f(y|M_i, \theta_i) + \log f(\theta_i|M_i) - \log f(\theta_i|M_i, y), \quad i = A,\ B, \tag{20}$$

where θ_i is the set of unknown parameters for model M_i, $f(y|M_i, \theta_i)$ is the likelihood of the model, $f(\theta_i|M_i)$ is the prior probability density, and $f(\theta_i|M_i, y)$ is the posterior probability density. The identity (20) holds for any value of θ_i, but following Chib (1995), Kim et al. (1998), Asai (2009), Nakajima (2012) and Nakajima and Omori (2012), we set θ_i at its posterior mean calculated using the MCMC samples to obtain a stable estimate of $f(y|M_i)$. 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 $f(y|M_i, \theta_i)$ can be estimated using the particle filter (see, for example, Pitt and Shephard (1999), Chib et al. (2002) and Omori et al. (2007)). For the posterior probability density $f(\theta_i|M_i, y)$, it can be estimated using the method developed by Chib (1995) and Chib and Jeliazkov (2001) using samples obtained through additional but reduced iterations of the MCMC algorithm.

First, we estimate de SVt model. Figures 16 – 20 show the MCMC estimation results of the SVt model for IGBVL, MERVAL, MEXBOL, IBOVESPA and IPSA stock returns data.

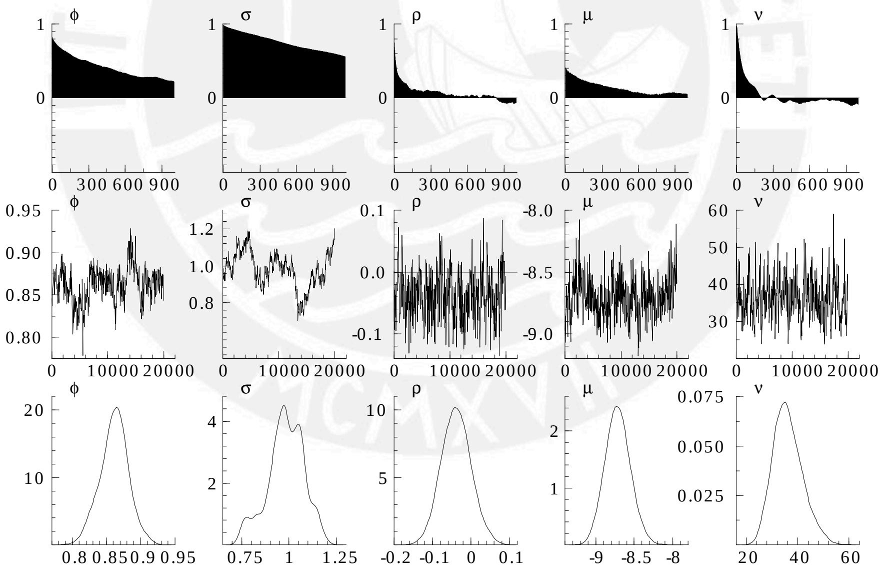

Figure 16. MCMC estimation results of the SVt model for IGBVL data (Peru). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

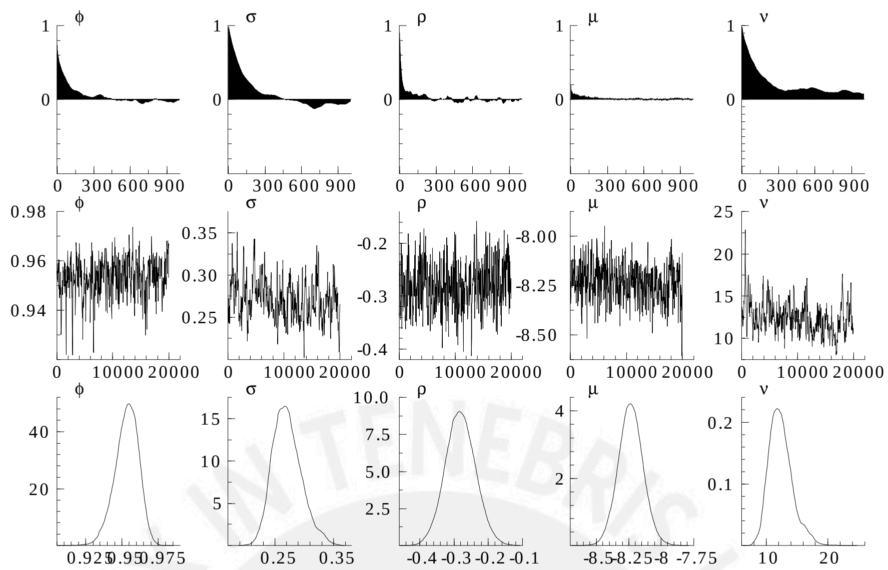

Figure 17. MCMC estimation results of the SVt model for MERVAL data (Argentina). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

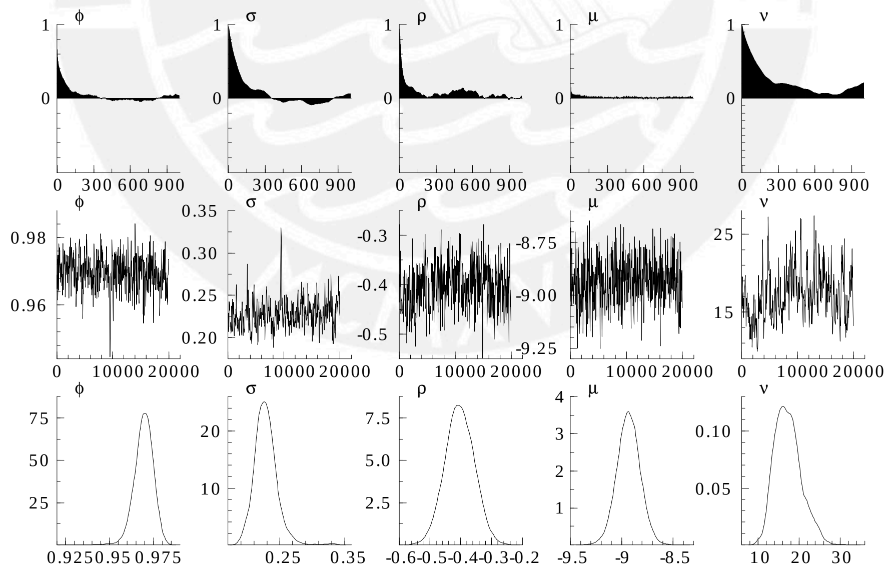

Figure 18. MCMC estimation results of the SVt model for MEXBOL data (Mexico). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

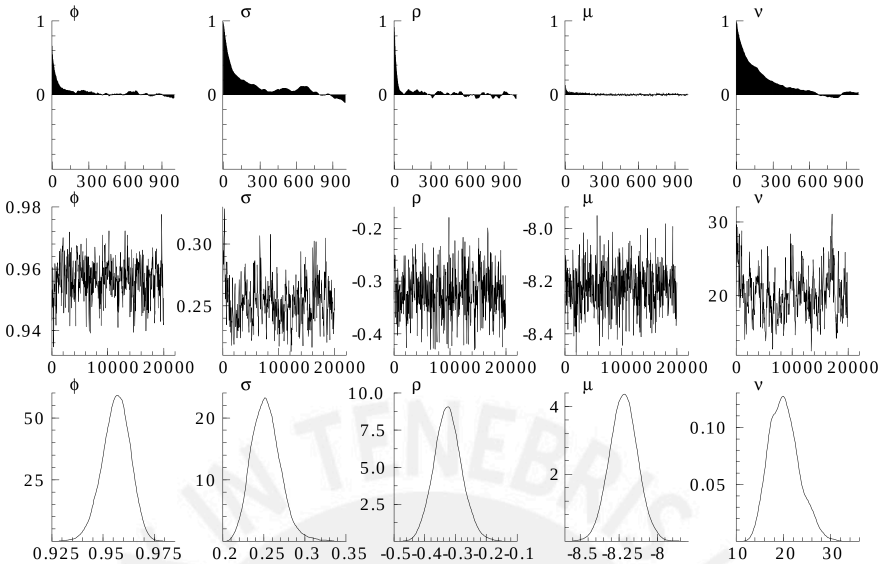

Figure 19. MCMC estimation results of the SVt model for IBOVESPA data (Brazil). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

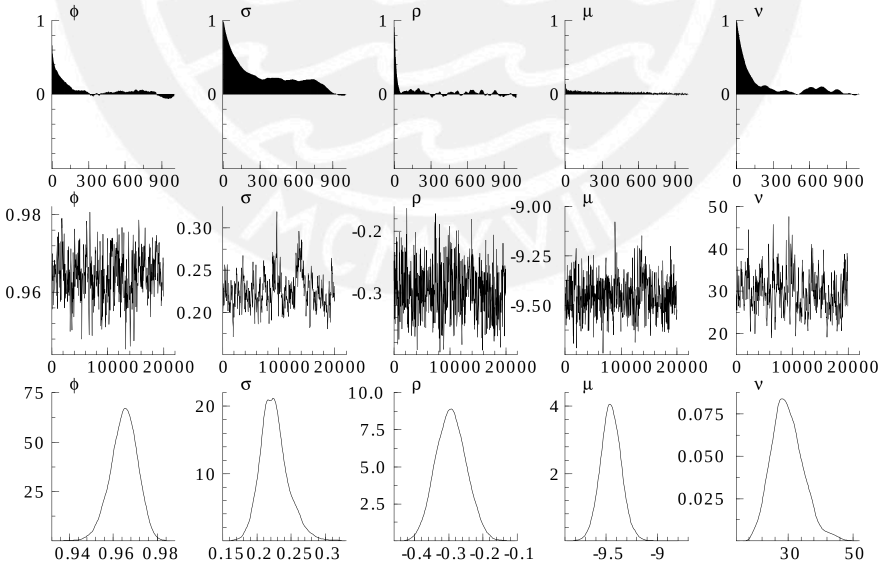

Figure 20. MCMC estimation results of the SVt model for IPSA data (Chile). Sample autocorrelations (top), sample paths (middle) and posterior densities (bottom).

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 stock return data. The posterior means of estimates parameter are very similar to the SVSKt model.

In order to compare the competing models, we estimate the log marginal likelihoods, $\log f(y|M_i)$, as follows: (i) the likelihood is estimated using the auxiliary particle filter with 10 000 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 $(\theta_i|M_i, y)$ is evaluated through 5 000 additional MCMC runs. Table 6 shows the estimates of the log marginal likelihoods and their standard errors. We choose the model that yields the largest log marginal likelihood. The SVSKt model outperforms the SVt model for IGBVL, MERVAL and IBOVESPA stock returns data and the SVt model outperforms the SVSKt model for MEXBOL and IPSA stock returns data. We can see that the *GH* Skew Student's t-disturbance distribution in the SV model (SVSKt model) is successful in describing the distribution of the daily stock return data for Peru, Argentina and Brazil and the symmetric Student's t-disturbance distribution in describing the distribution of the daily stock return data for MEXBOL and IPSA.

3.4 Prior Sensitivity Analysis

In spite of the computational expense of implementing, prior sensitivity analysis is an important tool in Bayesian inference because is important to assess the influence of the prior distribution on the final inference. In order to check prior sensitivity, the posterior distribution of parameters must be studied using a variety of prior distributions. As in Nakajima and Omori (2012), we are focusing on the skewness and heavy-taildness parameters, β and ν, to check the robustness of the model. We focus only on these parameters because we have assumed the values commonly used in the previous literature for the prior distributions of ϕ, σ, ρ and μ.

The prior sensitivity analysis take into account the following priors:

- Prior #1: $\beta \sim N(0,1)$, $\nu \sim Gamma(16, 0.8)\mathbf{1}(\nu > 4)$,
- Prior #2: $\beta \sim N(0,4)$, $\nu \sim Gamma(16, 0.8)\mathbf{1}(\nu > 4)$,
- Prior #3: $\beta \sim N(0,1)$, $\nu \sim Gamma(24, 0.6)\mathbf{1}(\nu > 4)$,
- Prior #4: $\beta \sim N(0,4)$, $\nu \sim Gamma(24, 0.6)\mathbf{1}(\nu > 4)$,
- Prior #5: $\beta \sim N(0,1)$, $\nu \sim Gamma(1.2, 0.03)\mathbf{1}(\nu > 4)$,

where the prior mean and prior standard deviation for $Gamma(16, 0.8)$, $Gamma(24, 0.6)$ and $Gamma(1.2, 0.03)$ are $(20, 5)$, $(40, 8)$ and $(40, 36.5)$, respectively. The prior #1 denote de prior distribution assumed in the previous estimations. The prior #5 for ν is rather flat compared to priors #1 to #4 and give less information on ν. Table 7 shows the parameter estimates: posterior means, the standard deviation, the 95% credible intervals and the inefficiency factors for β and ν.

Regarding IGBVL, we provide a prior sensitivity analysis focusing only in the priors #1, #2 and #5 because there are problems with the convergence of the chains of the MCMC algorithm with priors #3 and #4. The estimates for β are not affected by changing the priors considered. However, the estimates of ν (estimates are similar using prior #1 and #2) are affected by altering the prior for ν from prior #1 or prior #2 to prior #5. The estimates of ν get larger (from 36 to 161) implying lighter tails but similar skewness. The estimates of standard deviations for β and ν using the prior #5 are larger than the estimates using the prior #1 and #2.

Table 5. MCMC Estimation Results of the SVt Model for Latin American Stock Return Data

Parameter	Mean	S.D.	95% interval	Inefficiency
(i) IGBVL				
ϕ	0.8613	0.0209	[0.8183, 0.9024]	407.89
σ	0.9823	0.0964	[0.7655, 1.1565]	634.21
ρ	-0.0382	0.0386	[-0.1122, 0.0401]	104.68
μ	-8.7154	0.1639	[-9.0253, -8.3828]	156.46
ν	36.1646	5.7485	[26.2375, 48.8401]	105.50
(ii) MERVAL				
ϕ	0.9533	0.0081	[0.9358, 0.9674]	103.72
σ	0.2707	0.0244	[0.2279, 0.3256]	197.98
ρ	-0.2810	0.0434	[-0.3652, -0.1945]	52.60
μ	-8.2351	0.0948	[-8.4186, -8.0462]	20.95
ν	12.3573	1.9107	[9.2705, 16.8398]	253.81
(iii) MEXBOL				
ϕ	0.9694	0.0054	[0.9584, 0.9788]	74.32
σ	0.2282	0.0178	[0.1973, 0.2661]	160.80
ρ	-0.4037	0.0471	[-0.4951, -0.3126]	94.09
μ	-8.9333	0.1122	[-9.1537, -8.7119]	17.26
ν	17.2837	3.2351	[12.0182, 24.5254]	280.95
(iv) IBOVESPA				
ϕ	0.9564	0.0067	[0.9423, 0.9687]	59.67
σ	0.2528	0.0176	[0.2219, 0.2912]	169.99
ρ	-0.3244	0.0447	[-0.4091, -0.2330]	46.67
μ	-8.2198	0.0907	[-8.3970, -8.0400]	11.09
ν	20.1754	3.1531	[14.8120, 26.9762]	245.94
(v) IPSA				
ϕ	0.9649	0.0061	[0.9521, 0.9759]	85.17
σ	0.2253	0.0205	[0.1899, 0.2718]	283.75
ρ	-0.2928	0.0438	[-0.3754, -0.2059]	43.86
μ	-9.4536	0.1004	[-9.6466, -9.2478]	23.10
ν	29.8385	4.8740	[21.3740, 40.5813]	165.48

Table 6. Estimated Log Marginal Likelihoods (Log-ML)

(i) IGBVL		
SVSKt	9812.000	(1.582)
SVt	9781.776	(1.557)
(ii) MERVAL		
SVSKt	11469.963	(0.806)
SVt	11464.903	(0.684)
(iii) MEXBOL		
SVSKt	13377.385	(0.580)
SVt	13380.232	(0.704)
(iv) IBOVESPA		
SVSKt	11624.206	(0.996)
SVt	11614.349	(0.627)
(v) IPSA		
SVSKt	14583.355	(0.557)
SVt	14586.065	(0.653)

*Standard errors of the log-ML in parentheses.

MERVAL estimates for (β, ν) are not affected by changing the prior for β from prior #1 to prior #2, neither from prior #3 to prior #4. However, the estimates of (β, ν) are affected by altering the prior for ν from prior #1 to prior #3 (or from prior #2 to prior #4). The estimates of β get smaller (from -0.25 to -0.42) and the posterior means of ν get larger (from 12 to 20), implying greater skewness and lighter tails. The posterior standard deviations become larger reflecting the increase in the dispersion of the prior distribution for ν. Given less information on ν given by prior #5, the estimates of (β, ν) are similar to the estimates obtained by using priors #1 and #2.

MEXBOL estimates for β are not affected by changing the priors considered. The estimates of β are similar using the priors #1 - #5 (in the range from -0.1076 to -0.0753). However, the estimates of ν (estimates are similar using prior #1, #2 and #5) are affected by altering the prior for ν from prior #1 to prior #3 (or prior #2 to prior #4). The estimates of ν get larger (from 19.5 to 27) from prior #1, #2 and #5 to prior #3 and #4, implying lighter tails but similar skewness using the priors #3 to #4. The posterior standard deviations become larger from priors #1 and #2 to priors #3 and #4 and the estimates of standard deviations using prior #5 is the same for β comparing to priors #1 and #2 but larger for ν.

Regarding the IBOVESPA, the estimates for (β, ν) are not affected by changing the prior from prior #1 to prior #2 or to prior #5, however from prior #1, #2 or #5 to prior #3 (or from prior #1, #2 or #5 to prior #4) the estimates for (β, ν) are largely affected. The estimates of β and ν get larger from -0.03 to 0.07 and from 17.5 to 29 (average), respectively, implying a disturbance density that becomes less skewed and has lighter tails. The posterior standard deviations of (β, ν) become larger from prior #1, #2 or #5 to prior #3 (or from prior #1, #2 or #5 to prior #4).

Finally, the IPSA estimates for β are not largely affected by changing the priors considered. However, the estimates of ν are affected by altering the priors from prior #1 or #2 (the estimates for ν are similar with these priors) to prior #3 or #4 (the estimates of ν also are similar with these priors) or to prior #5. The posterior means of ν are 30.2 using the priors #1 and #2, 50 (average) using the priors #3 and #4, and 106.4 using the prior #5. This fact implies lighter tails but similar skewness. The posterior standard deviations of (β, ν) become larger from prior #1, #2 to prior #3 and #4, and the prior #5 has the largest posterior standard deviation (0.43 for β and 36.96 for ν).

Table 7. Prior sensitivity analysis for the SVSKt model

	Prior #1		Prior #2		Prior #3		Prior #4		Prior #5	
(i) IGBVL										
	-0.0286	(0.1508)	-0.0287	(0.1489)					-0.0274	0.3105
β	[-0.3247,	0.2678]	[-0.3203,	0.2675]					[-0.6400,	0.5881]
	16.87		15.71						9.72	
	35.6892	(5.4920)	36.7286	(5.5969)					161.0542	56.3557
ν	[25.9746,	47.2584]	[26.5760,	48.3670]					[76.0663,	297.5268]
	88.54		114.07						291.47	
(ii) MERVAL										
	-0.2464	(0.0811)	-0.2588	(0.0847)	-0.4123	(0.1556)	-0.4222	(0.1507)	-0.2522	(0.0826)
β	[-0.4180,	-0.0952]	[-0.4372,	-0.1019]	[-0.7668,	-0.1460]	[-0.7691,	-0.1748]	[-0.4309,	-0.1076]
	82.38		73.46		106.15		146.31		91.94	
	12.2135	(1.7827)	12.4566	(1.8322)	20.1648	(4.5284)	19.0319	(4.2291)	11.1433	(1.8805)
ν	[9.0445,	15.9292]	[9.4167,	16.4069]	[12.9397,	30.7610]	[12.7557,	28.7367]	[8.0934,	15.0799]
	286.25		320.19		379.97		410.92		309.20	
(iii) MEXBOL										
	-0.1076	(0.1194)	-0.1008	(0.1110)	-0.0818	(0.1559)	-0.0753	(0.1574)	-0.0855	(0.1194)
β	[-0.3389,	0.1326]	[-0.3168,	0.1224]	[-0.3795,	0.2325]	[-0.3754,	0.2402]	[-0.3099,	0.1692]
	37.20		36.08		40.87		50.87		59.73	
	19.8882	(3.2235)	18.1714	(3.0561)	27.6532	(5.1416)	26.5082	(5.0933)	19.5090	(5.3784)
ν	[14.2966,	27.2299]	[13.4678,	25.6527]	[19.1554,	39.1847]	[17.7599,	38.1351]	[12.7685,	34.4924]
	265.91		260.93		287.83		270.82		469.37	
(iv) IBOVESPA										
	-0.0342	(0.1107)	-0.0374	(0.1137)	0.0774	(0.2088)	0.0649	(0.1915)	-0.0366	(0.1273)
β	[-0.2437,	0.1954]	[-0.2449,	0.2044]	[-0.3077,	0.5243]	[-0.2755,	0.4832]	[-0.2672,	0.2518]
	38.07		44.98		50.19		63.74		94.28	
	17.5026	(2.6309)	17.6462	(3.0990)	31.4925	(6.4575)	28.3000	(5.6489)	18.6492	(4.5117)
ν	[13.1482,	23.3003]	[12.7563,	24.9240]	[19.5521,	46.0880]	[19.3273,	40.6572]	[13.0256,	30.7739]
	159.85		276.67		309.95		252.38		396.59	
(v) IPSA										
	-0.0852	(0.1872)	-0.0891	(0.1957)	-0.0808	(0.2880)	-0.1129	(0.2703)	-0.1381	(0.4383)
β	[-0.4475,	0.2858]	[-0.4835,	0.2834]	[-0.6762,	0.4567]	[-0.6307,	0.4476]	[-1.0289,	0.7027]
	37.72		56.60		56.82		25.53		36.74	
	30.2523	(5.1130)	30.2391	(4.7554)	49.2582	(8.4254)	50.6228	(9.0323)	106.4110	(36.9622)
ν	[21.1299,	41.1597]	[22.2936,	40.9700]	[34.8189,	67.3309]	[35.1501,	70.4570]	[48.2143,	189.1728]
	166.06		187.32		115.43		194.78		327.33	

The first row: posterior mean and standard deviation in parenthesis.
The second row: 95% credible intervals in square brackets.
The third row: inefficiency factor.

As in Nakajima and Omori (2012), we also observed that the posterior estimate of ν is sensitive to the choice of the prior distribution for ν and the posterior estimate of β is also sensitive to the choice of the prior distribution for ν because the skewness and heavy-tailedness of the *GH* skew Student's *t*-distribution are determined by β and ν simultaneously and not individually.

4 Conclusions

In this paper, we estimate a SV model incorporating both leverage effects and skewed heavy-tailed disturbances taking into account the *GH* Skew Student's *t*-distribution for a set of Latin American stock market indices using the Bayesian estimation method proposed by Nakajima and Omori (2012). We apply the SVSKt model to daily returns of five Latin American stock market indices: IGBVL, MERVAL, MEXBOL, IPSA and IBOVESPA and we also analyze the U.S. S&P500 returns to compare the results. The SVSKt model can be considered a flexible model to fit the returns and volatility characteristics because the SVSKt model is able to model substantially skewed and heavy tailed data and includes the SV model with Normal disturbances (SV-Normal) and the SV model with symmetric Student's *t*-disturbances (SVt).

The MCMC estimation results of the SVSKt model show that the sample paths of the iterations of parameters are stable, and the proposed estimation scheme works well for all indices except for the IGBVL (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. However, the results show that the IGBVL returns have low persistence in comparison to the volatility of the others stock indices of Latin American considered.

The results support the evidence that there are leverage effects in all indices considered but there is not enough evidence for the IGBVL. The estimates show that the leverage effect is more notable in MEXBOL and IBOVESPA, followed by MERVAL and IPSA. In the case of the IGBVL, the posterior mean estimate of ρ is also 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. Also, the results support the evidence of skewed heavy-tailed disturbances only for the MERVAL, symmetric heavy-tailed disturbances for the MEXBOL, IBOVESPA and IPSA, and symmetric Normal disturbances for the IGBVL.

Finally, volatility estimates for daily stocks returns show a similar pattern between them for the sample period considered. On the other hand, the model comparison between SVSKt and SVt model show that the SVSKt model outperforms the SVt model for IGBVL, MERVAL and IBOVESPA and the SVt model outperforms the SVSKt model for MEXBOL and IPSA.

Appendix

A. The *GH* Skew Student *t*-Distribution

This appendix includes some important properties of the *GH* skew Student-*t* distribution. For a more complete treatment, see Prause (1999), Eberlein and Hammerstein (2004) and Aas and Haff (2006). The *GH* Skew Student *t*-Distribution is a limiting case of the *GH* distribution, which was introduced in Barndorff-Nielsen (1977). The univariate *GH* distribution can be parameterized in several ways. Following, Prause (1999), Eberlein and Hammerstein (2004) and Aas and Haff (2006), the probability density function of a *GH* random variable x is given by:

$$f_{GH}(x;\lambda,\delta,\alpha,\mu_x,\beta) = \frac{(\alpha^2-\beta^2)^{\lambda/2}\mathbf{K}_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu_x)^2}\right)\exp\left(\beta(x-\mu_x)\right)}{\sqrt{2\pi}\alpha^{\lambda-\frac{1}{2}}\delta^{\lambda}\mathbf{K}_{\lambda}\left(\delta\sqrt{\alpha^2-\beta^2}\right)\left(\sqrt{\delta^2+(x-\mu_x)^2}\right)^{\frac{1}{2}-\lambda}} \qquad \text{(A.1)}$$

where $\mathbf{K}_j$ is the modified Bessel function[7] of the third kind of order j and $x \in \mathbb{R}$. $\alpha > 0$ determines the shape, $0 \leq |\beta| < \alpha$ determines the skewness, $\mu_x \in \mathbb{R}$ is a location parameter and $\delta > 0$ serves for the scaling. $\lambda \in \mathbb{R}$ characterizes certain subclasses and considerably influences the size of mass contained in the tails . The parameters must satisfy the conditions:

[7] The modified Bessel function of the third kind with order j, which we denote as $\mathbf{K}_j(\cdot)$, has the integral representation:

$$\mathbf{K}_j(x) = \frac{1}{2}\int_0^\infty w^{\lambda-1}\exp\left\{-\frac{1}{2}x(w+w^{-1})dw\right\}, \; x > 0.$$

Some properties of $\mathbf{K}_j(\cdot)$, taken from Abramowitz and Stegun (1972), are:

- $\mathbf{K}_j(x) = \mathbf{K}_{-j}(x)$,
- An asymptotic relations for small arguments x is given by :

$$\begin{aligned} \mathbf{K}_j(x) &\sim \Gamma(j)2^{j-1}x^{-j} \text{ as } x \to 0 \text{ and } j > 0, \\ \mathbf{K}_j(x) &\sim \Gamma(-j)2^{-j-1}x^{j} \text{ as } x \to 0 \text{ and } j < 0, \\ \mathbf{K}_0(x) &\sim -\ln(x). \end{aligned}$$

- An asymptotic relation for large arguments x is given by:

$$\mathbf{K}_j(x) \sim \sqrt{\frac{\pi}{2x}}\exp(-x) \text{ as } x \to \infty.$$

- If $j = n + \frac{1}{2}$, $n \in \mathbb{Z}$, $\mathbf{K}_j$ can be calculated as follows:

$$\mathbf{K}_{n+\frac{1}{2}}(x) = \sqrt{\frac{\pi}{2x}}\exp(-x)\left[1+\sum_{i=1}^{n}\frac{(n+i)!}{(n-i)!i!}(2x)^{-i}\right], \; n \in \mathbb{N},$$

- $\mathbf{K}_{-0.5}(x) = \mathbf{K}_{0.5}(x) = \sqrt{\frac{\pi}{2x}}\exp(-x)$.

$$\begin{aligned} \delta &\geq 0, |\beta| < \alpha \text{ if } \lambda > 0, \\ \delta &> 0, |\beta| < \alpha \text{ if } \lambda = 0, \\ \delta &> 0, |\beta| \leq \alpha \text{ if } \lambda < 0. \end{aligned} \tag{A.2}$$

The tails of the GH distribution behave as:

$$f_{GH}(x) \sim const\,|x|^{\lambda-1} \exp(-\alpha\,|x| + \beta x) \; as \; x \to \pm\infty, \; \forall\lambda, \tag{A.3}$$

hence, as long as $|\beta| \neq \alpha$, the GH distribution has two semiheavy tails.

The GH distribution may be represented as a Normal mean-variance mixture with Generalized Inverse Gaussian (GIG) distribution as a mixing distribution. This means that a GH variable X can be represented as:

$$X = \mu_X + \beta Z + \sqrt{Z}\epsilon, \qquad \epsilon \sim N(0,1), \qquad Z \sim GIG(\lambda, \delta, \gamma), \tag{A.4}$$

with ϵ and Z independent and $\gamma = \sqrt{\alpha^2 - \beta^2}$. It follows from (A.4) that $X \mid Z = z \sim N(\mu_X + \beta Z, Z)$. The density of the GIG distribution is given by:

$$f_{GH}(z; \lambda, \delta, \gamma) = \left(\frac{\gamma}{\delta}\right)^{\lambda} \frac{z^{\lambda-1}}{2\mathbf{K}_{\lambda}(\gamma\delta)} \exp(-\frac{1}{2}(\delta^2 z^{-1} + \gamma^2 z)). \tag{A.5}$$

Letting $\lambda = -\nu/2(\nu > 0)$ and $\alpha \to |\beta|$ in equation (A.1) (this is $\gamma = 0$), we obtain the GH Skew Student t-Distribution. Its probability density function is given by:

$$f_{GHskewt}(x; \nu, \delta, \mu_x, \beta) = \frac{2^{\frac{1-\nu}{2}} \delta^{\nu} \, |\beta|^{\frac{\nu+1}{2}} \, \mathbf{K}_{\frac{\nu+1}{2}}\left(\sqrt{\beta^2\left(\delta^2 + (x-\mu_x)^2\right)}\right) \exp\left(\beta\,(x-\mu_x)\right)}{\Gamma(\frac{\nu}{2})\sqrt{\pi}\left(\sqrt{\delta^2 + (x-\mu_x)^2}\right)^{\frac{\nu+1}{2}}}, \; \beta \neq 0, \tag{A.6}$$

and

$$f_{GHskewt}(x; \nu, \delta, \mu_\omega) = \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\pi}\delta\Gamma(\frac{\nu}{2})} \left[1 + \frac{(x-\mu_x)^2}{\delta^2}\right]^{-(\nu+1)/2}, \; \beta = 0. \tag{A.7}$$

where $\Gamma(.)$ is the gamma function. The density $f_{GHskewt}(\omega; \nu, \delta, \mu_\omega)$ in (A.7) is known as the noncentral Student's t-distribution with ν degrees of freedom, expectation μ_x, and variance $\delta^2/(\nu - 2)$.

The first four moments of a *GH* skew Student t-distributed random variable X are:

$$E(X) = \mu + \frac{\beta\delta^2}{\nu-2}, \tag{A.8}$$

$$Var(X) = \frac{2\beta^2\delta^4}{(\nu-2)^2(\nu-4)} + \frac{\delta^2}{\nu-2}, \tag{A.9}$$

$$Skewness(X) = \frac{2(\nu-4)^{1/2}\beta\delta}{\left[2\beta^2\delta^2+(\nu-2)(\nu-4)\right]^{3/2}}\left[3(\nu-2)+\frac{8\beta^2\delta^2}{\nu-6}\right], \tag{A.10}$$

$$Kurtosis(X) = \frac{6}{\left[2\beta^2\delta^2+(\nu-2)(\nu-4)\right]^2} \times \left[(\nu-2)^2(\nu-4)+\frac{16\beta^4\delta^2(\nu-2)(\nu-4)}{\nu-6}+\frac{8\beta^4\delta^4(5\nu-22)}{(\nu-6)(\nu-8)}\right]. \tag{A.11}$$

We observe that for the mean and variance to exist, $\nu > 2$ and $\nu > 4$, respectively. The variance is only finite when $\nu > 4$, as opposed to the symmetric Student's t-distribution. Skewness and (excess) kurtosis are defined only if $\nu > 6$ and $\nu > 8$, respectively.

It follows from equation (A.3) that in the tails, the *GH* skew t-density is given by:

$$f_{GHskewt}(x) \sim const\,|x|^{-\nu/2-1}\exp(-|\beta x|+\beta x) \text{ as } x \to \pm\infty, \tag{A.12}$$

Thus we have a heavy tail decaying as:

$$f_{GHskewt}(x) \sim const\,|x|^{-\nu/2-1} \text{ if } \left\{ \begin{array}{l} \beta < 0 \text{ and } x \to -\infty \\ \beta > 0 \text{ and } x \to +\infty \end{array} \right., \tag{A.13}$$

and a light tail decaying as

$$f_{GHskewt}(x) \sim const\,|x|^{-\nu/2-1}\exp(-2\,|\beta x|) \text{ if } \left\{ \begin{array}{l} \beta < 0 \text{ and } x \to +\infty \\ \beta > 0 \text{ and } x \to -\infty \end{array} \right.. \tag{A.14}$$

Thus the *GH* skew t-distribution has one heavy and one semiheavy tail. The heavy tail shows polynomial and the light tail exponential behavior. It is the only member of *GH* family of distributions having this property. Thus the *GH* skew student 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 student t-distribution are characterized uniquely by parameters β and ν, which determine jointly the degree of skewness and heavy tailedness. Finally, note that the heavy tail of the *GH* skew student t-distribution is heavier than the tails of the symmetric Student t-distribution, which have two tails decaying as polynomials and decay as:

$$f_{GHt}(x) \sim const\,|x|^{-\nu-1} \text{ as } x \to \pm\infty. \tag{A.15}$$

B. MCMC Algorithm for the SVSKt Model

This appendix includes each sampling step in detail of the MCMC algorithm proposed by Nakajima and Omori (2012) for the SVSKt model. For the prior distributions of μ and β, they assume $\mu \sim N(\mu_0, \nu_0^2)$ and $\beta \sim N(\beta_0, \sigma_0^2)$.

B.1 Generation of the parameters (ϕ, σ, ρ, μ) (steps 2–4)

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

$$\begin{aligned}\pi(\phi|.) &\propto \pi(\phi)\sqrt{1-\phi^2}\exp\left\{-\frac{(1-\phi^2)\overline{h}_1^2}{2\sigma^2} - \sum_{t=1}^{n-1}\frac{(\overline{h}_{t+1}-\phi\overline{h}_t-\overline{y}_t)^2}{2\sigma^2(1-\rho^2)}\right\} \\ &\propto \pi(\phi)\sqrt{1-\phi^2}\exp\left\{-\frac{(\phi-\mu_\phi)^2}{2\sigma_\phi^2}\right\}, \end{aligned} \tag{A.16}$$

where $\overline{h}_t = h_t - \mu$, $\overline{y}_t = \rho\sigma(y_t e^{-h_t/2} - \beta\overline{z}_t)/\sqrt{z_t}$, $\overline{z}_t = z_t - \mu_z$, $\mu_\phi = \frac{\sum_{t=1}^{n-1}(\overline{h}_{t+1}-\overline{y}_t)\overline{h}_t}{\rho^2\overline{h}_1^2+\sum_{t=2}^{n-1}\overline{h}_t^2}$ and $\sigma_\phi^2 = \frac{\sigma^2(1-\rho^2)}{\rho^2\overline{h}_1^2+\sum_{t=2}^{n-1}\overline{h}_t^2}$.

In order to sample from this conditional posterior distribution, Nakajima and Omori (2012) implement the Metropolis–Hastings (MH) algorithm. They propose a candidate, $\phi^* \sim TN_{(-1,1)}(\mu_\phi, \sigma_\phi^2)$, where $TN_{(a,b)}(\mu, \sigma^2)$ denotes the Normal distribution with mean μ and variance σ^2 truncated on the interval (a, b). Then, they accept it with the probability given by $\min\left\{\frac{\pi(\phi^*)\sqrt{1-\phi^{*2}}}{\pi(\phi)\sqrt{1-\phi^2}}, 1\right\}$.

Step 3. Because the joint conditional posterior probability density $\pi(\vartheta \mid \phi, \mu, \nu, h, z, y)(\equiv \pi(\vartheta|\cdot))$ of $\vartheta = (\sigma, \rho)'$ is given by $\pi(\vartheta|\cdot) \propto \pi(\vartheta)\sigma^n(1-\rho^2)^{\frac{n-1}{2}}\exp\left\{-\frac{(1-\phi^2)\overline{h}_1^2}{2\sigma^2} - \sum_{t=1}^{n-1}\frac{(\overline{h}_{t+1}-\phi\overline{h}_t-\overline{y}_t)^2}{2\sigma^2(1-\rho^2)}\right\}$, 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 there is a constraint, $R = \{\vartheta : \sigma > 0, |\rho| < 1\}$, on the parameter space of the posterior distribution, they consider the transformation ϑ to $\omega = (\omega_1, \omega_2)'$, where $\omega_1 = \log\sigma$, and $\omega_2 = \log(1+\rho) - \log(1-\rho)$, to generate a candidate using a Normal distribution. They first search for $\widehat{\vartheta}$ that approximately maximizes $\pi(\vartheta|.)$, and obtain its transformed value $\widehat{\omega}$. They next generate a candidate $\omega^* \sim N(\omega_*, \Sigma_*)$, where $\omega_* = \widehat{\omega} + \Sigma_* \left.\frac{\partial\log\widetilde{\pi}(\omega|\cdot)}{\partial\omega}\right|_{\omega=\widehat{\omega}}$ and $\Sigma_*^{-1} = -\left.\frac{\partial^2\log\widetilde{\pi}(\omega|\cdot)}{\partial\omega\partial\omega'}\right|_{\omega=\widehat{\omega}}$, where $\widetilde{\pi}(\omega|\cdot)$ is a transformed conditional posterior density. Then, they accept the candidate ω^* with probability $\min\left\{\frac{\pi(\vartheta^*|\cdot)f_N(\omega|\omega_*,\Sigma_*)|J(\vartheta)|}{\pi(\vartheta|\cdot)f_N(\omega^*|\omega_*,\Sigma_*)|J(\vartheta^*)|}, 1\right\}$, where $f_N(x|\mu,\Sigma)$ denotes the probability density function of a Normal distribution with mean μ and covariance matrix Σ, and $J(\cdot)$ is the Jacobian for the transformation, that is, $J(\vartheta) = |\frac{d\omega}{d\vartheta'}|_+ = \frac{2}{\sigma(1-\rho^2)}$. The values of (ϑ, ϑ^*) are evaluated at (ω, ω^*), respectively.

Step 4. The conditional posterior probability density $\pi(\mu \mid \phi, \sigma, \rho, \beta, \nu, h, z, y)(\equiv \pi(\mu|\cdot))$ is given by $\pi(\mu|\cdot) \propto \exp\left\{-\frac{(\mu-\mu_0)^2}{2\nu_0^2} - \frac{(1-\phi^2)\bar{h}_1^2}{2\sigma^2} - \sum_{t=1}^{n-1} \frac{\{(h_{t+1}-\mu)-\phi(h_t-\mu)-\bar{y}_t\}^2}{2\sigma^2(1-\rho^2)}\right\}$, from which Nakajima and Omori (2012) generate $\mu|\cdot \sim N(\widehat{\mu}, \sigma_\mu^2)$, where $\sigma_\mu^2 = \left\{\frac{1}{\nu_0^2} + \frac{(1-\rho^2)(1-\phi^2)+(n-1)(1-\phi)^2}{\sigma^2(1-\rho^2)}\right\}^{-1}$, and $\widehat{\mu} = \sigma_\mu^2\left\{\frac{\mu_0}{\nu_0^2} + \frac{(1-\rho^2)(1-\phi^2)h_1+(1-\phi)\sum_{t=1}^{n-1}(\bar{h}_{t+1}-\phi\bar{h}_t-\bar{y}_t)^2}{\sigma^2(1-\rho^2)}\right\}$.

B.2 Generation of skew-t parameters (β, ν, z) (steps 5–7)

Step 5. The posterior probability density $\pi(\beta \mid \phi, \sigma, \rho, \mu, \nu, h, z, y)(\equiv \pi(\beta|\cdot))$ is given by $\pi(\beta|\cdot) \propto \exp\left\{-\frac{(\beta-\beta_0)^2}{2\sigma_0^2} - \sum_{t=1}^{n} \frac{(y_t-\beta\bar{z}_t e^{h_t/2})^2}{2z_t e^{h_t}} - \sum_{t=1}^{n-1} \frac{\{\bar{h}_{t+1}-\phi\bar{h}_t-\rho\sigma(y_t e^{-h_t/2}-\beta\bar{z}_t)/\sqrt{z_t}\}^2}{2\sigma^2(1-\rho^2)}\right\}$, from which they generate $\beta|\cdot \sim N(\mu_\beta, \sigma^2)$ where $\sigma_\beta^2 = \left\{\frac{1}{\sigma_0^2} + \frac{1}{1-\rho^2}\sum_{t=1}^{n-1} \frac{\bar{z}_t^2}{z_t} + \frac{\bar{z}_n^2}{z_n}\right\}^{-1}$, and
$\mu_\beta = \sigma_\beta^2\left\{\frac{\beta}{\sigma_0^2} + \frac{1}{1-\rho^2}\sum_{t=1}^{n-1} \frac{y_t\bar{z}_t}{z_t e^{h_t/2}} + \frac{y_n\bar{z}_n}{z_n e^{h_n/2}} - \frac{\rho}{\sigma(1-\rho^2)}\sum_{t=1}^{n-1} \frac{(\bar{h}_{t+1}-\phi\bar{h}_t)\bar{z}_t}{\sqrt{z}t}\right\}$.

Step 6. Because, as in Step 3, it is not easy to sample directly from the posterior probability density of ν, $\pi(\nu|\cdot) \propto \pi(\nu)\prod_{t=1}^{n} \frac{(\nu/2)^{\nu/2}}{\Gamma(\nu/2)} z_t^{-\nu/2} \exp(-\frac{\nu}{2zt}) \exp\left\{-\sum_{t=1}^{n} \frac{(y_t-\beta\bar{z}_t e^{h_t/2})^2}{2z_t e^{h_t}} - \sum_{t=1}^{n-1} \frac{(\bar{h}_{t+1}-\phi\bar{h}_t-\bar{y}_t)^2}{2\sigma^2(1-\rho^2)}\right\}$, for $\nu > 4$, they draw a sample of ν using the MH algorithm based on the Normal approximation of the posterior probability density. They generate a candidate ν^* using a Normal distribution truncated on $(4, \infty)$.

Step 7. The conditional posterior probability density of the latent variable z_t is $\pi(z_t \mid \theta, h, y) \propto g(z_t) \times z_t^{-(\frac{\nu+1}{2}+1)} \exp(-\frac{\nu}{2z_t})$, and $g(z_t) = \exp\left\{-\frac{(y_t-\beta\bar{z}_t e^{h_t/2})^2}{2z_t e^{h_t}} - \frac{(\bar{h}_{t+1}-\phi\bar{h}_t-\bar{y}_t)^2}{2\sigma^2(1-\rho^2)}\mathbf{1}(t<n)\right\}$, where $\mathbf{1}(\cdot)$ is an indicator function. Using the MH algorithm, they generate a candidate $z_t^* \sim IG(\frac{\nu+1}{2}, \frac{\nu}{2})$ and accept it with probability $\min\{\frac{g(z_t^*)}{g(z_t)}, 1\}$.

B.3 Generation of volatility latent variable h (step 8)

Step 8. Nakajima and Omori (2012) extend the method developed by Omori and Watanabe (2008) for sampling h_t in the SVSKt model using the multi-move sampler, where the 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. The details of the multi-move sampler are described in the Appendix C.

C. Multi-Move Sampler for the SVSKt Model

Extending the algorithm of Omori and Watanabe (2008), Nakajima and Omori (2012) describe the multi-move sampler for sampling the volatility variable h in the SVSKt model. Defining $\alpha_t = h_t - \mu$, for $t = 0, \ldots, n$ and $\gamma = \exp(\mu/2)$, they consider the state space model with respect to $\{\alpha_t\}_{t=1}^n$ as:

$$
\begin{aligned}
y_t &= \{\beta \bar{z}_t + \sqrt{z_t}\epsilon_t\} \exp(\alpha_t/2), & t = 1, \ldots, n, \quad \text{(A.17)}\\
\alpha_{t+1} &= \phi \alpha_t + \eta_t, & t = 0, \ldots, n-1. \quad \text{(A.18)}
\end{aligned}
$$

Let $\widetilde{\Theta} = (\theta, \alpha_r, \alpha_{r+d+1}, z_r, \ldots, z_{r+d}, y_r, \ldots, y_{r+d})$. To sample a block $(\alpha_{r+1}, \ldots, \alpha_{r+d})$ from its joint conditional posterior density using MH algorithm, $(r \geq 0, d \geq 1, r+d \leq n)$, they sample disturbances

$$
(\eta_r, \ldots, \eta_{r+d-1}) \sim \pi(\eta_r, \ldots \eta_{r+d-1} \mid \widetilde{\Theta}) \propto \prod_{t=r}^{r+d} \frac{1}{\sqrt{2\pi}\widetilde{\sigma}_t} \exp\left\{-\frac{(y_t - \widetilde{\mu}_t)^2}{2\widetilde{\sigma}_t^2}\right\} \times \prod_{t=r}^{r+d-1} f(\eta_t) \times f(\alpha_{r+d}),
$$

where $\widetilde{\mu}_t = \left\{\beta \bar{z}_t + \rho_t \sqrt{z_t}(\alpha_{t+1} - \phi\alpha_t)/\sigma\right\} \exp(\alpha_t/2)\gamma$, $\widetilde{\sigma}_t^2 = (1-\rho_t^2) z_t \exp(\alpha_t)\gamma^2$, $f(\alpha_{r+d}) = \exp\left\{-\frac{(\alpha_{r+d+1} - \phi\alpha_{r+d})^2}{2\sigma^2}\right\} \mathbf{1}(r+d < n)$, and $\rho_t = \rho \mathbf{1}(r+d < n)$. To determine the block (r and d), they use the stochastic knots (see, for example, Shephard and Pitt (1997)). Let $\underline{\eta} = (\eta_r, \ldots, \eta_{r+d-1})'$ and $\underline{\alpha} = (\alpha_{r+1}, \ldots, \alpha_{r+d})'$. To construct a proposal density based on the Normal approximation of the posterior density of $\underline{\eta}$, they first define:

$$
\begin{aligned}
L &= \sum_{t=r}^{r+d} \left\{-\frac{\alpha_t}{2} - \frac{(y_t - \widetilde{\mu}_t)^2}{2\widetilde{\sigma}_t^2}\right\} + \log f(\alpha_{r+d}),\\
\delta &= (\delta_{r+1}, \ldots, \delta_{r+d})', \qquad \delta_t = \frac{\partial L}{\partial \alpha_t},\\
Q &= -E\left(\frac{\partial^2 L}{\partial \underline{\alpha} \partial \underline{\alpha}'}\right) = \begin{bmatrix} A_{r+1} & B_{r+2} & 0 & \cdots & 0 \\ B_{r+2} & A_{r+2} & B_{r+3} & \cdots & 0 \\ 0 & B_{r+3} & A_{r+3} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & B_{r+d} \\ 0 & \cdots & 0 & B_{r+d} & A_{r+d} \end{bmatrix},\\
A_t &= -E\left(\frac{\partial^2 L}{\partial \alpha_t^2}\right),\\
B_t &= -E\left(\frac{\partial^2 L}{\partial \alpha_t \partial \alpha_{t-1}}\right),
\end{aligned}
$$

for $t = r+2, \ldots, r+d$, and $B_{r+1} = 0$. For the second derivatives, they take the expectations with respect to y_t's and obtain

$$
\begin{aligned}
A_t &= \frac{1}{2} + \frac{1}{\widetilde{\sigma}_t^2}\left(\frac{\partial \widetilde{\mu}_t}{\partial \alpha_t}\right)^2 + \frac{1}{\widetilde{\sigma}_{t-1}^2}\left(\frac{\partial \widetilde{\mu}_{t-1}}{\partial \alpha_t}\right)^2 + \frac{\phi^2}{\sigma^2}.\mathbf{1}(t{=}r+d < n), \text{ and}\\
B_t &= \frac{1}{\widetilde{\sigma}_{t-1}^2} \frac{\partial \widetilde{\mu}_{t-1}}{\partial \alpha_{t-1}} \frac{\partial \widetilde{\mu}_{t-1}}{\partial \alpha_t}.
\end{aligned}
$$

Applying the second-order Taylor expansion to the log of the posterior density around the mode, $\underline{\eta} = \widehat{\underline{\eta}}$, they obtain an approximate Normal density as follows:

$$\begin{aligned}
\log \pi(\underline{\eta}|\widetilde{\Theta}) &\approx \widehat{L} + \frac{\partial L}{\partial \underline{\eta}'}\bigg|_{\underline{\eta}=\widehat{\underline{\eta}}} (\underline{\eta} - \widehat{\underline{\eta}}) + \frac{1}{2}(\underline{\eta} - \widehat{\underline{\eta}})' \, E(\frac{\partial^2 L}{\partial \underline{\eta} \partial \underline{\eta}'})\bigg|_{\underline{\eta}=\widehat{\underline{\eta}}} (\underline{\eta} - \widehat{\underline{\eta}}) + \sum_{t=r}^{r+d-1} (-\frac{1}{2}\eta_t^2) + (const.) \\
&= \widehat{L} + \widehat{\underline{\delta}}'(\underline{\alpha} - \widehat{\underline{\alpha}}) - \frac{1}{2}(\underline{\alpha} - \widehat{\underline{\alpha}})'\widehat{Q}(\underline{\alpha} - \widehat{\underline{\alpha}}) + \sum_{t=r}^{r+d-1} (-\frac{1}{2}\eta_t^2) + (const.) \\
&\equiv \log q(\underline{\eta}|\widetilde{\Theta}),
\end{aligned}$$

where $\widehat{L}$, $\widehat{\delta}$ and $\widehat{Q}$ is the value of L, δ and Q at $\underline{\alpha} = \widehat{\underline{\alpha}}$ (or, equivalently at $\underline{\eta} = \widehat{\underline{\eta}}$). It can be shown that the proposal density $q(\underline{\eta}|\widetilde{\Theta})$ is the posterior density of $\underline{\eta}$ for a linear Gaussian state space model given by (A.19)- (A.21). The mode $\widehat{\underline{\eta}}$ can be obtained by repeating the following algorithm until it converges:

1. Initialize $\widehat{\underline{\eta}}$ and compute $\widehat{\underline{\alpha}}$ at $\underline{\eta} = \widehat{\underline{\eta}}$ using the state equation (A.18) recursively.

2. Evaluate $\widehat{\underline{\delta}}$'s, $\widehat{A}_t$'s and $\widehat{B}_t$'s at $\underline{\alpha} = \widehat{\underline{\alpha}}$,

3. Let $\widehat{D}_{r+1} = \widehat{A}_{r+1}$ and $\widehat{b}_{r+1} = \widehat{\delta}_{r+1}$. Compute the following variables recursively for $t = r+2, \ldots, r+d$:
$$\widehat{D}_t = \widehat{A}_t - \widehat{D}_{t-1}^{-1}\widehat{B}_t^2, \; \widehat{K}_t = \sqrt{\widehat{D}_t}, \; \widehat{b}_t = \widehat{\delta}_t - \widehat{B}_t\widehat{D}_{t-1}^{-1}\widehat{b}_{t-1}, \text{ and } \widehat{B}_{d+r+1} = 0,$$

4. Define an auxiliary variable $\widehat{y}_t = \widehat{\gamma}_t + \widehat{D}_t^{-1}\widehat{b}_t$, where $\widehat{\gamma}_t = \widehat{\alpha}_t - \widehat{D}_t^{-1}\widehat{B}_{t+1}\widehat{\alpha}_{t+1}$, for $t = r+1, \ldots, r+d$, and $\widehat{\alpha}_{r+d+1} = \alpha_{r+d+1}$,

5. Consider the linear Gaussian state space model formulated by:
$$\begin{aligned}
\widehat{y}_t &= Z_t\alpha_t + G_t\zeta_t, \qquad t = r+1, \ldots, r+d, &\text{(A.19)}\\
\alpha_{t+1} &= \phi\alpha_t + H_t\zeta_t, \qquad t = r, \ldots, r+d, &\text{(A.20)}\\
\zeta_t &\sim N(0, I_2), &\text{(A.21)}
\end{aligned}$$
where $z_t = 1 + \phi\widehat{D}_t^{-1}\widehat{B}_{t+1}$, $G_t = (\widehat{K}_t^{-1}, \widehat{D}_t^{-1}\widehat{B}_{t+1}\sigma)$, and $H_t = (0, \sigma)$, for $t = r+1, \ldots, r+d$ and $H_0 = (0, \frac{\sigma}{\sqrt{1-\phi^2}})$. Apply the Kalman filter and the disturbance smoother to this state space model, and obtain the posterior mode $\widehat{\underline{\eta}}$ and $\widehat{\underline{\alpha}}$,

6. Go to 2.

In the MCMC sampling procedure, the current sample of $\underline{\eta}$ may be taken as an initial value of the $\widehat{\underline{\eta}}$ in Step 1. To sample $\underline{\eta}$ from the conditional posterior density, Nakajima and Omori (2012) implement the AR (Accept-Reject)-MH algorithm via the simulation smoother using the mode $\widehat{\eta}$ to obtain the approximated linear Gaussian state space model (A.19)- (A.21). See Omori and Watanabe (2008) and Takahashi et al. (2009) for the detail of the AR-MH algorithm.

Bibliography

Aas, K. and Haff, I. H. (2006). The generalized hyperbolic skew student t-distribution, *Journal of financial econometrics* **4**(2): 275–309.

Abanto-Valle, C. A., Bandyopadhyay, D., Lachos, V. H. and Enriquez, I. (2010). Robust bayesian analysis of heavy-tailed stochastic volatility models using scale mixtures of normal distributions, *Computational Statistics & Data Analysis* **54**(12): 2883–2898.

Abanto-Valle, C., Lachos, V. and Dey, D. (2013). Bayesian Estimation of a Skew-Student-t Stochastic Volatility Model, *Methodology and Computing in Applied Probability* pp. 1–18.

Abramowitz, M. and Stegun, I. A. (1972). *Handbook of Mathematical Functions, (10th ed.)*, Dover, New York.

Andersson, J. (2001). On the normal inverse Gaussian stochastic volatility model, *Journal of Business & Economic Statistics* **19**(1): 44–54.

Asai, M. (2008). Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactor volatility models, *Journal of Empirical Finance* **15**(2): 332–341.

Asai, M. (2009). Bayesian analysis of stochastic volatility models with mixture-of-normal distributions, *Mathematics and Computers in Simulation* **79**(8): 2579–2596.

Asai, M. and McAleer, M. (2011). Alternative asymmetric stochastic volatility models, *Econometric Reviews* **30**(5): 548–564.

Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size, *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* **353**(1674): 401–419.

Barndorff-Nielsen, O. E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling, *Scandinavian Journal of statistics* **24**(1): 1–13.

Berg, A., Meyer, R. and Yu, J. (2004). Deviance Information Criterion for comparing stochastic volatility models, *Journal of Business & Economic Statistics* **22**(1): 107–20.

Black, F. (1976). Studies of stock price volatility changes, *Proceedings of the 1976 Meetings of the American Statistical Association, Business and Economics Statistics Section*, pp. 177–181.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity, *Journal of econometrics* **31**(3): 307–327.

Bollerslev, T., Engle, R. F. and Nelson, D. B. (1994). ARCH models, *Handbook of Econometrics* **4**: 2959–3038.

Broto, C. and Ruiz, E. (2004). Estimation methods for stochastic volatility models: a survey, *Journal of Economic Surveys* **18**(5): 613–649.

Cappuccio, N., Lubian, D. and Raggi, D. (2004). MCMC Bayesian estimation of a skew-GED stochastic volatility model, *Studies in Nonlinear Dynamics & Econometrics* **8**(2).

Cappuccio, N., Lubian, D. and Raggi, D. (2006). Investigating asymmetry in U.S. stock market indexes: evidence from a stochastic volatility model, *Applied Financial Economics* **16**(6): 479–490.

Chib, S. (1995). Marginal likelihood from the gibbs output, *Journal of the American Statistical Association* **90**(432): 1313–1321.

Chib, S. (2001). Markov chain monte carlo methods: computation and inference, *Handbook of Econometrics* **5**: 3569–3649.

Chib, S. and Jeliazkov, I. (2001). Marginal likelihood from the metropolis–hastings output, *Journal of the American Statistical Association* **96**(453): 270–281.

Chib, S., Nardari, F. and Shephard, N. (2002). Markov chain monte carlo methods for stochastic volatility models, *Journal of Econometrics* **108**(2): 281–316.

Choy, S. T. B., Wan, W. Y. and Chan, C. M. (2008). Bayesian Student-t stochastic volatility models via scale mixtures, *Advances in Econometrics, Special issue on Bayesian Econometrics Methods* **23**: 595–618.

Delatola, E.-I. and Griffin, J. E. (2013). A bayesian semiparametric model for volatility with a leverage effect, *Computational Statistics & Data Analysis* **60**: 97–110.

Deschamps, P. J. (2012). Bayesian estimation of generalized hyperbolic skewed student garch models, *Computational Statistics & Data Analysis* **56**(11): 3035–3054.

Eberlein, E. and Hammerstein, E. A. v. (2004). Generalized hyperbolic and inverse gaussian distributions: limiting cases and approximation of processes, *Seminar on Stochastic Analysis, Random Fields and Applications IV*, Springer, pp. 221–264.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica* **50**(4): 987–1007.

Engle, R. F. (1995). *ARCH: Selected Readings*, Oxford University Press.

Ghysels, E., Harvey, A. and Renault, E. (1996). Stochastic Volatility, *in* G. Maddala and R. C. (eds), *Statistical Methods in Finance*, Handbook of Statistics 14, Elsevier Science, Amsterdam.

Harvey, A., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models, *The Review of Economic Studies* **61**(2): 247–264.

Jacquier, E., Polson, N. G. and Rossi, P. E. (1994). Bayesian analysis of stochastic volatility models, *Journal of Business & Economic Statistics* **12**(4): 371–89.

Jacquier, E., Polson, N. G. and Rossi, P. E. (2004). Bayesian analysis of stochastic volatility models with fat-tails and correlated errors, *Journal of Econometrics* **122**(1): 185–212.

Jones, M. C. and Faddy, M. J. (2003). A skew extension of the t-distribution, with applications, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **65**(1): 159–174.

Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models, *The Review of Economic Studies* **65**(3): 361–393.

Liesenfeld, R. and Jung, R. C. (2000). Stochastic volatility models: conditional normality versus heavy-tailed distributions, *Journal of Applied Econometrics* **15**(2): 137–160.

Meyer, R. and Yu, J. (2000). BUGS for a Bayesian analysis of stochastic volatility models, *The Econometrics Journal* **3**(2): 198–215.

Nakajima, J. (2012). Bayesian analysis of generalized autoregressive conditional heteroskedasticity and stochastic volatility: Modeling leverage, jumps and heavy-tails for financial time series, *Japanese Economic Review* **63**(1): 81–103.

Nakajima, J. and Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models, *Computational Statistics & Data Analysis* **53**(6): 2335–2353.

Nakajima, J. and Omori, Y. (2012). Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GH skew Student's t-distribution, *Computational Statistics & Data Analysis* **56**(11): 3690–3704.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica* **59**(2): 347–70.

Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference, *Journal of Econometrics* **140**(2): 425–449.

Omori, Y. and Watanabe, T. (2008). Block sampler and posterior mode estimation for asymmetric stochastic volatility models, *Computational Statistics & Data Analysis* **52**(6): 2892–2910.

Pitt, M. K. and Shephard, N. (1999). Filtering via simulation: Auxiliary particle filters, *Journal of the American statistical association* **94**(446): 590–599.

Prause, K. (1999). *The Generalized Hyperbolic Model: Estimation, Financial Derivatives, and Risk Measures*, PhD thesis, University of Freiburg.

Scott, D. (2014). Package "SkewHyperbolic". Manual for the R package SkewHyperbolic. Available online at: http://cran.r-project.org.

Shephard, N. (1996). Statistical aspects of arch and stochastic volatility, *Time series models*, Springer US, pp. 1–67.

Shephard, N. (2005). *Stochastic Volatility: Selected Readings*, Oxford University Press.

Shephard, N. and Pitt, M. K. (1997). Likelihood analysis of non-gaussian measurement time series, *Biometrika* **84**(3): 653–667.

Silva, R. S., Lopes, H. F. and Migon, H. S. (2006). The extended generalized inverse Gaussian distribution for log-linear and stochastic volatility models, *Brazilian Journal of Probability and Statistics* **20**: 67–91.

Takahashi, M., Omori, Y. and Watanabe, T. (2009). Estimating stochastic volatility models using daily returns and realized volatility simultaneously, *Computational Statistics & Data Analysis* **53**(6): 2404–2426.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes - A study of the daily sugar prices 1961-75, *Time Series Analysis: Theory and Practice* **1**: 203–226.

Taylor, S. J. (1986). *Modelling Financial Time Series*, John Wiley, Chichester (UK).

Taylor, S. J. (2005). *Asset Price Dynamics, Volatility and Prediction*, Princeton University Press.

Taylor, S. J. (2008). *Modelling Financial Time Series (Second Edition)*, World Scientific Publishing.

Trojan, S. (2013). Regime Switching Stochastic Volatility with Skew, Fat Tails and Leverage using Returns and Realized Volatility Contemporaneously, *Economics Working Paper Series 1341*, University of St. Gallen, School of Economics and Political Science.

Tsiotas, G. (2012). On generalised asymmetric stochastic volatility models, *Computational Statistics & Data Analysis* **56**(1): 151–172.

Wang, J. J., Chan, J. S. and Choy, S. B. (2011). Stochastic volatility models with leverage and heavy-tailed distributions: A Bayesian approach using scale mixtures, *Computational Statistics & Data Analysis* **55**(1): 852–862.

Wang, J. J. J. (2012). On asymmetric generalised t stochastic volatility models, *Mathematics and Computers in Simulation* **82**(11): 2079–2095.

Wang, J. J. J., Choy, S. T. B. and Chan, J. S. K. (2013). Modelling stochastic volatility using generalized t distribution, *Journal of Statistical Computation and Simulation* pp. 340–354.

Watanabe, T. and Omori, Y. (2004). A multi-move sampler for estimating non-gaussian time series models: Comments on shephard & pitt (1997), *Biometrika* **91**: 246–248.

Yu, J. (2005). On leverage in a stochastic volatility model, *Journal of Econometrics* **127**(2): 165–178.